

# This Week's Citation Classic®

Sneath P H A & Sokal R R. *Numerical taxonomy: the principles and practice of numerical classification*. San Francisco: Freeman, 1973. 573 p.

[Medical Research Council Microbial Systematics Unit, Univ. Leicester, England and Dept. Ecology and Evolution, State Univ. New York, Stony Brook, NY]

Numerical taxonomy establishes classification of organisms based on their similarities. It utilizes many equally weighted characters and employs clustering and similar algorithms to yield objective groupings. It can be extended to give phylogenetic or diagnostic systems and can be applied to many other fields of endeavour. [The SCI® indicates that this book has been cited in over 2,015 publications.]

P.H.A. Sneath  
Department of Microbiology  
University of Leicester  
Leicester LE1 7RH  
England

June 6, 1986

I became interested in classification as a young researcher with the Medical Research Council in London in the late 1950s. I had been studying a poorly known bacterial genus, *Chromobacterium*, common in soil and water, that caused a rare human infection, which I had observed in Malaya while serving with the Royal Army Medical Corps. The question arose as to how many species there were. I read up on taxonomic theory and concluded that there was very little logic in how classifications were made; this led to attempts to sketch a logical outline for the process.

One problem proved intractable: how should one weight the different attributes of organisms? The realization that attributes should logically have equal weights, because their apparent importance arose from their mutual correlations, came to me suddenly while riding on a red double-decker London bus at Mill Hill Circus. Soon a paper illustrated using *Chromobacterium* was ready—initially with manual calculations, but later using a computer<sup>1</sup> at the nearby firm of Elliott Brothers Ltd. One day, the distinguished virologist Sir Christopher Andrewes—who was a keen entomologist—put his head in the door

to say that there was someone else “who was also mad enough to apply computers to taxonomy,” and he put me in touch with Robert Sokal at the University of Kansas, where we met in 1959. He was working with C.D. Michener on the systematics of bees.<sup>2</sup> Our initial collaboration largely took place when Sokal was on sabbatical at the Galton Laboratory in London University; it led to our first book on numerical taxonomy.<sup>3</sup> This “showed that it could be done,” one might say, and attracted both attention and controversy.<sup>4</sup>

Our second book (the *Classic*) was not a second edition but was entirely reworked. We reorganized the logical structure of numerical taxonomy and tried to place the successive steps in proper sequence. Many topics, such as homology, that had been glossed over before were given detailed discussion. The major advances had been in two areas, numerical methods for reconstructing phylogenies and numerical techniques for identification. Sokal took a special interest in phylogeny, which has since become an extremely controversial field.<sup>5</sup> My own interests were turning to numerical identification; this is now an intensely practical field in microbiology,<sup>6</sup> where computer techniques are becoming the norm for modern diagnostic instruments. Both of these newer fields rest on a firm basis of numerical taxonomy in matters of symbolism, choice of organisms and characters, coding and scaling of attributes, resemblance measures, and methods for displaying taxonomic structure.

One reason the book has been so widely cited is that it was very comprehensive, covering all aspects of classification including nonbiological applications. It contained numerous formulae and descriptions of algorithms but was not too heavily mathematical. It also reviewed critically the voluminous literature of the preceding 10 years.

1. Sneath P H A. The application of computers to taxonomy. *J. Gen. Microbiol.* 17:201-26, 1957. (Cited 390 times.)

2. Michener C D & Sokal R R. A quantitative approach to a problem in classification. *Evolution* 11:130-62, 1957. (Cited 65 times.)

3. Sokal R R & Sneath P H A. *Principles of numerical taxonomy*. San Francisco: Freeman, 1963. 359 p. (Cited 1,555 times.)

4. Sokal R R. Citation Classic. Commentary on *Principles of numerical taxonomy*. San Francisco: Freeman, 1963. 359 p. in *Current Contents/Agriculture, Biology & Environmental Sciences* 13(46):14, 15 November 1982.

5. Feisenstein J. Numerical methods for inferring evolutionary trees. *Quart. Rev. Biol.* 57:379-404, 1982.

6. Goodfellow M, Jones D & Priest F G. *Computer-assisted bacterial systematics*. London: Academic Press, 1985. 443 p.

1A-15 L E