C. Tolmie

REFIGURING ANTHROPOLOGY

First Principles Of Probability & Statistics

David Hurst Thomas
American Museum of Natural History

Waveland Press, Inc.
Prospect Heights, Illinois

For information about this book, write or call:

Waveland Press, Inc.
P.O. Box 400
Prospect Heights, Illinois 60070
(312) 634-0081

For permission to use copyrighted material, the author is indebted to the following:

FIG. 2.1. By permission of the Trustees of the British Museum (Natural History).
TABLE 2.3. (p. 24) From *Physical Anthropology: An Introduction* by A. J. Kelso. Reprinted by permission of the publisher, J. B. Lippincott Company. Copyright © 1974. (p. 25) Reproduced by permission of the Society for American Archaeology from *Memoirs of the Society for American Archaeology*, Vol. 11, 1956.
FIG. 3.1. From Hulse, Frederick S. *The Human Species: An Introduction to Physical Anthropology*. Copyright © 1963 by Random House, Inc.

FIG. 3.2. From Dozier, Edward P., *The Pueblo Indians of North America*. Copyright © 1970 by Holt, Rinehart and Winston, Inc. Reproduced by permission of Holt, Rinehart and Winston. [This book reissued 1983 by Waveland Press, Inc.]
FIG. 3.4. Reproduced by permission of the Society for American Archaeology from *American Antiquity*, Vol. 35 (4), 1970.
FIG. 3.5. Reproduced by permission of the American Anthropological Association from the *American Anthropologist*, Vol. 73 (3), 1971.
FIG. 13.14. From *Biometry* by Robert R. Sokal and F. James Rohlf, W. H. Freeman and Company. Copyright © 1969.

Second Printing
The 1976 version of this book was entitled *Figuring Anthropology*.

ISBN 0-88133-223-2

Printed in the United States of America.

2 What Are Anthropological Data?

● *The subject of anthropology is limited only by man.*
—A. Kroeber

2.1 INTRODUCTION

This chapter begins in backward fashion. Rather than considering what data *are,* we must first peruse what data *are not.* For one thing, the word "data" is not singular; "datum" is the singular form, although this term is rarely used in statistical contexts. So, statements about a set of measurements are properly worded "the data are . . ." rather than "the data is . . . ."

More importantly, even though anthropology is generally defined as the study of people, *people do not constitute the data of anthropology.* People are people. What anthropologists study are observations *about* people rather than the people themselves. Along this same line, skulls and bones are not the data of physical anthropology any more than artifacts, temple tombs, or housepits are the data of archaeology. Data are not people, objects, or things; data are counts, measurements, and observations *made on* people, objects, and things. Twenty Neanderthal crania are not in themselves a set of data; the cranial capacities, cranial lengths, or nasal widths of these skulls comprise the data. There are no data until an anthropologist observes them. Data do not passively exist. Data must be generated.

A couple of major points follow from this active definition of data. For one thing, those who would accuse anthropologists of "using people as data" neither understand what data are nor how they are collected. Clearly, anthropologists manipulate observations (the *real* data), not people.

Secondly, data can hardly be destroyed once they are generated. A society might become extinct, an archaeological site may be bulldozed to make way for a parking lot, a skeletal series might be lost or destroyed, but all these disasters occur only to people or objects or things, not to data. One case in point is the strange, unfinished saga of the famous Peking man fossils, which has been related in marvelous detail by Harry L. Shapiro (1971). At this writing, the fossils are lost to science, their whereabouts unknown. The fossils disappeared in 1941 when a Japanese invasion of the Chinese mainland seemed imminent. Professor Franz Weidenreich, who had been studying the fossils at the Peking Union Medical College, was forced to flee to the United States, crating the *Sinanthropus* fossils for secret transport. But they disappeared. It appeared for a while as though the fossils might have been lost overboard during loading onto the *S.S. President Harrison*. But new evidence has recently come to light indicating that the fossils safely reached Camp Holcombe in Chinwangtao, where Japanese soldiers confiscated and apparently ransacked the crates. The fossils may have been discarded as worthless junk. But there is also the possibility that some of the crates—those destined for shipment to a Swiss warehouse, the Pasteur Institute, and the homes of reliable citizens in Tientsin—were stored in Chinese warehouses, or perhaps even taken to Tokyo as war booty (as was the Solo skull, which had been taken from G. H. R. von Koenigswald in a Japanese prison camp). Perhaps the fossils miraculously reached the United States and will someday surface. But, even assuming the worst—that the fossils were totally destroyed—the *data* on *Sinanthropus* have not been lost to science because detailed measurements, photographs, and observations were published by Weidenreich in 1943.

In order to actually "destroy data," one would have to destroy every original and every copy of the published descriptions, a virtual impossibility. Whatever the fate of the fossils, the Peking man *data* remain as viable as the day the finds were first analyzed. Of course, if the original fossils could be found, then *new* data could be generated, using modern techniques developed in the three decades since the fossils disappeared.

Data can also take a number of forms. Data can be recorded as counts or as measurements or as observations. Data can also be generated from other data, as with ratios and powers. Data can be in the form of variables and constants, variates and populations, samples and statistics. This chapter is concerned primarily with how the individual observations—the "raw" data—are transformed into statistically *meaningful forms*, a discussion which supplies the foundation for much of statistical theory.

