

Well-Structured Biology

Numerical Taxonomy's Epistemic Vision for Systematics

BECKETT STERNER

To appear in: Patterns of Nature, edited by Andrew Hamilton. University of California Press, 2014.

The history of twentieth-century systematics is full of periodic calls for revolution and the battles for dominance that followed. At the heart of these disputes has been a disagreement over the place of evolutionary theory in the field: some systematists insist that it is central to their methodology, while others argue that evolution can only be studied from an independent foundation (Hull 1988; Vernon 2001; Felsenstein 2001, 2004). Despite the importance of these battles, another trend across the twentieth century is increasingly relevant as it assumes center stage in current methodology: the value and costs of integrating mathematics and computers into the daily research practices of biology (Hagen 2001, 2003; Vernon 1993; Agar 2006; Hine 2008; Hamilton and Wheeler 2008; Strasser 2010). These two features of the history of systematics are related, and we can easily observe today how mathematics forms the common language within which systematists pose their fundamental disagreements. Indeed, the fates of revolution and of mathematics in twentieth-century systematics were inseparable. Moreover, their interdependence exemplifies a general relation between scientific change and the symbolic formalization of a body of practices. I will argue for these two theses from the historical perspective of the founding and articulation of numerical taxonomy (NT) roughly between 1950 and 1970, a period that marks the first major attempt to revolutionize the whole of systematics based on the formalization of the key concepts and techniques in the field.

The central issue will prove to be understanding the character of a particularly *mathematical* revolution in a biological science; that is, what does it look like when a group of scientists set out to reenvision an entire field of biology in symbolic and formal terms? It is important to note at the outset, however, that the integration of mathematics into biology has often proceeded in piecemeal fashion without such a holistic vision. Indeed, systematists began to incorporate the ideas of population sampling and other statistical methods, such t-tests for statistical significance, into their research long before NT became prominent in the 1960s (Hagen 2003). In addition, Ronald Fisher's work at Rothamsted Station beginning in the 1920s opened the door for thousands of biologists to use statistical models as a valuable way to increase the precision and complexity of experiments without

undue worry about needing theoretical explanations for the experimental phenomena. Also, the advent of quantitative traits such as serotypes in midtwentieth-century taxonomy encouraged a certain mathematical facility without requiring a radical reconception of what a character is in general. Usually these piecemeal adoptions of mathematical techniques take an instrumentalist perspective, such as in much of contemporary bioinformatics: math is a tool for getting things done rather than a potential reflection of any deeper truth about biology.

Nonetheless, we can clearly see historical moments when the adoption of symbolic and formal techniques had consequences rippling out to encompass the entire field's practices. The importance of statistical thinking, for example, is unavoidable in the development of the Modern Synthesis. In fact, NT was in a sense an unintended offspring of the groundbreaking statistical work of Sewall Wright and Fisher, both of whom proved an important influence on Robert Sokal, cofounder of NT with Peter Sneath. As we will see, whereas the Modern Synthesis's mathematics failed to penetrate into systematics, Sokal and Sneath reenvisioned the field according to a nonevolutionary yet still statistical foundation. By demonstrating that mathematics had no easy favorite in the systematists' disputes over the proper place of evolution in their field, Sokal and Sneath also opened the door to future arguments over how to reconstruct evolutionary history in terms of mathematical demonstrations and formal philosophical arguments (Farris 1969).

Revolutions are not much if they are just philosophy, however, and the character of a mathematical revolution must center on the daily practices of systematists. How can we uncover the inseparable relation between the mathematical view of biology held by Sokal and Sneath and the revolutionary character of their work for systematics? Although many perspectives are possible on the historical and philosophical significance of NT, I will borrow from Jon Agar the approach of examining the history as an overarching reorganization of scientific work. While my immediate purpose is somewhat orthogonal to his, we are both working within the same field of problems.

Agar touches on the history of systematics insofar as he is pursuing a more general topic, "What difference did computers make?" (Agar 2006). Aiming to deflate or at least analyze the hyperbolic claims often made for the transformative effect of computers on science, Agar investigates their impact on a number of concrete historical cases. He argues that there are three modes in which computers can affect an ongoing scientific research project: "the computer program could closely follow the existing organization of computation (and associated work)"; computerization "could be used as an opportunity for reorganizing methods of computation (and associated work), even though it was still feasible to use manual

or mechanical means”; or last, it “could be used to reorganize methods of computation with new methods that, in principle or in practice, would be impossible to accomplish by manual or mechanical means” (Agar 2006, 873).

He finds that none of his cases fit this last mode, and he speculates that it has indeed never happened. Numerical taxonomy, somewhat implicitly, he places in the second category of using computerization as an opportunity to reorganize the work process without a discontinuous leap from old practices. Keith Vernon offers an account of NT that corroborates this classification insofar as both Sokal and Sneath arrived at their mathematical innovations independent of the use of computers, although they quickly recognized them as a valuable resource (Vernon 1988). Agar limits the scope of his case studies, however, to computerization within a single laboratory or institution, and it is important that what was continuous development in Sokal and Sneath’s own research appeared as a discontinuous leap to the perspective of almost any other systematist at the time. What was relatively continuous growth for a small group of innovators appeared as a universal, radical new take on the field when Sokal and Sneath went public with their research program.

Numerical taxonomy as a reorganization of work therefore has its significance as a historical process in which a small group of scientists universalized their originally quasi-private methods into a research program for every systematist. I will not be able to pursue Agar’s interest in the impact of computers on this process further, though. Instead, I will focus on how the symbolic formalisms of mathematics were central to Sokal and Sneath’s universalization of their local practices into the larger community. That is, I will be concerned with how their effort to arrive at a mathematical method of classification that was universal to every type of organism intrinsically implicated them in a qualitative reorganization of the work of all systematists. This chapter therefore expands Agar’s topic to consider how local reorganizations of work can expand beyond individual labs or institutions to communities as wholes. This sort of process—in which a provincial prototype of a research strategy is generalized into a (putatively) universal program—is a common one that happens on many different scales. For convenience, I want to label the thinking that guides this process an “epistemic vision,” briefly defined as a feasible strategy for solving ill-structured problems by the principled reorganization of work. I will have a bit more to say about the nature of epistemic visions below, but the major part of this chapter will focus on elucidating the particular vision that Sokal and Sneath had rather than epistemic visions in general.

An excellent place to examine how Sokal and Sneath universalized their work into a general research program is their manifesto and the bible of NT, *Principles of Numerical Taxonomy*, published in 1963 (Sokal 1963). The publication of *Principles* in 1963 is a critical intersection point of the integration of mathematics

into systematics and the contentious ideological debates of the field. Although Sokal and Sneath, along with a number of interested colleagues, had already published similar arguments in a couple journal articles, the scope, provocativeness, and coherence of the book made it a flashpoint of debate as well as a critical reference source. In the book, Sokal and Sneath set out a new theory and method of classifying organisms based solely on a concept of similarity that, at least initially, excluded the use of evolutionary theory. In order to carry this out, they reenvisioned the character traits of organisms, which form the empirical observations upon which classifications are based, as quantitative variables sampled from a larger population of traits ultimately contained in the genome. In NT, characters gained essential statistical properties that guaranteed the convergence of a classification as the number of measurements increases. In this way, a required reorganization of the order of research—classify using similarity alone before proceeding to evolutionary considerations—is inseparably linked to new statistical methods.

My particular conclusions about NT's contributions to twentieth-century systematics focus on how certain aspects of the new organization of work Sokal and Sneath proposed became entrenched even as opposing schools of thought criticized NT's larger ambitions. The way that morphological characters are treated even today owes something to how NT first reorganized the process of classification to treat characters as inherently statistical in nature. This heritage depends specifically on how Sokal and Sneath distinguished two self-sustaining and autonomous lines of work within systematics: methodology, that is, the study of numerical taxonomy in general, and taxon-specific research, the study of particular groups of organisms, within which NT was supposedly the universal method of classification. The character of how they distinguished methodology from taxon-specific research remains influential today, even if the content of the method and of taxon research have changed dramatically. In other words, I believe one of NT's most important legacies for the field has been the institutional entrenchment of a distinction between types of research that hinges on particular assumptions and techniques in the process of classification.

ARTICULATING A GENERAL PROCESS OF CLASSIFICATION

The problem of a universal method for systematics, if there even was just *one*, had engaged researchers for some time before Sokal and Sneath entered the scene.¹ Systematics, such as it was before 1960, was a fractious field where entomologists might feel little communion with other zoologists, let alone botanists or microbiologists (Johnson 2009). Few ready resources were available to support systematic inquiry into the common techniques used across communities studying different taxa, and advancing a substantive perspective on the field as a whole was a considerable achievement. The efforts of Ernst Mayr, George Gaylord Simpson,

and Willi Hennig were influential for Sokal and Sneath's 1963 book, *Principles*, both as scaffolding for their own ideas and as critical targets. The work of Mayr and Simpson on describing a general step-wise process for classification, in particular, is important background to understanding how Sokal and Sneath sought to reorganize classificatory work.

