CHAPTER 1

Introduction

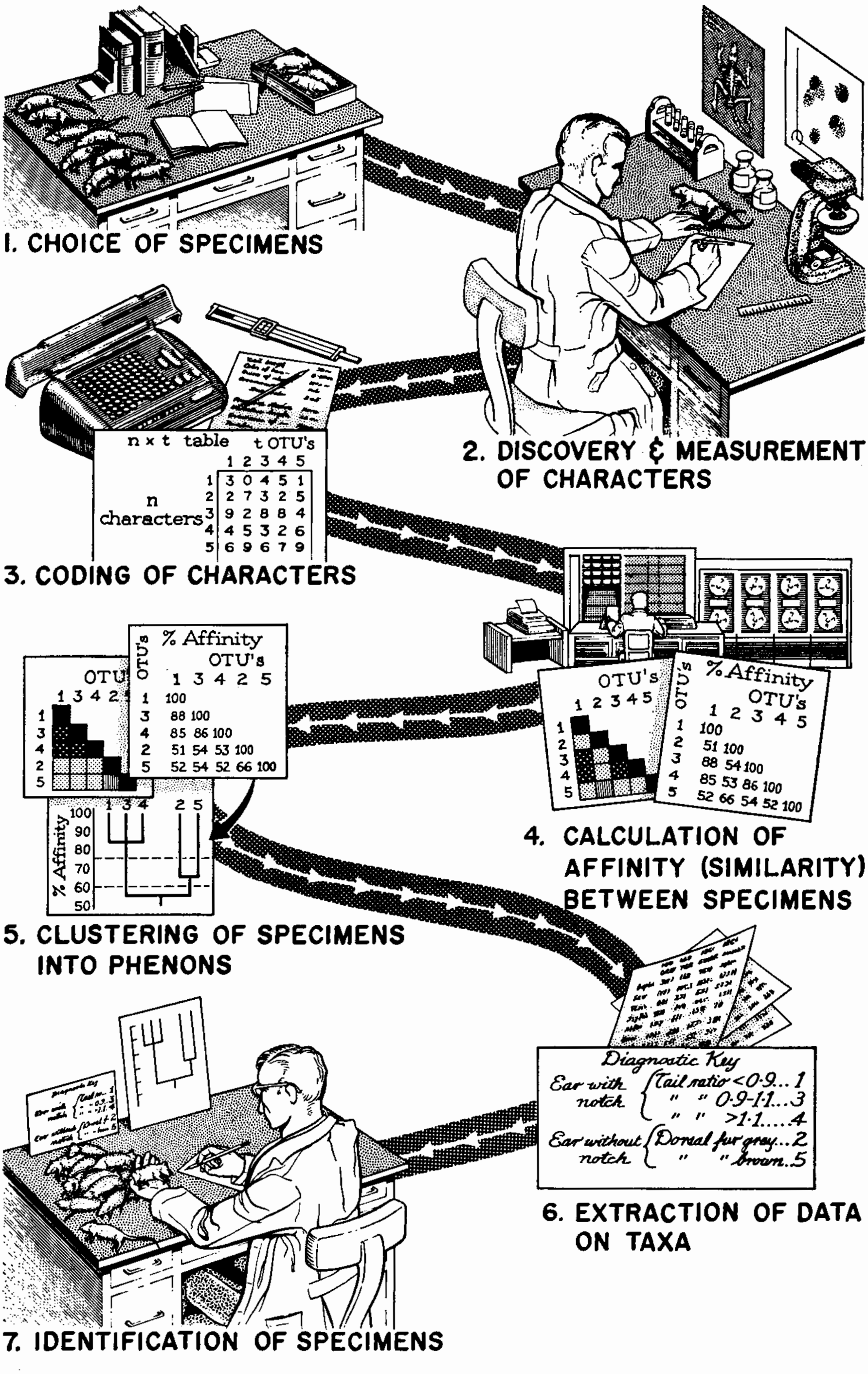
1.1. SUMMARY OF CONTENTS

The contents of this book fall into three main parts. The first part provides a background to the field of taxonomy in general and an introduction to numerical taxonomy in particular. In Chapter 2 we discuss in some detail the historical development of the conceptual basis of classification and our criticisms of current taxonomic practices. Chapter 3 treats various quantitative methods in taxonomy, other than those described in detail in subsequent chapters. This is followed in Chapter 4 by a discussion of the aims and principles of numerical taxonomy.

The central part of the book is arranged on a plan which closely reflects the successive steps followed by taxonomists, unconsciously for the most part, in performing the classificatory process. In Chapter 5 are discussed the choice of characters for numerical taxonomy, the coding of these characters for mathematical manipulation, and the selection of organisms for study. The estimation of taxonomic resemblance between organisms follows in Chapter 6. Chapter 7 considers the grouping of organisms into taxa on the basis of these resemblances.