2.2 CONSTANTS AND VARIABLES

● *Freedom is the freedom to say that two plus two makes four. If that is granted, all else follows.*—G. Orwell

Philosopher Bertrand Russell once pointed out that the reason some people have so much trouble understanding what *X* means is that *X* doesn't mean anything at all. *X* is nothing but a symbol, and symbols assume meaning only after they are assigned to a particular characteristic. Symbols such as *X*, *Y*, and *w* have no natural or necessary relationship to their assigned referent. Symbols are merely arbitrary notations.

Anthropology is often concerned with the meaning and significance of symbols. In fact, there is even a movement within modern ethnology whose adherents term themselves "symbolic anthropologists." Symbols pervade our everyday life to such an extent that we cannot think without them. Leslie White (1940) has gone so far as to term symbols "basic units of all human behavior and civilization . . . the symbol is the universe of humanity."

But if symbols are truly basic to humanity, then why do so many humans (especially social science undergraduates) seem to fear X? X is just a symbol. Why should X be any more terrifying a symbol than, say, a stop sign, the peace sign, or the Star of David? A recent report by the Mathematical Association of America cautioned statistics instructors to allow for "symbol shock" suffered by many introductory college students. The problem, of course, is that X belongs to a very special class of symbols—*mathematical symbols*—and a large segment of the Western population has been covertly programmed to fear the encroachment of mathematical logic and methodology. Nowhere is this aversion to mathematical symbolism seen more clearly than in the application of statistical methods to social science. Perhaps by looking more closely at what statistical symbols actually symbolize, we can make some inroads at dispelling "symbol shock."

The most elementary use of mathematical symbols is to denote a constant. In this case, the symbol and its referent have a one-to-one relationship.

> A *constant* is a quantity (denoted by a symbol) which can assume only one value.

Mathematics is rife with constants, often assigned a conventional symbol.

$\pi = 3.14159265$	(ratio of diameter to circumference)
$e = 2.71828183$	(base of natural logarithms)

Constants are named according to scientific convention rather than because of any natural isomorphism between the symbol and its characteristic. The Greek letter π is no more suitable to designate 3.14159265 than is any other symbol.

Anthropology has its share of constants too, but the symbols for these constants are more tractable than those of mathematics. Since each specimen of *Homo sapiens* has exactly two ears, for example, we could say that 2 is a constant for our species. Ego has precisely two biological parents or four biological grandparents, so these numbers are other biological constants. But anthropological constants need not always be so trivial. *Naroll's constant*, for example, attempts to relate the floor area of a settlement to the size of human population living in that area. In general, the population of a given settlement is about one-tenth of the floor area, expressed in square meters (Naroll 1962a):

$$\text{population} = 0.10 \text{ (floor area in square meters)}$$

Naroll's constant ($c = 0.10$) is particularly useful to archaeologists; once the floor area of a prehistoric settlement or structure is known, the prehistoric population can be estimated. At the Thomas Riggs site in South Dakota, for example, a longhouse was excavated which covered about 260 square meters. Applying Naroll's constant,

$$\begin{aligned}\text{population} &= 0.10\ (260\ \text{m}^2)\\ &= 26.0 \text{ individuals}\end{aligned}$$

Naroll's constant thus estimates the prehistoric population of the structure at the Thomas Riggs site to be about 26 people. The longhouse contained four hearths, so if each hearth served a single nuclear family, the average family size must have been about 6.5 people. This estimate corresponds closely with the known ethnographic information for this area, and provides some support for the proposition that Naroll's constant is *in fact* relatively constant (that is, invariant).

Another useful constant in anthropology is *Shapiro's constant for cranial deformation* (Shapiro 1929). Artificial deformation of the skull is a cultural trait worldwide in its distribution, with the primary occurrence in pre-Columbian America (especially among the Classic Maya), Peru, and the American Southwest. The specifics vary from region to region, but in all cases the infant's skull was altered to correspond with the local ethnic conception of beauty. Sometimes the skull was simply flattened by pressure from a cradleboard, while other groups deformed the frontal portion of the skull as well. The difficulty for physical anthropologists is that the cultural deformation of crania renders many anthropometric measurements virtually useless for comparison to undeformed skulls. Because of the frequency of artificial deformation, the craniometry of several areas of the world was simply unknown.

Shapiro attacked this problem by reasoning that while the shape of the cranial vault is drastically altered by deformation, the facial and frontal areas are left essentially unchanged, even in highly deformed skulls. Thus, the diameter from basion to nasion ought to reflect undeformed cranial characteristics even in deformed skulls, although the cranial length would be much too short (Fig. 2.1). This relatively constant basion–nasion diameter could then be used to correct the deformed length. Shapiro tested his idea upon a series of 1400 undeformed skulls from throughout the world, and computed a "constant":

$$\text{cranial length} = 1.49\ (\text{basion-nasion diameter})$$

For every unit change in cranial length, there is a corresponding change of roughly 1.49 units in the basion–nasion diameter. To obtain the corrected head length, the difference between the averages of the basion–nasion diameter in deformed and undeformed crania is multiplied by Shapiro's constant ($c = 1.49$). This product is then subtracted from the undeformed head length to obtain an estimate for length in deformed crania. Earnest Hooton (1930:39) provides the following example to illustrate how Shapiro's constant enabled him to correct for cranial deformation in the skeletons from the Pecos Pueblo in central New Mexico:

Cranial length (undeformed males), mm	175.74
Cranial length (deformed males)	(?)
Basion–nasion diameter (undeformed males), mm	102.70
Basion–nasion diameter (deformed males), mm	101.58

The difference in basion–nasion diameter between the deformed and undeformed crania is

$$101.58 - 102.70 = -1.12\ \text{mm}$$

which corrects to -1.67 mm when multiplied by Shapiro's constant ($c = 1.49$).