A leading light of the New Systematics, Ernst Mayr published in 1942 his influential contribution to the Modern Synthesis, *Systematics and the Origin of Species*. The book aimed to fix a new purpose for classifications beyond the practical needs of description and identification: to aid in the study of speciation and evolution. Mayr's stated goal was for classificatory work to be a step along the way to analyzing patterns of variation and speciation, but this aim required resources and huge collection sizes far beyond the means of most taxonomists. Mayr also included a section titled "The Procedure of the Systematist" (Mayr 1942, 11), where he notes the paucity of introductions to "the practical methods of taxonomic work" and recommends that a beginner should find a "lucid" treatise that "set an example of method" and would serve to begin the education-by-apprenticeship model characteristic of taxonomy at the time. Although Mayr does go on to describe several stages in the procedures of systematics, including collection, identification, and description, his exposition quickly leaves practical method behind for a conceptual discussion of types, nomenclature, and rank. Nonetheless, his effort to organize the work of systematics into a series of consecutive stages was one of several important attempts to which Sokal and Sneath would respond in *Principles*.

Publishing *Principles of Animal Taxonomy* almost two decades later, in 1961, G.G. Simpson went beyond Mayr in articulating the actual practice of classification, although he also embraced its open-ended nature. In the section "What Do Taxonomists Do?" Simpson writes that "in general, he does the following things, more or less in this sequence but usually with a good deal of overlapping, jumping, and backtracking." I summarize the steps here: the taxonomist (1) selects the organisms, (2) records observations on them, (3) sorts them into different units (local populations, species, etc.), (4) compares the units' characteristics, (5) interprets their relationships, (6) makes inferences about an evolutionary pattern, (7) translates these conclusions into hierarchic rankings, and finally (8) selects appropriate names for the units (Simpson 1961, 108–109). Simpson was a proponent of the artistic side of taxonomy, and while he adds new detail to the description of classificatory work, he also leaves considerable creative freedom to the taxonomist regarding how to achieve the final product (Hagen 2003).

A final example of methodological description will help flesh out the background to NT's criticisms. In Willi Hennig's 1950 book, translated into English much later, in 1966, he discusses "reciprocal illumination" as the method of systematics.

Since both a whole and its parts have specific properties and coherences, one can investigate the parts and then learn something about the whole by recalibrating one's results to its nature, and vice versa. As Hennig describes it, "The phylogenetic affiliations are, so to speak, the whole. With its parts—the relations of morphological, ecological, physiological, geographic, etc. similarity, which reflect the phylogenetic affiliations of their bearers—[the whole] is subjected to the method of checking, correcting, and rechecking" (Hennig 1966, 23). Thus one can use geographic distribution as a way of checking the reliability of phylogenetic relationships (Hennig 1966, 148). Hennig quotes the ethnographer Mühlmann to help refute a common objection against reciprocal illumination: "The objection that this is circular reasoning has been discussed by Dilthey; it applies logically, but not to practical investigation" (Hennig 1966, 22). To paraphrase, while Hennig believed that a logical translation of his scientific research might appear to be deriving its premises from its conclusions, this circularity does not apply in the face of the empirical uncertainty of practice. That is, logical circularity in the world of precise deductive methods does not necessarily imply a poor research strategy in the world of messy, ambiguous ideas.

Learning from exemplars, applying artistic expertise, and reinterpreting data on the go are characteristic modes of work based on practical expert knowledge about particular organisms. Each of these modes is well fitted to coping with the open-ended diversity of concrete specifics about which little is known beforehand. Consequently, from this perspective there is no single procedure for classifying that can be simultaneously perfectly fitted for arthropods and fungi because the specifics of each are at least as important as their commonalities. Moreover, any two entomologists, for example, could make classifications using different procedures and principles. As a result, Mayr's goal of studying variation and speciation helped harmonize the work of some systematists but still left a lot of leeway (Johnson 2005, 2009). In addition, although Simpson's list of steps is helpful, almost all the practical knowledge of the expert is contained in the "overlapping, jumping, and backtracking" of the actual process. As long as the structure of systematic method was so open-ended, it was difficult even for leading researchers such as Mayr or Simpson to provide much helpful guidance.

Indeed, a variety of research groups across the world struggled to make do under the umbrella of the New Systematics. When there was no evolutionary evidence about the history of traits, the evidence was ambiguous, or a simple classification was desired independent of evolutionary hypotheses, no exemplars of practice led the way for researchers. As Vernon has documented, the shortcomings of the New Systematics appear to have produced a remarkable, independent agreement among critics about what was best for classificatory work to proceed outside the Modern Synthesis. As a result, Sokal and Sneath's research program was not the only one to advance a controversial emphasis on segregating evolutionary theory from the

classification process: similar ideas sprang up in numerous locations around the world in the time between 1950 and 1963, when *Principles of Numerical Taxonomy* was first published. I will use the general label “phenetic” methods to describe their resulting proposed solutions.²

DETERMINING THE METHODOLOGICAL ORDER OF TAXONOMY

While the Modern Synthesis in general would be hamstrung without statistical concepts, it was not the research program’s statistics with which the pheneticists took issue; rather, the particular problem of phenetics was how to solve a class of problems without making use of evolutionary theories. Statistics would in fact prove critical to the phenetic approach in general, insofar as it provided new, more formalized exemplars for how to infer classifications. The point of phenetics, putting inferences about similarity before historical genealogy, went hand-in-hand with a new perspective on what material went into the start of classificatory work and how it could be used. Phenetics, therefore, represents a first step for NT as a reorganization of work; it was both a general response to intellectual conditions at the time and the ground upon which Sokal and Sneath developed their particular approach.³

For our purposes, what is most important about the phenetic approach is how it articulates an alternative problem situation that encompasses the same general process as Simpson’s eight steps but differs on the start, end, means, and goal of the process. In order to describe this new structuring of classification, it will help to bring in Herbert Simon’s discussion of well versus ill-structured problems (Simon 1973). Simon and others developed this distinction (independently of systematics) in order to study the tractability of various challenges for artificial intelligence, and the general scope of their description of problem structures offers a useful and appropriate perspective for our situation here. I’ll proceed in this section by first summarizing the relevant history of phenetics’ independent origins and then examining them in more depth in relation to the new structure placed on classification.

In the 1950s, working outside an evolutionary perspective was a thoroughly “ill-structured problem.” There were few answers to important questions such as, How should a researcher select and characterize the data that went into the classification? What techniques should he use to go from initial observations to groupings of taxa? How could he know when the classification was finished? Without a general, uniform structure to the problem of nonevolutionary classification, it remained baffling in complexity, offering too many choices with evaluative criteria that were too weak to determine a clear answer. The innovations of phenetics would contribute several key insights, including that all characters that were not logically redundant were acceptable, and they should each be weighted equally in judging similarity.

Phenetic methods developed independently across a set of three initial groups comprising five researchers in all: A. J. Cain and G. A. Harrison, Sokal and C. D. Michener, and Sneath.⁴ These three groups published the seminal papers of what would become NT in a thirteen-month period, between June 1957 and July 1958. The reasons for this striking simultaneity of independent innovations deserve further research, but Vernon finds that the innovations were produced by reactions of outsiders to the general state of taxonomy at the time. The interests of each individual help us to understand the scope of and motivations behind what became a radical change in systematic practice, so it is worth summarizing their motivations and the problems they faced.

Of Cain and Harrison, Harrison was the outsider to taxonomy and Cain an expert who dissented from aspects of the Modern Synthesis. Cain had already spent considerable time studying the geographic variation of the land snail *Cepaea nemoralis* with the geneticist P. Sheppard, and this research served as a basis for his arguments against the Modern Synthesis's emphasis on drift as opposed to selection. Also, the research highlighted the problems with using phylogenetic methods in data-poor case studies as there were few fossils available for study. Cain leveraged his difficulties into general arguments against evolutionary classifications that used poor data and unanalyzed notions of taxonomic importance.

When Harrison, a budding biological anthropologist, took a course from Cain at Oxford, his interest in craniometry sparked a collaboration that led to their innovative theories about taxonomy. Harrison was interested in the Mount Carmel skulls but struggled with how anthropologists lacked a justifiable way to identify clear evolutionarily significant and measurable characters. Over the course of several years, Cain and Harrison arrived at the idea that any characters not correlated with each other by necessity were acceptable, and that weighting each character equally would allow the calculation of affinities without evolutionary input. Thus they separated out the judgment of affinity, "phenetics," from the judgment of evolutionary history, phylogenetics.