The final part deals with the implications of numerical taxonomy for systematic research. We discuss the relation of numerical taxonomy to phylogeny and paleontology in Chapter 8 and to nomenclature and diagnostic keys in Chapter 9. Possible future developments and similar work in subjects other than the classification of living creatures are discussed in Chapter 10. An appendix has been added, in which details of the mathematical methods of numerical taxonomy are given, illustrated by elementary examples.

A FLOW CHART OF NUMERICAL TAXONOMY



1.2. GUIDE TO THE USE OF THE BOOK

The summary of contents given above will itself serve as the most useful guide to the use of this book, but it may be well to emphasize a number of additional points. First, no critical systematist will employ a method of whose validity he is unconvinced. The justification for our standpoint is to be found particularly in Chapters 2 and 4; the former is mostly a criticism of current practices, and the latter is an attempt to construct logical alternatives to these. We would suggest that these chapters be read before embarking on the later ones; not only do they provide the justification for the methods, but there is also some danger in the unimaginative application of the procedures if the principles on which they are based have been imperfectly understood.

Second, the computational chapters will be most profitably read if reference is made to the pertinent examples in the Appendix, as indicated by the cross-references. The mathematical treatment has been made as simple as possible, and the examples should aid its painless assimilation. We have preferred to err on the side of prolixity in the interests of making the computations easily comprehensible, and only a slight knowledge of elementary mathematics and statistics is necessary.

Third, it should be emphasized that for any considerable body of data the assistance of computers is essential, not indeed because of the difficulty of the mathematical manipulations (for these can readily be performed with paper and pencil or desk calculators if only a few organisms are studied) but because with extensive data the analyses are too time-consuming to be practicable without computational aids. The Appendix therefore contains some notes on computer programs and the preparation of data for them.

1.3. DEFINITIONS OF CERTAIN TERMS

The adequate definition of taxonomic terms would almost require a book by itself. Many terms are used in so many different senses that we have had great difficulty in using them in a consistent and exact manner, and no doubt there are still a great many ambiguities which we have overlooked. The meanings of terms and symbols as they are used in this book can be looked up via the Index. There are, however, several which are employed so frequently that we have deemed it desirable to present them here at the outset.

Classification, systematics, and taxonomy are often used interchangeably. In recent years, especially in the United States, there has been a trend toward assigning separate meanings to these terms. In this sense they have been well defined by Simpson (1961) and we follow his usage here.

Systematics. ". . . *is the scientific study of the kinds and diversity of organisms and of any and all relationships among them*" (Simpson, 1961, p. 7). This definition uses systematics in its widest sense, concerning itself not only with the arrangement of organisms into taxa and with naming them, but also with the causes and origins of these arrangements.

Classification. We have adapted Simpson's definition (1961, p. 9), which restricted itself to zoological classification, to more general usage. *Classification is the ordering of organisms into groups (or sets) on the basis of their relationships, that is, of their associations by contiguity, similarity, or both.* We have restricted the definition above 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. Classification as defined above is the name of a process; however, it has also been used for the end product of this process. Thus the result of classification is a classification. The term classification has also been employed, mainly in fields outside biology, in the restricted sense of putting entities into distinct classes as opposed to arraying them in a continuous spectrum, cline, or other arrangement showing no distinct divisions. We have not restricted the term in this manner.

Taxonomy. ". . . *is the theoretical study of classification, including its bases, principles, procedures and rules*" (Simpson, 1961, p. 11). By this definition the bulk of the subject matter of our book is concerned with taxonomy, and for the same reason we have called our subject numerical taxonomy, rather than numerical systematics. Taxonomy, like classification, has also been used to designate the end products of the taxonomic process. Since this is a generally accepted usage, we will occasionally employ it in this sense.

There may be confusion over the term "relationship." This may imply relationship by ancestry (of which there are several kinds—see Section 8.1.2), or it may simply indicate the overall similarity as judged by the characters of the organisms without any implication as to their relationship by ancestry. For this meaning of overall similarity we have used the term "affinity," which was in common use in pre-Darwinian times. We may also distinguish this sort of relationship from relationship by ancestry by calling it *phenetic relationship*, employing the convenient term of H. K. Pusey as used by Cain and Harrison (1960b), to indicate that it is judged from the phenotype of the organism and not from its phylogeny.

We use the term *taxon* (plural *taxa*) as an abbreviation for taxonomic group of any nature or rank, as suggested by Lam (in Lanjouw, 1950) and Rickett (1958).