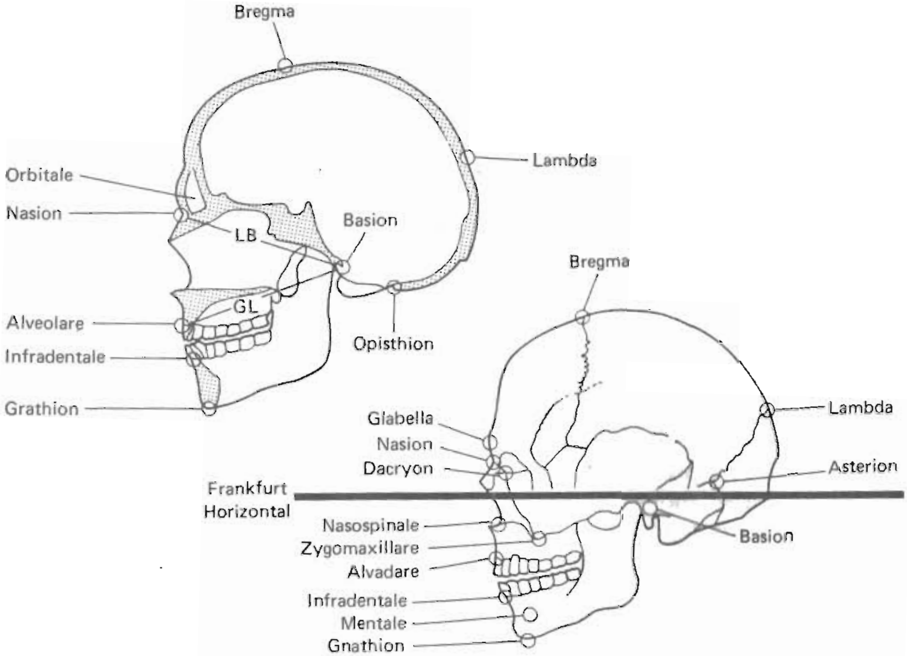


Fig. 2.1 Principal craniometric points and measurements of the human skull (after Brothwell 1963:fig. 33).

When this correction factor is added to the undeformed cranial length, the corrected head length of the deformed series can be estimated to be 174.07 mm. Hooton used Shapiro's constant to correct for deformation in over 300 crania, which otherwise would have been useless for further comparative analysis.

Symbols can thus denote a handy set of constants, but symbols are even more frequently used in statistics to denote *variables*. Most people know that variables are what scientists observe, and variables are most commonly assigned conventional letter symbols such as X, Y, or Z. What makes variables different from constants is that while *constants* must remain *constant*, *variables vary*.

> A *variable* is a measurable quantity (represented by a symbol) which is free to assume more than one value.

By this definition, variables must always have *at least two* "states" or potential values.

> A *variate* is an individual measurement of a variable.

Variables are thus abstractions, while the real data of anthropology consist of variates—the observations themselves and their measurements.

Cranial capacity is an example of a variable, which we can symbolize by the term X (or any other symbol). When dealing with a series of measurements, the variates are often denoted by the symbol X_i (read as "X sub i"), where i is a subscript specific to each variate. The cranial capacities of four individual Neanderthal skulls can be conveniently recorded in this shorthand (data from Coon 1971a: table 39):

Specimen	Symbol		Variate, cc
La Ferrassie I	X_1	=	1641
Spy 1	X_2	=	1525
Spy 2	X_3	=	1425
Gibraltar	X_4	=	1300

Using this symbolic form of notation, one can summarize even massive sets of variates with relative ease.

All variables vary, but it is *important* to note that all variables do not share the same underlying mathematical structure. Some variates are exact *observations*, while other variates are mere approximations of unknown measurements. A variable is called *discrete* if it can assume only certain fixed, predetermined values. "Number of teeth per species" is a discrete variable, for instance. We know that some primitive mammals had 44 teeth and that some varieties of South American monkeys have 36 teeth, while others have 32 teeth. Both humans and apes generally have 32 teeth. This variable is discrete because tooth number can assume only certain values—in this case, positive integers—even though the number of teeth varies between species. We can eliminate on logical grounds the possibility of ever discovering a mammalian skull with exactly 33.38 teeth. Teeth may be broken, of course, but we can be certain that, regardless of the true number before breakage, the total number was a positive integer. Thus, discrete variables are always exact measurements.

Most discrete variables common to anthropology are counts: the number of sacral vertebrae in gorilla, the number of sites located in an archaeological survey, the number of individuals whom ego calls FaBr, the total number of female children born in society *Y* during the last calendar year. Each case involves exact whole number measurements, but not all discrete variables are counts. Expected Mendelian genetic frequencies can generally assume only a few fixed values, such as 3:1, 2:1, or 9:3:3:1. These ratios are not free to assume all possible intervening values, so Mendelian ratios are also discrete variables.