Sneath started off even further from evolutionary biology than Harrison: he trained in medicine and was working in bacteriology when he first struggled with the problems of taxonomy. He was preparing an M.D. thesis in London on clinical cases of infection by *Chromobacterium* when he began to wonder if his classification of thirty-eight strains into two broad groups meant the groups were different species. He quickly discovered the chaotic state of bacterial taxonomy at the time, especially the seemingly arbitrary emphasis on some characters above others and unjustified attempts at phylogenetic classifications. In consultation with the botanist J. S. L. Gilmour, Sneath worked to articulate a nonphylogenetic method for classifying bacteria that had a more logical way to assess quality and

method. As with Harrison and Cain, the failings of taxonomy's tool kit, including the phylogenetic approach, inspired Sneath to attempt a reorganization of classification to make it more coherent, justifiable, and self-complete with respect to evolution.

In contrast, Michener and Sokal were motivated by a different dynamic opposing the objectivity of mathematical methods to the subjectivity of evolutionary methods. Sokal arrived in the Department of Entomology at the University of Kansas after a fellow graduate student at the University of Chicago had converted him to applying mathematical methods in biology. Sokal had taken a biometry course from Sewall Wright at Chicago and over time learned or taught himself cutting-edge methods in statistics. At Kansas, however, he was a colleague of five other entomologists, all of whom were traditional taxonomists with little appreciation for mathematics. Sokal in fact had no initial interest in taxonomy—his dissertation was on ecological variation—but he was spurred by the tenor of the department to make a bet one day at lunch that he could do a better job of classifying organisms with mathematics than the traditional approach.

Michener, the department heavyweight, provided Sokal with data on 122 characters of 98 bee species and then helped him evaluate the results of the new mathematical procedure Sokal created. Michener was an exemplar of the New Systematics, and he produced authoritative work on bees that relied on the traditional “by eye” approach to classification. While Michener knew little mathematics himself, it made a certain intuitive sense to him, and he was open to Sokal's innovations where many systematists were not. In fact, when Sokal's results turned out to make a slight improvement on Michener's prior results, he embraced them and coauthored a lengthy paper with Sokal. What neither of them expected, however, was that Sokal's approach did not rely on evolution; as Vernon relates, it took a couple years before they realized the study was purely one of affinity, not of history. In parallel to Cain and Harrison, Michener and Sokal therefore concluded that taxonomy was in fact based on a judgment of similarity first and evolution second. They ultimately diverged, however, on whether evolutionary considerations should be incorporated into the method, and Sokal preferred keeping the static analysis of similarity pure.

The three groups therefore originated in a common concern of the outsider: to make use of the products of systematics, that is, its classifications or phylogenies, without having to choose among the potentially idiosyncratic commitments and methods of systematics' practitioners. In order to better interpret the consequences of phenetics for the problem situation of nonevolutionary classification, it will help to borrow Simon's work on describing the relation between well and illstructured problems. Simon listed six criteria that go into making a particular problem well versus ill structured, although ultimately the distinction is one of

degree rather than kind. I quote them at length because they indicate very richly what sorts of knowledge, techniques, and criteria are necessary to define a process of calculation that does not involve Simpson's "overlapping, jumping, and backtracking."⁵ Our immediate interest in this section is point 4, although 1, 2, and 3 are also relevant for considering NT's distinctive merits in later sections.

1. There is a definite criterion for testing any proposed solution, and a mechanizable process for applying the criterion.
2. There is at least one problem space in which can be represented the initial problem state, the goal state, and all other states that may be reached, *or considered*, in the course of attempting a solution of the problem.
3. Attainable state changes (legal moves) can be represented in a problem space, as transitions from given states to the states directly attainable from them. But considerable moves, whether legal or not, can also be represented—that is, all transitions from one considerable state to another.
4. Any knowledge that the problem solver can acquire about the problem can be represented in one or more problem spaces.
5. If the actual problem involves acting upon the external world, then the definition of state changes and of the effects upon the state of applying any operator reflect with complete accuracy in one or more problem spaces the laws (laws of nature) that govern the external world.
6. All of these conditions hold in the strong sense that the basic processes postulated require only practicable amounts of computation, and the information postulated is effectively available to the processes—i.e., available with the help of only practicable amounts of search. (Simon 1973, 183)

The common situation facing all three groups was that of desiring a classification whose *method of production* was intelligible even to outsiders. Harrison, Sneath, and Sokal each in his own way found the guildlike mastery of systematists obstructive or questionable. One of the greatest merits of setting universal rules about the admissibility and weighting of characters—that is, any logically nonredundant character is acceptable and all are weighted equally—was to explicitly delimit what prior knowledge could be used to make a classification. Under these strictures, the set of characters available for a classification therefore fulfilled the dictum, per item 4, that "any knowledge that the problem solver can acquire about the problem can be represented in one or more problem spaces." Methodological rules therefore ensured the sufficiency of the pheneticists' abstracted representation of what initial knowledge went into their classifications.

Where Simpson sought a general description of systematists' methodology but explicitly recognized its ultimately ill-structured nature, pheneticists pursued a

similar program but in the critical vein, emphasizing a need for new rigor. My interest in this chapter is how that rigor manifested itself in the reorganization of classificatory work. Each of the different groups, beginning from independent experiences of frustration with contemporary systematics, arrived at analogous resolutions that allowed them to proceed in solving what they had initially set as their problem.⁶ Broadly speaking, these resolutions to otherwise narrow problems took on greater significance for all of systematics because they were posed as a critique and positive alternative to the “whole” of the field already envisioned by Mayr, Simpson, and many others in the New Systematics. Only through the test of time had it become clear that the practical gaps in the methodological prescriptions of the New Systematics were systematic in nature, in that the gaps included a general class of problems rather than merely a hodgepodge of disparate cases.⁷ Moreover, these problems proved sufficiently fundamental in nature that their local solutions supported further research by pheneticists to generate a new view of the field.

In the introduction I offered a name for the sort of thinking that guides this fairly common process of moving from local to general problems. I termed them “epistemic visions,” defined as a feasible strategy for solving ill-structured problems by the principled reorganization of work. Now that we have seen a concrete instance of how putting judgments of similarity before evolutionary history helped make classification a better-structured problem for the early pheneticists, we can say more about epistemic visions in general, although a full discussion would take too much space. The local problems the early pheneticists faced in developing their thoughts were relevant to the whole field in part because their challenge required finding a new *way* to make what is a nearly universal product of every systematics research project: some sort of branching diagram that organized the species (or other units) into mutual relationships. The key concepts of the field are intrinsic to carrying out this work, in that it is only through ideas such as character, species, similarity, and homology that the scientists can *know* what they are *doing*.⁸ Hence articulating a different way of classifying intrinsically implicates a shift in concepts because it is through these concepts that the method becomes generally intelligible. In our present context, we can uncover these linked changes in concepts and practices through careful attention to the particular epistemic obstacles systematists faced, the universal assumptions they made about their objects of study to justify certain procedures, and also how they represented their work process as a whole (more on this last item later). Simon’s characterization of welland ill-structured problems therefore offers a heuristic frame for us to track what about a work process is being reorganized.

In general, an epistemic vision refers to a plan of action, a strategy, that a person has thought through as a way of making some previously insoluble class of problems tractable. Often, as was the case for the early pheneticists, one will

stumble onto a solution for a local version of the general problem first, often not even appreciating its true scope, and only then, upon reflection, begin to ask how the local situation can be understood as a more general one.⁹ Through this process of reflection, the scientists revise their conceptual understanding of their work in such a way as to end up claiming that something similar to the local solution would occur more broadly. Their reflection thus essentially involves them in *envisioning how* to demonstrate in action the truth of this new conceptual understanding. For Sokal and Sneath, their vision of a new sort of taxonomy engaged them in a multipronged research program whereby they aimed to demonstrate the correctness of their approach using theoretical arguments, concrete comparisons against older methods, simulations, and mathematical proofs.

There are a host of important questions that are raised by the place of epistemic visions in scientific change, but I will focus here only on its relation to the significance of Sokal and Sneath's reorganization of classificatory work. Beyond the general phenetic move to classifications based on similarity using any nonredundant character with equal weight, Sokal and Sneath articulated a particular approach with its own unique consequences: they added enough further structure to the problem to separate out two different types of work that were tied together in a functional relationship. Numerical taxonomy supported *methodology* as a full-time research project, in addition to the traditional production of concrete classifications of various organisms. In the rest of this chapter I articulate the further structure Sokal and Sneath added to the process of classification and describe what parts of it have become entrenched, forming NT's major historical contribution to methodology in systematics.