Variables are termed *continuous* if the variates can logically assume any interval of measurement. The precision of observations generally determines which measurement interval should be applied, and these intervals are mere approximations. Discrete variables are always exact, but a continuous variable can never be exact because its measurement is an approximation. The cranial capacity of the La Ferrassie I Neanderthal skull, for instance, was estimated at 1641 cc (cubic centimeters). Unlike a discrete variable, cranial capacity is perfectly free to assume any possible value measurable in cubic centimeters. There would be no objection to a variate of 1641.2 cc or even 1641.34920593 cc, provided the measurements could be sufficiently accurate.

Continuous variables generally involve variates measured by common physical units: time, length, or mass. Body stature, body weight, the average score on the Graduate Record Examination, population density, daily caloric intake, and birth rate are all continuous variables because the final figure is determined by the accuracy of measurement rather than by logic.

Sometimes a continuous variable might be purposely defined as discrete. "The relative degree to which a society depends upon animal husbandry," for example, is a continuous variable measured as a percentage. But because of the difficulty in estimating accurately, and also to facilitate coding upon computer punch cards, column 10 in the *Ethnographic Atlas* (Murdock 1967) lists the variable "relative dependence upon animal husbandry" in ten categories:

0 0–5 percent
1 6–15 percent
2 16–25 percent
.
.
.
9 86–100 percent

The opposite situation can also occur: A seemingly discrete variable can be refined into a continuous variable, and this is precisely what happened in the case of measuring human skin color. For decades, physical anthropologists characterized skin color by discrete categories such as "very light," "light," "intermediate," "dark," and "very dark." But a reflectometer has recently been used to measure light reflectance from human skin. The darker the skin, the more light is absorbed; therefore skin color can now be characterized by "the percent of light reflectance," a continuous variable which allows for greater precision and objectivity. There are even laboratory methods for determining the precise amount of melanin present in human skin, but these techniques remain impractical for massive population surveys.

The refinement of discrete variables into more precise, continuous variables is sometimes considered an obvious sign of progress in science. But one must remember that while each method of testing may purport to characterize a variable such as "skin color," a change in measuring technique usually involves a new operational definition of the variable. The skin color determined by visual inspection does not exactly correspond to skin color measured by reflectometer readings. Caution must be exercised when comparing findings resulting from different techniques.

● *SYMBOL, Nn. Something that is supposed to typify or stand for something. Many symbols are mere "survivals"—things which having no longer any utility continue to exist because we have inherited the tendency to make them; as funereal urns carved on memorial monuments. They were once real urns holding the ashes of the dead. We cannot stop making them, but we can give them a name which conceals our helplessness.*
—A. Bierce

2.3 OPERATIONALLY DEFINING THE VARIABLES OF ANTHROPOLOGY

Much of the literature of modern science is involved with describing conditions and outcomes of experimentation. As long as the scientist has properly reported his experiments, other investigators should be able to repeat the initial procedures and obtain similar results. Science is grounded in establishing the *repeatability of results*, and these procedures apply to some anthropological research, especially in genetic and dietary studies conducted by physical anthropologists. But as Pelto (1970:48–49) has pointed out, anthropological data are not always collected as the result of experimentation, but are collected frequently through the systematic observation of unusual, aperiodic events such as ceremonies, kinship interactions, subsistence practices, and even disasters such as floods or fires. When dealing with phenomena of this sort, anthropologists cannot provide for true repeatability of observation, no matter how well the field techniques are described. In fact, the very nature of some anthropological research—particularly in archaeology and hominid paleontology—involves the destruction of archaeological and paleontological sites during the process of data extraction. Pelto suggests that rather than attempt to refine experimental repeatability, anthropologists should address themselves to a more realistic proposition: "If another observer had been at the particular event, and if he used the same technique, would he have obtained the same results?" (Pelto 1970:49). The true test of adequate definition and technique in anthropology often involves *objectivity* rather than strict repeatability.

Some anthropologists, of course, still object to quantification and statistical manipulation of anthropological data in any form. Social phenomena are too complex, too subjective—the argument goes—to be approached in a "scientific" (that is, objective) manner. What these skeptics overlook is that the larger the errors involved, the more imperative become statistical methods. Statistics is often called the *science of variability*, so clearly the mere presence of error in no way vitiates use of statistical procedures. In fact, statistical methods were developed to meet the needs of those who must deal with imperfect data.

But all treatments of anthropological data, whether statistical or otherwise, can have no more validity than the basic definition of concepts. The most important criterion for adequate *operational definition* requires one to specify the procedures or processes through which data have been generated. Operations should be so specified that the same procedures can be repeated "in an unbiased manner by an intelligent person after a period of training" (Krumbein and Graybill 1965:69). But it is impractical (and, in fact, impossible) to define every term operationally, since there must always be certain "primitive" terms which remain undefined. Physicists have difficulty in defining absolutes such as time, length, and mass. But anthropological definitions can neatly sidestep such difficulties by simply taking given primitive terms and using them as undefined physical terms to build operational criteria relevant to anthropology (Harris 1964:3–6). Rather than consider at length the theory behind a good operational definition, let us examine some practical attempts by anthropologists to clarify their definitions operationally.

● *No member of a crew is praised for the rugged individuality of his rowing.*—R. Emerson

2.3.1 Operationally Defining Acculturation

In a study of rural Buganda in Uganda, Robbins and Pollnac (1969) attempted to establish the relationship between drinking patterns and acculturation. Not only is alcoholic consumption in Buganda considered to be an overt symptom of psychic disorder, but also the actual mechanisms involved in drinking behavior are thought to be indications of more far-reaching social changes. The fieldworker can readily define and observe alcoholic consumption, but the problem of operationally defining acculturation is a more elusive task. Robbins and Pollnac decided that "degree of acculturation" should be considered as two major aspects in the overall acculturation process: the self-identification of informants with Western society (as seen through the use of material items) and general exposure to Western behavior and values (through formal education). To achieve a rough approximation of the various acculturative factors operating within Bugandan society, the researchers devised a survey questionnaire involving a variety of economic and social topics. This questionnaire was administered to 109 randomly selected househoulds in six rural villages of Buganda.

From these results, Robbins and Pollnac abstracted criteria to distinguish traditional from acculturated households. The pilot study provided investigators with empirical evidence on Bugandan acculturation. The scale items consisted of 25 discrete variables, each readily observable by the ethnographer (Table 2.1). Each item is an indicator of westernization, and can be answered only by one of two possible responses: acculturational or traditional (this is why they are *discrete* variables). Ownership of common items of material culture—clocks, radios, stoves, and the like—indicates Western influence, while the relative degree to which the *kanzu* (native dress) was worn indicates traditional behavior. The sum of the acculturative responses thus proves an operational measure of westernization within any Bugandan household.

Using this scheme, independent workers should be able to scale any household from "highly acculturational" to "highly traditional" with a high degree of accuracy and repeatability. This ordinal scale (see Section 2.4.2) could in turn be compared to various aspects of drinking behavior (beverage preference, degree of alcoholic consumption, and various material aspects involved with drinking, such as bar furniture). Not only does the Robbins-Pollnac acculturation scale satisfactorily describe the processes used to rank households, but the items are also explicitly defined so the results can be repeated by other independent investigators.

2.3.2 Operationally Defining Projectile Point Attributes

The initial step in most archaeological analyses involves classifying the artifacts into rough categories. This preliminary classification is undertaken for several purposes: to condense the data, to establish time markers for dating the sites, to determine functional artifact types, perhaps even to reflect prehistoric "mental templates" reputed to exist in the mind of the maker. Regardless of the motive for classification, archaeological typology is always based upon the relatively fine-grained analysis and grouping of variables.

TABLE 2.1 **Acculturation Scale Items for Bugandan Households (after Robbins and Polinac 1969: table 1).**

1. The ability to read Luganda.
2. The ability to speak English.
3. The ability to read English.
4. Education of spouse, one or more years.
5. Education of spouse, four or more years.
6. Education of spouse, nine or more years.
7. Owns clock.
8. Owns watch.
9. Owns radio.
10. Owns iron.
11. Owns stove.
12. Respondent thinks it is proper for the husband and wife to eat at the same table.
13. Wearing *Kanzu* (native dress) at home with relatives and friends or when visiting relatives and friends (trad.).
14. Wearing *Kanzu* to work, to the local market, to towns and cities (trad.).
15. Wearing *Kanzu* all of the time (trad.).
16. Prefers drinking from a glass instead of gourd.
17. Has been to the bank to do business.
18. Goes to the cinema.
19. Purchases and reads magazines.
20. Visits Kampala (city).
21. Likes to straighten hair.
22. Presence of photographs on the inside walls.
23. Education of respondent one or more years.
24. Education of respondent four or more years.
25. Education of respondent nine or more years.

Sometimes these variables are quite crude and can be readily defined operationally: artifact length and width (as measured by vernier calipers), weight (as determined by a three-beam laboratory balance), color (as measured by a Munsell color chart), and so forth. But few archaeologists limit their analyses to such straightforward and easily defined attributes. Significant variables are more often rather subjective in nature, and observation becomes a matter of past experience rather than objectivity. Consider the projectile points in Fig. 2.2. Is artifact (a) basally indented? What about artifact (b)? Is projectile point (e) a basally notched point, or is it corner-notched? What about artifact (d)? Such questions cannot be answered objectively until we know how the variables are defined operationally. Such impressionistic variables are often defined by the naming process itself: basal indention means just that; corner-notched points appear just as the name implies; side-notched artifacts are notched from the side. Not only are these definitions circular, but they are also not operational because we are not told just how to determine the amount of basal indention or corner-notching in given artifacts.

Operational definitions must tell other researchers exactly how the variables are defined and measured:

$$\text{basal indention} \equiv \text{basal indention ratio} = \frac{\text{axial length}}{\text{total length}}$$