BRACKETING METHOD OFF FROM TAXON-SPECIFIC RESEARCH

As contrasted to the descriptions of method offered within the New Systematics that we saw above, the methodological innovations of NT were twofold. The first was rigorous enforcement of classification as a strictly linear process, born out of the general phenetic move to put classifying *before* inferring phylogeny and instantiated in practice almost exclusively by computer programs. Second, Sokal and Sneath devoted additional attention to specific steps in the overall process: the three stages in the middle of the process were where they made their statistical and methodological contribution, as is reflected by the technical sophistication and length of treatment received by those stages in *Principles of Numerical Taxonomy*. It was in the second innovation that NT made its distinctive mark on the history of systematics: the particular way it articulated and justified a methodology universal to all taxa, drawing on conceptual resources from statistics, logical empiricism, and genetic determinism. While many of the historical specifics of Sokal and Sneath's vision have dropped away, the boundaries and relations of methodology to taxon-specific work have remained largely unchanged, but this organization is

neither necessary nor unproblematic for the field. Indeed, some systematists appear to be struggling with this legacy of NT in just the past few years.

Sokal and Sneath's first innovation in reorganizing the work of taxonomy was to insist strictly that even in practice one must proceed

Table 9.1 Comparison of the procedures of Simpson and Sokal and Sneath

Step	Simpson		Step	Sokal & Sneath
1.	Select specimens	=	1.	Select specimens
2.	Record observations	≈	2.	Discover & measure characters
3.	Sort into units	X		
4.	Compare units	→	3.	Code characters
5.	Interpret relations	→	4.	Calculate affinity
6.	Infer evol. patterns	→	5.	Cluster specimens
7.	Make rankings	≈	6.	Extraction of data
8.	Select names	=	7.	Identify specimens

Note: Sokal and Sneath have transformed the middle of the process, steps 4–6 (denoted by arrows in the middle column), while leaving the extremes relatively untouched. Step 3 was eliminated because Sokal and Sneath treated all clustering problems as independent of which units in the Linnaean hierarchical ranking system are used.

linearly instead of reciprocally. In particular, NT split classification into stages not unlike those of Simpson but required that each stage depend on only the output of the previous stage.¹⁰ Later stages could not directly affect the execution of earlier stages unless the taxonomist decided to start the process anew with different specimens, characters, or algorithms. In addition, their insistence on the equal weighting of characters, at the least the first time through the classification process, precluded much of the backtracking and reciprocal illumination Simpson and Hennig found so valuable.

Figure 9.1 shows the emphatically linear process Sokal and Sneath articulated, which was enforced in part by normative strictures in *Principles* and by the linear, input-output nature of computer programs. If we compare the steps of NT's process to Simpson's, we see that the steps given in figure 9.1 closely mirror the eight given by Simpson but show important differences. See table 9.1 for a summary. Step 3 in Simpson, the forming of specimens into groups, has disappeared for Sokal and Sneath because their general procedure was intended to be the same no matter to what hierarchical level the clustering applied, for example, genera or varieties. They introduced the term *operational taxonomic unit* (OTU; see steps 3–5 in fig. 9.1) in order to side step the loaded issues of species and ranks. Thus the expert knowledge about geographic variation, species concepts, or any determinate ecological relation between the specimens is hidden or diffused into the remaining stages in NT's procedure.

In addition, step 6 in Simpson, the inference of evolutionary pattern, has been transformed into a nonevolutionary analog. The replacement is justified from Sokal and Sneath's point of view because such evolutionary conclusions putatively could be based only on an already complete classification. Thus the inference of evolutionary patterns, especially in a reflective method, invited a vicious circle. Hence instead of stage 6 in Simpson we have stage 5 of Sokal and Sneath, where the

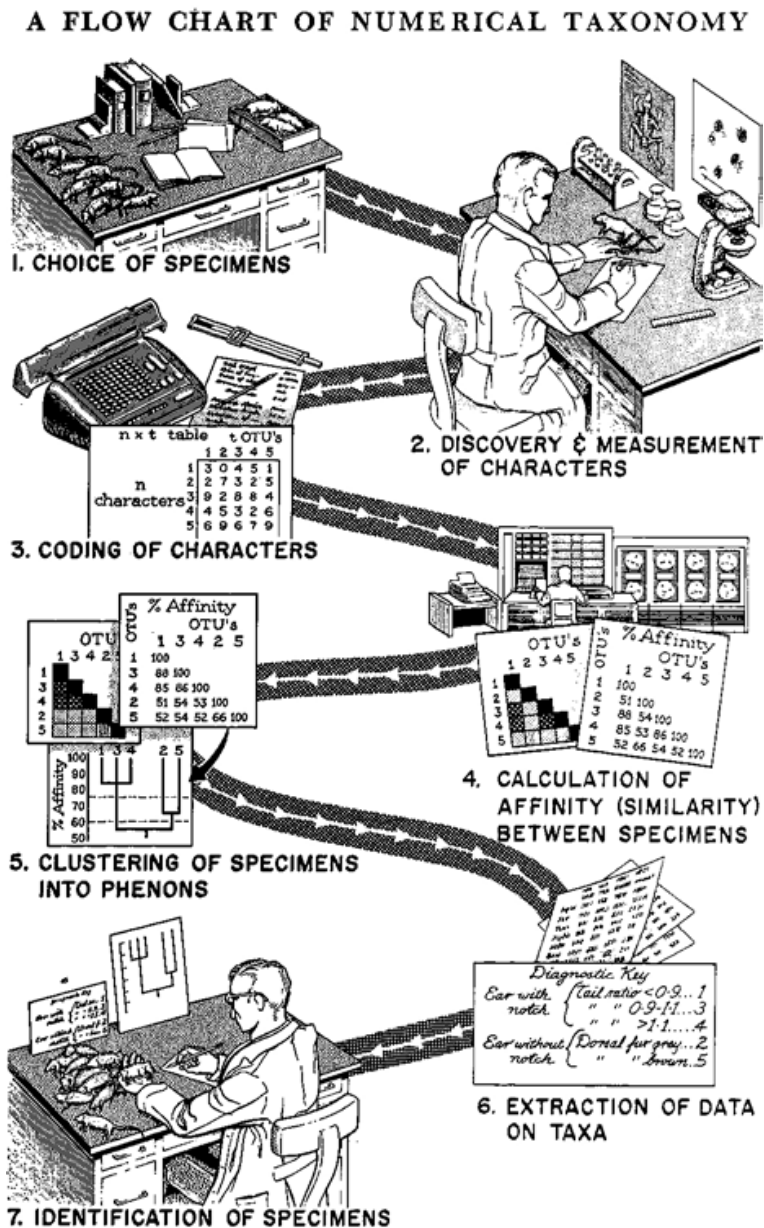


FIGURE 9.1. The seven stages of carrying out a classification process in numerical taxonomy. Sokal and Sneath's innovations compared to previous descriptions of taxonomic procedure lie in stages 3 through 5.

taxonomist infers clusters by affinities, that is, measures of similarity between the

characters of specimens.

The heart of NT's technical innovation as of 1963 was stages 3–5, which reflect in order its statistical approach to taxonomic characters, the articulation of affinity as a mathematical function, and the use of clustering algorithms to produce a branching classification diagram. The corresponding three chapters in *Principles* (chaps. 5–7) transform steps in Simpson's account from a “descriptive generalization” to a technically involved, explicit procedure (Sokal 1963, 21). In contrast, Sokal and Sneath have changed almost nothing about the extreme stages (1 and 8 in Simpson), and their stance toward stages 2 and 7 was concerned only to ensure that the bracketing off of methodology was secured, leaving a wide range of decisions about character choice and ranking to the particular systematist. For example, while Sokal and Sneath did have points to make about the selection of specimens and character measurement, these issues receive a much briefer discussion in *Principles* that lacks extensive mathematical symbolization (e.g., their take on nomenclature, the logical independence of characters, or species as a useful choice for operational taxonomic units). In addition, the technical issues of how to build diagnostic keys for identification embodied by step 6 in figure 9.1 receive little discussion in *Principles* and were treated by other authors outside NT.¹¹

Focusing on this second innovation of NT for now, we will see that the transformation of the *middle* of the process of classification carried a far-reaching consequence for the organization of work. On the one hand, Sokal and Sneath's work opened well-structured research problems about the statistical nature of characters, measuring affinity, and clustering OTUs. These problems could and did demand the full attention of many researchers who made careers out of studying methods of classification, irrespective of any knowledge about particular organisms.¹² On the other hand, methodology was now sharply distinguished from the work of producing actual classifications, which were the end products of the great majority of taxonomists.¹³ These systematists' work specialized on particular organisms, and their results rarely aimed for a universal impact on the work of other taxonomists. I will call the former group methodologists and the latter group taxon specialists.¹⁴ This differentiation of research work demonstrates how clearly Mayr's 1942 suggestion about learning taxonomy through exemplars had been challenged by the claim of numerical taxonomists to a method universal for all taxa.