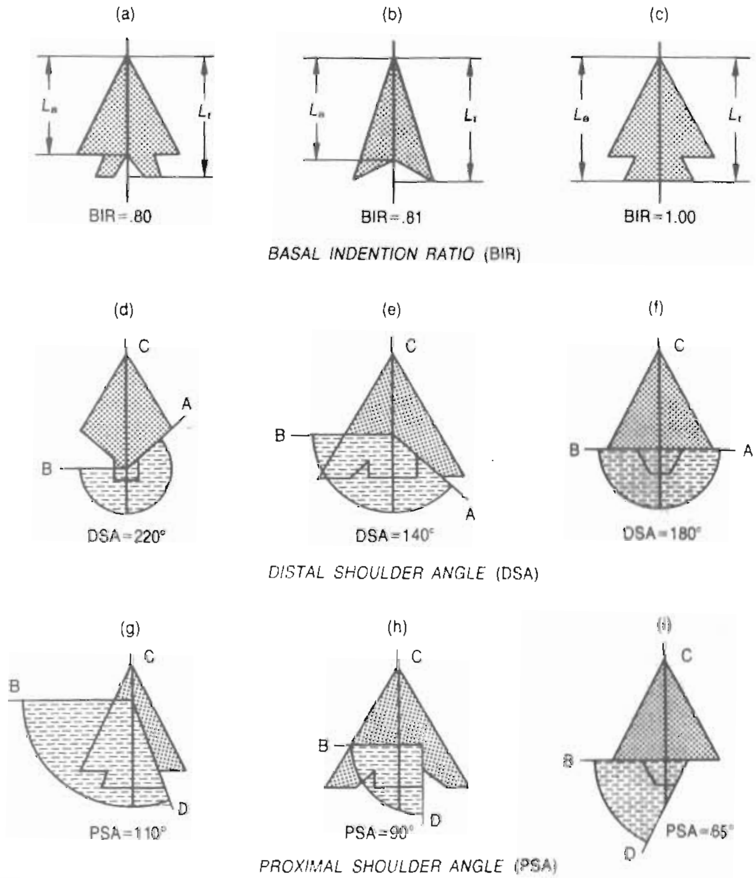


Fig. 2.2 Operational definitions of projectile point attributes (after Thomas 1970: fig. 2).

The variable "basal indention" is defined as the *ratio* (BIR) of two other continuous variables, "axial length" and "total length." "Total length" (L_t), in turn, is defined as the longest dimension of the projectile point. "Axial length" (L_a) is the distance along the line of symmetry (the midline). "Length" is a primitive term and hence is undefined. We'll let the physicists worry about that! Basal indention is thus operationally defined by two simple and relatively objective measurements. The measurement BIR = 1.00 [Fig. 2.2(c)] indicates that the point is not at all basally indented, but BIR = 0.80 [Fig. 2.2(a)] denotes a high degree of basal indention. The smaller the BIR, the more an artifact is basally indented. After a little practice, any novice can determine a BIR measurement as accurately as the most grizzled dirt archaeologist.

Notching position is a trickier attribute to define operationally. Rather than deal with position of notching directly, we must define a totally new variable, the Distal Shoulder Angle (DSA). In this case, DSA has not been assigned a written definition, but is defined graphically in Fig. 2.2. This case illustrates that good operational definitions need only be objective, and not necessarily stated in words. Discussing projectile point morphology in terms of such operational variables obviates the use of vague categories such as "corner-notching," "side-notching," and so forth, since we now have more precise methods of expression. Furthermore, DSA is on the interval scale of measurement (defined in Section 2.4.3) and is amenable to more sophisticated quantitative analysis than are mere categories.

2.3.3 Operationally Defining Cranial Measurements

Nowhere in anthropology has the operational definition of variables progressed so far as in craniometry—the measurement and analysis of human skulls. In fact, there are even those who suggest that craniometry has progressed *too far* (for example, Boyd 1950) and that excessive zeal in measuring human skulls has unduly overshadowed other lines of study in physical anthropology. Perhaps this is true, for surely no portion of mammalian anatomy has been more extensively probed, poked, X-rayed, and photographed than the human skull.

The recording of cranial measurements seems to date back to Petrus Camper, an eighteenth-century Dutch anthropologist who measured the degree of facial angle in order to determine the amount of prognathism of the human skull (Hulse 1963:427). But Camper's work and most subsequent attempts were rather clumsy approximations and quite lacking in precision. The problem of establishing objective cranial measurements was discussed in 1884 at the Congress of the International Anthropological Association in Frankfurt. A standardized plane of reference for craniometry, the so-called Frankfurt Horizontal (FH), was established as the plane defined by the left and right poria, and the right orbitale (see Fig. 2.1). Since that time, the phrase "measured in the Frankfurt horizontal" has assured that common cranial measurements are indeed comparable.

But the definition of the Frankfurt horizontal is itself dependent upon the objective placement of the two cranial reference points: the porion and the orbitale. Krogman has operationally defined these terms as follows (1962: 316–317):

orbitale: The lowest point on the lower margin of the orbit.
porion: Most lateral point on the roof of the external auditory meatus (bony ear hole).

Both definitions contain primitive terms ("lowest point," "lower margin," "most lateral point") which must themselves be undefined. Figure 2.1 presents these and other cranial reference points in common use by modern physical anthropologists.

2.4 LEVELS OF MEASUREMENT

If anthropologists measured only skulls, then almost all relevant variables could be expressed in metric units such as millimeters, cubic centimeters, or degrees of arc. But the study of mankind embraces a range of topics and specialized subject matter, and such purity of measurement does not exist. In fact, most observations in social science are not measurements at all, strictly speaking, but are rather counts or ranked orderings. In order to cope adequately with the diversity of anthropological data, it is necessary to consider the common scales of measurement: their definition, their properties, and their restrictions.

2.4.1 Nominal Scale

The most elementary measuring operation involves the sorting of individual objects into homogeneous categories. This procedure (*classification*) is a necessary first step in nearly all social, biological or physical sciences, and many branches of science—especially the social sciences—have yet to evolve beyond the stage of primitive item classification. Although classification must never be treated as an end in itself, few worthwhile projects in science can proceed until individual phenomena can be treated in operational classes rather than as mere raw variates.

In nominal scales, symbols are assigned to categories which represent the range of possible values any given variate might take. These symbols are often words, such as "male" and "female" or "left" and "right," but other symbols such as numbers, pictures, colors, or even simple signs (+, −, and *) can label nominal categories. Creating a nominal scale involves merely assigning symbols to categories, subject to only one rule: Do not assign the same symbol to different categories, or different symbols to the same category (Stevens 1951). In a more formal sense, a nominal scale requires only that the classification be *exhaustive* (classify all possible items in the array) and *mutually exclusive* (classify each item into only one category). Beyond these elementary restrictions, anything goes on the nominal scale.

The nominal scale is anthropology's most primitive form of measurement, but this simplicity must not be allowed to obscure the practical difficulties and pitfalls involved in applying the nominal scale to real data. Numerals, for instance, are sometimes used to symbolize a set of nominal classes. The mere application of a numeral to a class, however, in no way justifies the use of ordinary arithmetic operations on that scale. It is arithmetically possible, for example, to add several license plate numbers, but the outcome is logically absurd. The arithmetic manipulation of any nominal scale, regardless of the symbol, is meaningless because nominal scales reflect only differences in *kind*, not of *degree*.

Some variables can be characterized in a number of different ways—that is, on several different nominal scales. "Blood type," for instance, has been observed on literally dozens of operational scales: ABO system, MN system, P system, Lutheran system, Lewis system, Duffy system, Kidd system, and Diego system, to name but a few. Each scheme defines its own categories and each involves a different nominal scale. These scales may or may not be independent. Which scale is selected usually depends upon a variety of factors such as the exact objectives of research, feasibility, cost, and previous experience. But once a given scale has been selected, the operational definitions of a successful classification divide all possible variates into exhaustive and mutually exclusive categories. Table 2.2 presents several examples of other nominal scales commonly used in modern anthropology.

TABLE 2.2 Some Common Nominal Scales in Anthropology.

Variable	Symbol	Common Operational Definition
"Regional identification" (col. 1, *Ethnographic Atlas*)	A	Africa, exclusive of Madagascar and the northern and northeastern portions of the continent.
	C	Circum-Mediterranean, including Europe, Turkey, and the Caucasus, the Semitic Near East, and northern and northeastern Africa.
	E	East Eurasia, excluding Formosa, the Philippines, Indonesia, and the area assigned to the Circum-Mediterranean but including Madagascar and other islands in the Indian Ocean.
	I	Insular Pacific, embracing all of Oceania as well as areas like Australia, Indonesia, Formosa, and the Philippines that are not always included therewith.
	N	North America, including the indigenous societies of this continent as far south as the Isthmus of Tehuantepec.
	S	South America, including the Antilles, Yucatan, and Central America as well as the continent itself.
"Slavery" (col. 71, *Ethnographic Atlas*)	H	Hereditary slavery present and of at least modest social significance.
	I	Incipient or nonhereditary slavery, that is, where slave status is temporary and not transmitted to the children of slaves.
	O	Absence or near absence of slavery.
	S	Slavery reported but not identified as hereditary or nonhereditary.
"Ground plan of dwelling" (col. 80, *Ethnographic Atlas*)	C	Circular.
	E	Elliptical or elongated with rounded ends.
	P	Polygonal.
	Q	Quadrangular around (or partially around) an interior court.
	R	Rectangular or square.
	S	Semicircular.

"Blood group (ABO system)"	A	Antigen A present
	B	Antigen B present
	AB	Antigens A and B present
	O	Neither antigens A nor B present
"Epicanthic fold"	Present	". . . consists of a fold of skin that when the eye is open, comes down over and runs on a line with the edge of the upper eyelid" (Kelso 1970:253–254).
	Absent	—
"European Mousterian" (after Bordes 1968: chapter 8)	*Typical Mousterian*	Handaxes rare, percentage of scrapers between 25–50%; points well developed and carefully worked; limaces (double-pointed scrapers retouched all around) rare; few backed knives; notched flakes and denticulates; relatively small percentage of Levalloisian technique.
	Quina-Ferrassie Mousterian	50–80% scrapers; few or no handaxes; no backed knives; relatively few denticulates; non-Levalloisian technique.
	Denticulate Mousterian	No typical handaxes; no typical backed knives; few or no points; 5–25% scrapers (of poor quality); 35–55% denticulated tools; notched flakes common; technique may or may not be Levalloisian.
	Mousterian of Acheulean tradition, Type A:	8–40% handaxes; 20–40% scrapers; prevalence of points; some notched flakes; fairly large number of denticulates; some tools of Upper Paleolithic type; backed knives.
	Type B:	2–8% handaxes; great development of denticulates and backed knives; occasional double burins.