The intended universality of method in NT is evident in a number of ways. One simple example is how Sokal and Sneath redefine classification:

We have adapted Simpson's definition (1961, p. 9), which restricted itself to zoological classification, to more general usage. . . . We have restricted [this definition] to organisms, since this book is primarily intended for the biological taxonomist. However, there are

many methods of classification, including numerical taxonomy, which are equally applicable to concepts and entities other than organisms. (Sokal and Sneath 1963, 3)

In other words, NT is a general method of classification that extends beyond living organisms and is most certainly not restricted to any group of them. This universality is also reflected in the term *operational taxonomic unit* that Sokal and Sneath used to refer to data taken from taxonomic units at any level of the Linnaean hierarchy (Sokal and Sneath 1963, 120–121). The issue is again evident in the fact that Sokal and Sneath looked for a definition of characters that would “allow the broad treatment needed for a *general* theory of systematics” (Sokal and Sneath 1963, 63; my emphasis). Hence NT is a universal method in the sense that its concepts, principles, and procedures are putatively applicable to all living organisms (and more), and although particular procedures may prove more apt in a given study than others, they are still general in form.

Numerical taxonomy’s putative universality for all kinds of organisms depended in practice on two important objects that bracket off the taxon specialists’ work from the methodologists’. That is, the objects are the transitions where the process moves up into the abstractions of the computational methods and then back down into the concreteness of the organisms, and they are the character matrix and dendrogram, prominently visible in steps 3 and 5 in figure 9.1. The two objects function as boundary object between lines of work of methodologists and taxon specialists (Star and Griesemer 1989). For NT, the character matrix and dendrogram are boundary objects that join together but also hold apart the expertises of the methodologist and the taxon specialist.

In a more concrete sense, the boundary objects help articulate illstructured problems into well-structured ones by making the problems’ structure and solutions mostly independent of each other. (This is analogous to Simon’s [1962] near decomposability concept.) By establishing the character matrix and dendrogram as pivotal objects between lines of work, Sokal and Sneath can claim to legitimately have abstracted away aspects of practice, making the organisms a separate subject of study from the algorithm that analyzes them. In other words, they consciously set out a strategy in *Principles* for how to re-organize research work to separate out what is a good method vs. what is a good classification of the arthropods, for example. Figure 9.2 superimposes this reorganization of work I have described onto figure 9.1.

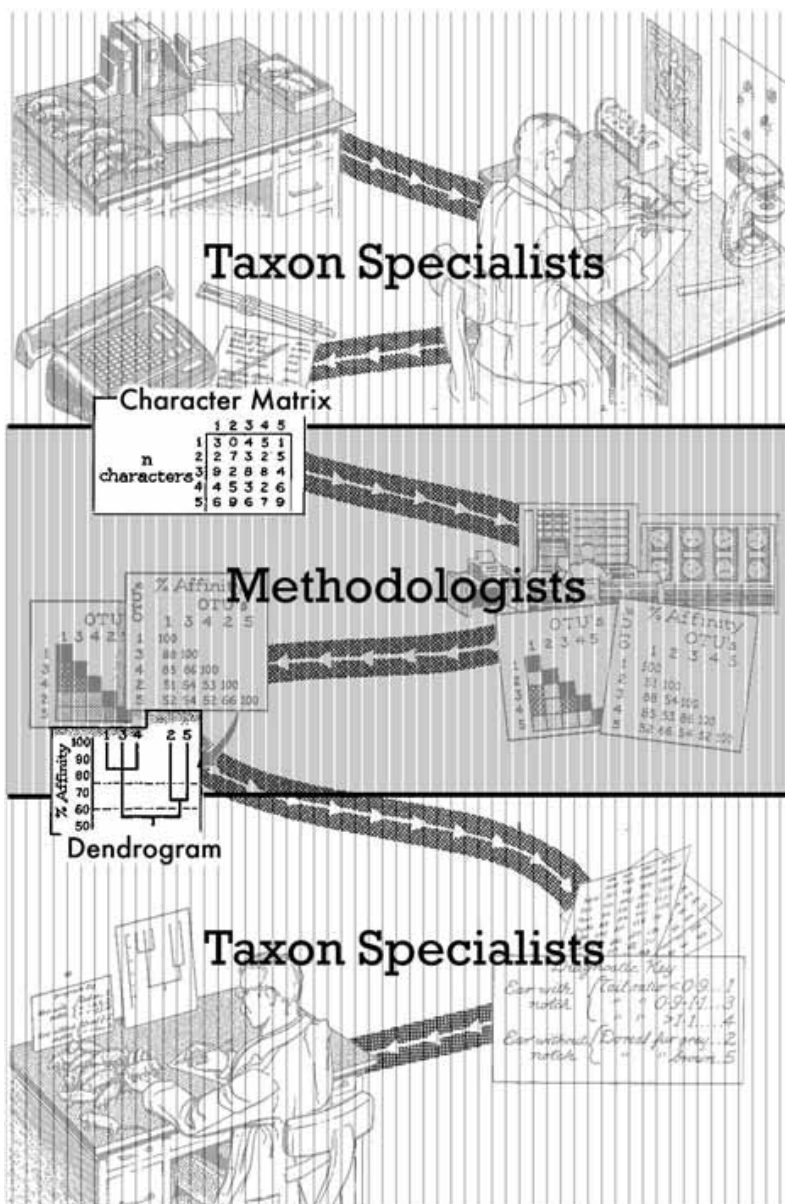


FIGURE 9.2. The workflow of numerical taxonomy edited to emphasize its reorganization of work for methodologists and taxon specialists. Methodologists focus their expertise on the middle stages, while taxon specialists are experts in the first and last stages. These subspecialties are bracketed off from each other by the character matrix and dendrogram, which are highlighted as well.

Sokal and Sneath took seriously the issue of formally defining a character in such a way that its nature and representation does not depend on the particular organism from which it came. In other words, their approach effectively aimed to establish that *in principle* no taxonspecific expertise was needed to classify the taxa once a character matrix had been produced, although they acknowledged “only specialists in the various groups will be in a position to define and describe unit taxonomic characters in the organisms they are studying” (Sokal and Sneath 1963, 92).¹⁵ How, then, did they ground their method in the nature of characters while avoiding the

“speculative” hypotheses of evolutionary theory and expertise of taxon-specific researchers? Sokal and Sneath brought their own distinctive set of conceptual resources from outside systematics, and together they succeeded at reposing the question of how to judge the similarity of two organisms in a well-structured fashion as a problem of statistical estimation.¹⁶

From the perspective of NT, character measurements are simply instances of matches or mismatches in the overall similarity function, and neither the quantitative nor the qualitative nature of what the characters *are* is important to their value in the classification process. Indeed, the work of the taxon specialist is effectively recast as one of developing and improving new ways of measuring the genome through phenotypic traits, at least insofar as the methodology of classification was concerned. The meaning of resemblance in NT is independent from the results and concepts of most of systematics: it is ultimately a function of matches between genomes that Sokal and Sneath have defined to be unaffected by any theories about homology or evolution. The problem of classification therefore becomes an essentially statistical one: devise methods for accurately estimating the parameter indicating the proportion of matching traits and then find ways to best represent those resemblances in a branching classification diagram. This way of posing the problem of classification made it well-structured yet complex enough to involve multiple subproblems that were nonetheless all solvable together using standard procedural algorithms on a computer.

Returning to my characterization of the problem situation facing any nonevolutionary taxonomist, we can now see how Sokal and Sneath sought to resolve aspects 1, 2, and 3 of Simon’s list of the properties of a well-structured problem. Briefly summarized, these are (1) a definite criterion for testing proposed solutions; (2) a problem space with which they could represent the start, end, and middle states of the classification process; and (3) state changes (“legal moves” in the game), which can also be represented in the problem space. Commensurate with their initial status as outsiders to systematics, Sokal and Sneath drew on external resources to reorganize classificatory work, including mathematical statistics, logical positivism, genetic determinism, and information theory.¹⁷ Gilmour’s positivist ideas about the naturalness of classifications would prove important for item 1, along with an adaptation of the correlation coefficient from statistics, and the combination of information theory with genetic determinism was critical for supporting the algebraic coding of characters and the matrix representation of the units of classification in a multidimensional trait space, fulfilling items 2 and 3.

One major source of inspiration came from Sneath’s interactions with Gilmour and Gilmour’s 1940 essay, “Taxonomy and Philosophy,” in *The New Systematics* (Vernon 1988; Gilmour 1940). Gilmour’s intention was to bring the

epistemological ideas of logical positivism to taxonomy “for an examination of the purpose and method” of the field (Gilmour 1940, 463). Under this interpretation, the work of taxonomy “consists in clipping together the mass of sense-data collectively classed as ‘living things’ into a logically coherent pattern for the purpose of making inductive generalizations concerning those data” (Gilmour 1940, 465). He goes on to define a “natural” classification as “that grouping which endeavors to utilize *all* the attributes of the individuals under consideration, and is hence useful for a very wide range of purposes,” in contrast to an “artificial” classification that uses few attributes and fits only a narrow range of purposes (Gilmour 1940, 472).

Sokal and Sneath fleshed out this perspective in greater detail by offering operationalized definitions of character and homology. They initially defined a unit character “as an attribute possessed by an organism about which one statement can be made,” with the understanding that “attributes are formally logical constructs, since they will change if the technique of observation changes” (Sokal and Sneath 1963, 63). The states of characters can be recoded as binary bits according to information theory, and Sokal and Sneath proceed to give a survey of calculations one can make about how many bits of information an organism contains in its genome. However, since gene sequences were rarely observable directly at the time, Sokal and Sneath give a “working definition” of a unit character that uses only phenotypic data: “we may define a unit character as *a taxonomic character of two or more states, which within the study at hand cannot be subdivided logically, except for subdivision brought about by changes in the method of coding*” (Sokal and Sneath 1963, 65; emphasis in original). Since logical divisions depend on the availability of instruments to measure states, a unit character is a set of possible states that no measurement is able to further refine, perhaps because we do not have the capability to observe it in fine enough detail.

The information view of characters therefore sees them as observations ultimately representable in the form of binary yes or no states. Sokal and Sneath believed this view was also consonant with the genetic theory of the time: they claimed that “we may tentatively identify our taxonomic bits with the genetic code.” In addition, “it is generally considered that only genetically determined characters should be used in orthodox taxonomy, and with this we concur” (Sokal and Sneath 1963, 92).¹⁸ (What remains unclear, though, is precisely how they proposed to select only genetically determined characters.) Nonetheless, the focus on genes as the sole (or perhaps primary) cause of taxonomic characters was central to their statistical theory of why clustering based on affinity functions was a valid approach for taxonomy.

Given this causal identification of observable taxonomic characters with genomes, however, there arises a problem of characterizing the worth of nongenetic

characters as indicators of the similarity of different genomes. In response, Sokal and Sneath advanced several basic hypotheses that essentially justified why traditional characters in systematics could be treated as offering an unbiased sample of the genomes. The first two “fundamental hypotheses” (Sokal and Sneath 1963, 84) are the nexus and nonspecificity hypotheses. The nexus hypothesis states that “every taxonomic character is likely to be affected by more than one genetic factor and that, conversely, most genes affect more than one character.” The nonspecificity hypothesis assumes “that there are no distinct large *classes* of genes affecting exclusively one class of characters such as morphological, physiological, or ethological, or affecting special regions of the organism such as head, skeleton, leaves” (Sokal and Sneath 1963, 84–85; emphasis in original). The main point here is that no character is particularly special: all have complex relations to genes, and there are no groupings of genes that solely affect certain groupings of characters. As Sokal and Sneath put it, “There are no a priori grounds for favoring one character over another.” Alternatively, one could say that there is effectively no relevant structure to an organism’s genotype-phenotype map: detailed knowledge of that structure is not necessary to correctly classify the organism based on its genetic nature while using only phenotypic traits. While they devote a substantial discussion to the evidence for and against these hypotheses, the relevant outcome for us here is that they believed both hypotheses were largely correct for the majority of cases (Sokal and Sneath 1963, 85, 91).

Only given the result that characters are generally of equal worth as reflections of the genomes could Sokal and Sneath proceed to lay the fundamental ground for why NT should work. What they argued is that although it appears even one hundred characters are unlikely to represent a sample of more than 15 percent of a given genome under a simple model, what matters is not “the absolute magnitudes and qualitative nature of the characters sampled but [the] relative degrees of difference between them” (Sokal and Sneath 1963, 112–113).¹⁹ The reason is that NT treats “resemblance as a proportion of characters agreeing (matching) out of the total number being compared” (114). This proportion is then taken as an estimate of the actual proportion of matching traits that exists between two operational taxonomic units. (The issue of how to treat unmatchable characters is not discussed.)

This leads Sokal and Sneath to the matches asymptote hypothesis: “as the number of characters increases, the value of the similarity coefficient becomes more stable,” because as the sample size increases, the proportion better reflects that found in the total population of traits being sampled (Sokal and Sneath 1963, 114). The reasons that numerical taxonomic classification produces a true representation of the relation between organisms are therefore that there exists an objective set of pairwise matching relations between the operational taxonomic units’ genomes, that resemblance is solely an increasing function of the proportion of matches, and

that a (large enough) set of taxonomic characters provides an effectively unbiased estimate of that proportion. The nexus and nonspecificity hypotheses are crucial to establishing that almost any sufficiently large sample of characters should be adequate to producing an unbiased estimate of the true proportion of the total population of characters.²⁰

An additional, key factor supporting the autonomy of NT's method was Sokal and Sneath's belief in its ability to produce maximally general-purpose classifications. Compared to their treatment of the statistical properties of characters, however, their justifications for why one should expect this benefit were more vague and intuitive. Following Gilmour's idea of natural classifications, Sokal and Sneath noted that "the nature of a taxonomy depends on its purpose . . . and we choose one way because we think it is best for some purpose. If the purpose is restricted, then the classification conveys less information than a general or 'natural' one" (Sokal and Sneath 1963, 12). Hence they held "the view with Gilmour that a 'natural' taxonomy is a general arrangement intended for general use by all scientists" (12). However, their concept of purpose was the somewhat vague notion of how many things a classification can predict: "a system of classification is more natural the more propositions there are that can be made regarding its constituent classes" (19).

Broadly speaking, NT would maximize this measure of naturalness because it used the equal weighting of many characters, which was more "informative" than unequal weighting (akin to Gilmour's sense of "artificial") or the use of few characters. In addition, the use of mathematical techniques on a computer was critical: as Sneath put it in his early paper, "A classification based on the number of similar features [i.e., equally weighted] and their correlation, in other words the classification which would emerge if one fed one's data into a calculating machine, would in theory be the best classification which could be made with the available data" (Sneath 1957, 198).²¹ This very general picture of what was required to make a good classification was a crucial piece of support for the autonomy of methodology, precisely because it meant that the problem of evaluating competing methods could be separated from the questions and nature of particular taxa.

CONCLUSION

The separability of methodological questions from taxon-specific research that Sokal and Sneath articulated quickly ran into problems. Numerical taxonomy faced challenges to its claims to superiority as a methodology both on its own terms and on the broader grounds that it did not in fact serve all purposes equally well.²² Numerical taxonomy's fate in fact differed from one group of organisms to another: it was highly appreciated within the microbiology community, moderately so in botany, and only occasionally in zoology.²³ Emerging in part from that last group, systematists who wanted their classifications to represent

evolutionary history intrinsically rather than post hoc began to develop their own mathematical and computational approaches. In the two decades after the 1963 publication of *Principles*, a cladistic school of classification (and phylogenetics) coalesced and disagreed strongly with some of the tenets of NT. Although numerical taxonomists and cladists often initially mutually tolerated each other, over time their relations soured to the extent of giving rise to a period known as the “systematics wars” (Hagen 2001). However, despite this vitriol and the varying judgments on NT’s merits across taxonomic specialties, key aspects of NT’s methodology and its autonomy survived.

The impact of NT’s new methodology was not limited simply to its direct success or failure with taxon specialists, however, because its methodological framework—that is, the boundary objects distinguishing methodology’s proper domain as lying between the character matrix and dendrogram in figure 9.2—has remained largely fixed. In other words, the post-NT systematics wars have been concerned with the substance and interpretation of method, not its general organization and distinction from taxon-specific research as Sokal and Sneath first articulated. In this way, the battle for dominance operated within a framework that depended intrinsically on the use and concepts of mathematics, and therefore the integration of mathematics proceeded even as the battle between research programs intensified.

To be clear, it is obvious that cladistics has departed significantly from the steps indicated by figure 9.1; however, the form of the boundary objects between methodology and taxon-specific research, a data matrix and a branching diagram, have remained constant. As Brent Mishler has recently pointed out, “Paradoxically, despite the logical preeminence of *data matrix construction* in phylogenetic analysis, by far the greatest effort in phylogenetic theory has been directed at the second phase of analysis, the question of how to *turn a data matrix into a tree*” (Mishler 2005, 57; my emphasis). Mishler’s point exactly parallels the distinction Sokal and Sneath drew between methodology and taxon-specific work, as I argued above. Indeed, Hagen has documented that early work by James Farris shared a number of methodological tenets with NT: when Farris proposed Wagner trees and networks instead of the Camin-Sokal technique for making cladograms, “Farris did not justify it on philosophical distinctions between phenetic and cladistic techniques. He was in fundamental agreement that the solution to Hennig’s dilemma lay in the mathematical analysis of large numbers of characteristics and alternative phylogenies. . . . His was simply a better solution to a computational problem established by Camin and Sokal’s earlier paper” (Hagen 2001, 300). Although a full account of NT’s contribution to an entrenched division of work within systematics requires much further investigation, Farris’s agreement with NT about the importance of large numbers of characteristics hints at a narrative of the “statistical view” of morphological characters in systematics that

crosses continuously over ideological divides.

The theoretical foundation of methodology as articulated by Sokal and Sneath and its weak points remain relevant to contemporary problems in systematics: the issue of how to distinguish methodology from taxon-specific research in a functional way is *still* a pressing problem.²⁴ Despite the fact that the tree is ultimately only as good as the characters that go into it, the steps that transform character matrices into dendrograms have continued to receive the near-exclusive attention of methodologists. Indeed, at the heart of the issue is what constitutes a good character, including properties such as independence from other characters and the absence of too much noisy variation. Since “a systematist spends 95% of his/her time gathering and analyzing character data and less than 5% time turning the assembled data matrix into a tree,” Mishler argues, the problem of what makes a good character should be paramount for any taxon specialist interested in methodology (Mishler 2005, 59).

Sokal and Sneath provided the issue with a careful and deep conceptual treatment, and it seems that any revision or reform of how they reenvisioned the process of classification needs to have worked through the merits and weaknesses of their approach. In particular, Sokal and Sneath introduced the idea that sets of characters represent an unbiased “sample” of some underlying nature of the organism, and that a large enough sample should provide statistical convergence in the resulting tree—fifty years later, what if anything justifies this “statistical view” of characters?

Indeed, our attention to the reorganization of classificatory work carried out by Sokal and Sneath has illustrated how important the relationship is between the actions of scientists and their concepts. Just as the initial challenge facing NT was to take its broad “philosophy” of classification and demonstrate its effectiveness in practice, over time foundational concepts can become unmoored from their practical grounding. Sokal and Sneath at least had an argument for what in the nature of morphological characters (i.e., their origins in a determinative, simple genome) should justify an expectation of convergence on a single tree, yet as their argument proved insufficient historically, there does not seem to be a clear alternative to replace it.

From the perspective of epistemic visions, the historical trajectory of NT illuminates the need for an extra dimension to Agar’s question, “What difference did computers make?” Agar focused on the consequences of computers for single laboratories or institutions, but the fate of NT’s methodological innovations occurred properly at the scope of the entire field. My concern here has been to trace how Sokal and Sneath managed to generalize their local innovations into a universalist methodology, and it is commensurately important to ask what

difference computers made to *this* process, that is, beyond their impact on the prehistory of NT as a revolutionary effort. I can only point toward an answer here, but the obvious direction is to examine computers not as tools of calculation but as infrastructure and communication media. Sokal and Sneath spent considerable effort developing the computing infrastructure necessary to support the distribution of numerical taxonomic techniques, including the publication of algorithms in a specialized journal and the sale of magnetic tapes containing these programs across the United States.²⁵ To my knowledge, the existence of an (at least putatively) universal pool of shared, objective methods that could be published as contributions to the field and exchanged as commodities had no precedent in systematics. Only with the ever-growing accessibility of computers in the 1960s and an exchange network for software programs could NT methodologists make their work usable by taxon specialists, who typically did not have the mathematical and technical background to develop NT's techniques for themselves. Computers provided essential social infrastructure for the spread of the numerical taxonomic revolution, indeed also for its enduring methodological legacy.

Acknowledgments

I would like to thank Andrew Hamilton for his support in pursuing this project, and I also want to acknowledge my appreciation of the extensive comments and suggestions I received from Chris Diteresi, Trevor Pearce, Eli Gerson, Marcie Holmes, Zoe Nyssa, Jennifer Karlin, William Wimsatt, William Sterner, Margot Browning, Erin Barringer, and the CHSS Graduate Workshop. This work was funded by an NSF Graduate Research Fellowship.

NOTES

1. I should briefly clarify a historical ambiguity contained in referring to the practice of systematics over the twentieth century. For the sake of simplicity, and because I do not believe it directly affects the argument here, I am going to project backward the label “systematics” to a diverse field of heterogeneous composition and name at the turn of the century. At the time, the work of natural history was being redescribed or transformed into taxonomy or systematics, and the people doing this work were “amateurs” more often than what we would recognize as professional scientists today. Although these transformations of name, work, and profession are of great importance to the history of biology, they must remain outside the scope of this chapter. Instead, I will, as do Sokal and Sneath, use taxonomy and systematics broadly in the same sense as Simpson in 1961: taxonomy is “the theoretical study of classification, including its bases, principles, procedures and rules,” while systematics “is the scientific study of the kinds and diversity of organisms and of any and all relations among them” (Simpson 1961, 7–9). Taxonomy is therefore generally an activity within systematics, although for exactly what reason is contested.

2. The term *phenetic* is A.J. Cain's coinage, and here I use it to mean any method of classification that insists on making judgments using similarity alone at first and then potentially involving evolutionary theories or interpretations.
3. Of course, Sokal and Sneath also had a hand in preparing this ground, as they collaborated with A.J. Cain and others early on to develop the phenetic perspective (Vernon 1988).
4. My account below is based on Vernon (1988). Also, although Gilmour influenced Sneath considerably, Sneath published separately, in contrast to the collaborations of Cain and Harrison and Sokal and Michener, and therefore I follow Vernon in listing him as a separate individual.
5. Simon in fact wants to argue that AI can also deal with overlapping and backtracking problems because those ill-structured problems are not different in kind from well-structured ones, only different in degree. His solution, however, involves positing several notoriously difficult to realize capacities of computers, such as a "noticing and evoking mechanism" that is crucial to successfully redefining the problem space as new difficulties and information crop up. As these capacities are not required for solving a well-structured problem, one might wonder if he needs a qualitative leap in computing functionality in order to show how ill-structured problems differ only in degree.
6. Of course, one reason for the similarity in phenetic perspectives articulated in the 1960s was that Sokal, Sneath, and Camin began to collaborate, yet the point stands that they embarked on closely parallel trajectories before interacting (Vernon 1988). Also, it will prove relevant to my later discussion to note that the resolution of what were initially specific concrete problems proved to be of more interest to Sokal and Sneath than the original problems themselves.
7. William Wimsatt's (2007) take on theories as heuristics all the way down is particularly relevant here to understanding the *positive* relationships that different revolutionary camps have to the discipline as a whole. Heuristics by nature are flawed, and when they fail in systematic fashion, we have the potential to learn from the errors (i.e., "metabolize them"). If this chapter succeeds, it is because revolutions in systematics illuminate major functional failings of their historical predecessors for the historian and philosopher in a way that reflects not only conceptual changes but also how concepts are implicated in the daily practical needs of the scientists.
8. I mean "know" in a generally pragmatic sense here. As Wittgenstein put it, "But is a blurred concept a concept at all? — Is an indistinct photograph a picture of a person at all? Is it even always an advantage to replace an indistinct picture by a sharp one? Isn't the indistinct one often exactly what we need?" (Wittgenstein 2001, 34e).
9. Mark Wilson's (2006) discussion of Heaviside's adventurous use of calculus to model electrical currents in circuits, and the applied mathematicians who made sense of it only decades later, makes a similar point, though Wilson focuses on the "seasonality" of concepts as abstracted from their part in organizing the research practices of a scientific

community.

10. Sokal and Sneath in fact cited Simpson, Hennig, and Remane as having provided notable efforts at “analytic studies of systematic principles,” but none of them resolved the problem of “circular reasoning” in a satisfying fashion for Sokal and Sneath (Sokal and Sneath 1963, 2).

11. For a later review emphasizing plants, see Pankhurst 1978.12. Vernon (2001) discusses some of the second generation of NT.13. Gerson (1998) has a related distinction between taxon specialties and

analytic specialties, e.g., population genetics. The main difference between analytic specialties and methodologists is that methodologists are strictly speaking a subspecialization within taxonomy rather than an outside specialty.

14. While these particular terms are not actor’s categories, similar distinctions had existed long before the publication of *Principles*. For example, see Simpson’s discussion of Gregg’s distinction between “taxonomy proper,” which is analogous to taxon-specific research, and “methodological taxonomy” (Gregg 1954; Simpson 1961).

15. They had also earlier on written that “it is possible to formulate a list of characters and their states that do not presuppose a prior knowledge of homologies” (74), thereby bypassing a great portion of the relevant expertise of the taxon specialist. In fact, Sokal later went further and attempted to demonstrate that even the discovery and measurement of characters required no specialized knowledge or training. (It did not work as hoped; see Sokal and Rohlf 1966, 1970; Rohlf 1965.)

16. I will not touch here on their emphasis on OTUs as a way of avoiding rank-dependent problems in step 1 of their classification process, but this is an additional, important reason for why NT could separate itself from taxonspecific research.

17. Olivier Rieppel has investigated these issues for cladistics, and my section here is a parallel, preliminary account of similar problems. See Rieppel 2003, 2007.

18. They also added, “however, a study to investigate the influence of environment using numerical taxonomic methods could legitimately include environmentally determined characters.”

19. Note that although Sokal and Sneath discard the factor asymptote hypothesis as “not very fruitful,” they do not reject the reasoning behind it that treats characters as forming a “random sample of the genome of the operational taxonomic unit.”

20. Sokal and Sneath make a few hedges here involving requiring sampling characters across an organism’s life cycle and including a broad variety of whatever characters show variation. See pp. 92–93 and others.

21. While Sokal and Sneath did have a way of checking how well the dendrogram diagram represented the initial measurements of affinities (the co-phenetic correlation

coefficient; see Sokal and Sneath 1963, 189–194), it would require another step of logic to demonstrate that the whole process maximized the number of possible propositions that could be made using the classification.

22. James S. Farris (1969) found that the co-phenetic correlation coefficient, the primary measure of optimality NT had used, did not in fact measure what Sokal and Sneath thought it did.

23. About thirty years after *Principles* first appeared, approximately 3,000 papers had been published applying NT to bacterial classifications, and the subspecialty had indeed used NT to increase the objectivity and repeatability of its character measurements and classifications. According to a similar estimate for botany, about 500 to 1,000 publications had appeared applying NT to plants in the same time period. Far fewer appeared within zoology. See Vernon 2001; Hagen 2001; Hull 1988; Goodfellow, Jones, and Priest 1985; Duncan 1981.

24. Winther (2009) has set out to address this problem, among others. 25. See, e.g., Rohlf 1969.

REFERENCES

Agar, J. 2006. "What Difference Did Computers Make?" *Social Studies of Science*, 36: 869–907.

Duncan, T. 1981. "Numerical Phenetics: Its Uses in Botanical Systematics." *Annual Review of Ecology and Systematics*, 12: 387–404.

Farris, J.S. 1969. "On the Co-Phenetic Correlation Coefficient." *Systematic Zoology*, 18: 279–285.

_____. 1970. "Methods for Computing Wagner Trees." *Systematic Zoology*, 19: 83–92.

_____. 1972. "Estimating Phylogenetic Trees from Distance Matrices." *American Naturalist*, 106: 645–668.

Farris, J.S., Kluge, A.G., and Eckardt, M.J. 1970. "A Numerical Approach to Phylogenetic Systematics." *Systematic Zoology*, 19: 172–189.

Felsenstein, J. 2001. "The Troubled Growth of Statistical Phylogenetics." *Systematic Biology*, 50: 465–467.

_____. 2004. *Inferring Phylogenies*. Sunderland, MA: Sinauer Associates.

Galison, P. 1997. *Image and Logic: A Material Culture of Microphysics*. Chicago: University of Chicago Press.

Gilmour, J.S.L. 1940. "Taxonomy and Philosophy." In J. Huxley (ed.), *The New Systematics*. New York: Oxford University Press, 461–474.

- Goodfellow, M., Jones, D., and Priest, F.G., eds. 1985. *Computer-Assisted Bacterial Systematics*. New York: Academic Press.
- Gerson, E. M. 1998. "The American System of Research: Evolutionary Biology, 1890–1950." Ph.D. dissertation, University of Chicago.
- Gregg, J. R. 1954. *The Language of Taxonomy*. New York: Columbia University Press.
- Hagen, J.B. 1984. "Experimentalists and Naturalists in Twentieth-Century Botany: Experimental Taxonomy, 1920–1950." *Journal of the History of Biology*, 17: 249–270.
- _____. 2001. "The Introduction of Computers into Systematic Research in the United States during the 1960s." *Studies in the History and Philosophy of Biology and Biomedical Sciences*, 32: 291–314.
- _____. 2003. "The Statistical Frame of Mind in Systematic Biology from *Quantitative Zoology* to *Biometry*." *Journal of the History of Biology*, 36: 353–384.
- Hamilton, A., and Wheeler, Q.D. 2008. "Taxonomy and Why History of Science Matters for Science: A Case Study." *Isis*, 99: 331–340.
- Hennig, W. 1966. *Phylogenetic Systematics*. Trans. D. D. Davis and R. Zangerl. Urbana: University of Illinois Press.
- Hine, C. 2008. *Systematics as Cyberscience: Computers, Change, and Continuity in Science*. Cambridge, MA: MIT Press.
- Hull, D. L. 1988. *Science as Process: An Evolutionary Account of the Social and Conceptual Development of Science*. Chicago: University of Chicago Press.
- Huxley, J., ed. 1940. *The New Systematics*. New York: Oxford University Press.
- Johnson, K. 2005. "Ernst Mayr, Karl Jordan, and the History of Systematics." *History of Science*, 43: 1–35.
- _____. 2009. "The Return of the Phoenix: The 1963 International Congress of Zoology and American Zoologists in the Twentieth Century." *Journal of the History of Biology*, 42: 417–456.
- Kuhn, T.S. 1962. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.
- Mayr, E. 1942. *Systematics and the Origin of Species: From the Viewpoint of a Zoologist*. Cambridge, MA: Harvard University Press.
- Mayr, E., and Provine, W. B. 1980. *The Evolutionary Synthesis: Perspectives on the Unification of Biology*. Cambridge, MA: Harvard University Press.

- Mishler, B. D. 2005. "The Logic of the Data Matrix in Phylogenetic Analysis." In V. A. Albert (ed.), *Parsimony, Phylogeny, and Genomics*. Oxford: Oxford University Press, 57–70.
- Pankhurst, R.J. 1978. *Biological Identification: The Principles and Practice of Identification Methods in Biology*. Baltimore, MD: University Park Press.
- Rainger, R., Benson, K.R., and Maienschein, J., eds. 1988. *The American Development of Biology*. Philadelphia: University of Pennsylvania Press.
- Rieppel, O. 2003. "Semaphoronts, Cladograms, and the Roots of Total Evidence." *Biological Journal of the Linnean Society*, 80: 167–186.
- _____. 2007. "The Nature of Parsimony and Instrumentalism in Systematics." *Journal of Zoological Systematics and Evolutionary Research*, 45 (3): 177–183.
- Rohlf, J.F. 1965. "A Randomization Test of the Nonspecificity Hypothesis in Numerical Taxonomy." *Taxon*, 14: 262–267.
- _____. 1969. "GRAFPAC, Graphic Output Subroutines for the GE 635 Computer." *Kansas Geological Society Computer Contributions*, 36.
- Simon, H. 1962. "The Architecture of Complexity." *Proceedings of the American Philosophical Society*, 106: 467–482.
- _____. 1973. "The Structure of Ill Structured Problems." *Artificial Intelligence*, 4: 181–201.
- Simpson, G.G. 1961. *Principles of Animal Taxonomy*. New York: Columbia University Press.
- Sneath, P.H.A. 1957. "Some Thoughts on Bacterial Classification." *Journal of General Microbiology*, 17: 184–200.
- Sokal, R.R., and Rohlf, J.F. 1966. "Random Scanning of Taxonomic Characters." *Nature*, 210: 461–462.
- _____. 1970. "The Intelligent Ignoramus: An Experiment in Numerical Taxonomy." *Taxon*, 19: 305–319.
- Sokal, R. R., and Sneath, P. H. A. 1963. *Principles of Numerical Taxonomy*. San Francisco: W.H. Freeman.
- Star, S. L., and Griesemer, J. "Institutional Ecology, 'Translations' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907–39." *Social Studies of Science*, 19 (3): 387–420.
- Stebbins, G. L. 1950. *Variation and Evolution in Plants*. New York: Columbia University Press.

Strasser, B.J. 2010. "Collecting, Comparing, and Computing Sequences: The Making of Margaret O. Dayhoff's *Atlas of Protein Sequence and Structure*, 1954–1965." *Journal of the History of Biology*, 43: 623–660.

Vernon, K. 1988. "The Founding of Numerical Taxonomy." *British Journal of the History of Science*, 21: 143–159.

_____. 1993. "Desperately Seeking Status: Evolutionary Systematics and the Taxonomists' Search for Respectability, 1940–1960." *British Journal of the History of Science*, 26: 207–227.

_____. 2001. "A Truly Taxonomic Revolution? Numerical Taxonomy, 1957– 1970." *Studies in the History and Philosophy of Biology and Biomedical Sciences*, 32: 315–341.

Wimsatt, W.C. 2007. *Re-Engineering Philosophy for Limited Beings: Piecewise Approximations to Reality*. Cambridge, MA: Harvard University Press.

Winther, R.G. 2009. "Character Analysis in Cladistics: Abstraction, Reification, and the Search for Objectivity." *Acta Biotheoretica*, 57: 129–162.