As mentioned earlier, operational definitions vary in both precision and accuracy. The ABO blood-grouping system is virtually infallible, since classification depends strictly upon the comparison of the antigen-antibody reaction of a drop of blood. "Sex" is another nominal scale which is relatively clear-cut, although there are a few borderline cases. But dichotomizing a lithic tool assemblage into "core" and "flake" tools, or estimating the predominant familial organization within a given society can introduce a high degree of personal judgment and intuition. Obviously, such cases involve nominal scales which are neither totally exhaustive nor completely mutually exclusive. But these difficulties arise as a result of inadequate definition of categories (a strictly anthropological matter) rather than from any inherent insufficiency in the mathematical grounding of nominal scales.

2.4.2 Ordinal Scale

An *ordinal* (or *ranked*) scale involves a relevant ordering of discrete categories into a meaningful sequence, obviously a significant logical advance over mere classification on the nominal level. Ordinal categories rank classes along a continuum, but the distance between each category in the continuum is either unknown or undefined because of imprecise measurement techniques or some quality inherent in the variable. It is impossible to specify *how much* of a variable each ordinal category represents.

Ordinal scales possess all properties of nominal scales, but have the additional property of *asymmetry*. That is, if the ordinal relations indicate that $A > B$ and $B > C$, then it must follow that $A > C$. In the case of ethnographic settlement pattern (Table 2.3), we know that since B (fully migratory) is less sedentary than H (separated hamlets), and since H is less sedentary than N (neighborhoods of dispersed homesteads), then B must also be less sedentary than N. But note that there is no indication in this—or in any other—ordinal scale regarding the *magnitude* of difference between categories. "Separated hamlets" are not necessarily twice as sedentary as "fully migratory," any more than "separated hamlets" are half as sedentary as "neighborhoods of dispersed homesteads." Because equal distances between ordinal categories can never be assumed, it is improper to add two ordinal scores, or to attempt to take an average of an ordinal scale. A large body of "rank-order statistics" has evolved to handle problems on the ordinal scale (as discussed in Chapters 12 and 14).

The line between nominal and ordinal scales occasionally becomes fuzzy and indistinct. Peruvian archaeologists, for example, sometimes classify their sites as "ceramic" or "preceramic." This distinction can be applied as a simple descriptive label, without necessary implications of chronological priority. This is a nominal scale. But when presence of pottery is used to seriate sites into a temporal sequence, then a rank ordering is definitely implied and the scale becomes ordinal. These distinctions are discussed in more detail when nonparametric statistics are considered.

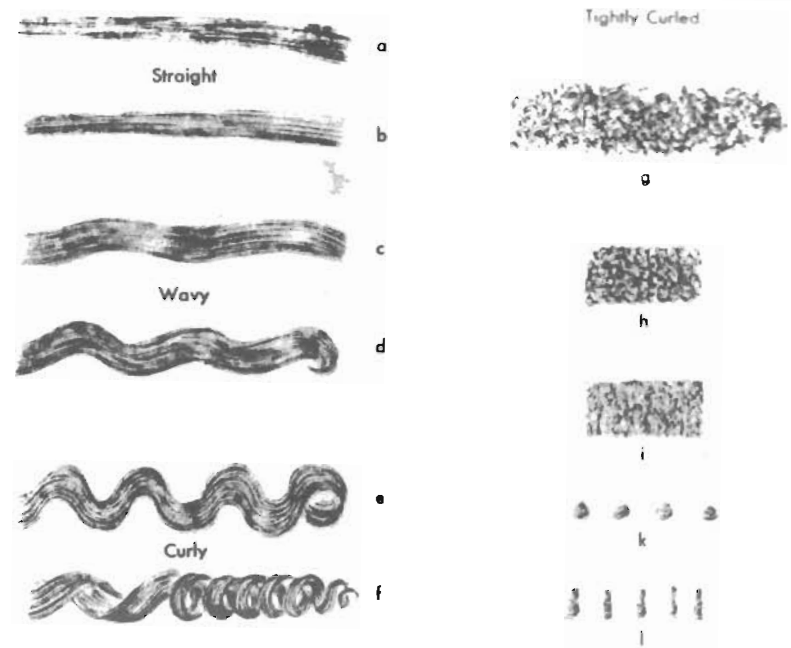
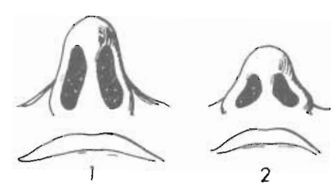
2.4.3 Interval Scale

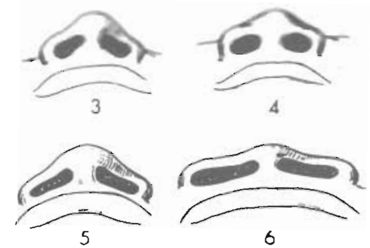
A scale of measurement is termed *interval* if it possesses all the properties of an ordinal scale but also implies *equal distances between the symbols*. A weather-

TABLE 2.3 Some Common Ordinal Scales in Anthropology

Variable	Symbol	Common Operational Definition
"Settlement pattern" (col. 30, *Ethnographic Atlas*)	B	Fully migratory or nomadic bands.
	H	Separated hamlets where several such form a more or less permanent single community.
	N	Neighborhoods of dispersed family homesteads.
	S	Seminomadic communities whose members wander in bands for at least half the year but occupy a fixed settlement at some season or seasons, for example, recurrently occupied winter quarters.
	T	Semisedentary communities whose members shift from one to another fixed settlement at different seasons or who occupy more or less permanently a single settlement from which a substantial proportion of the population departs seasonally to occupy shifting camps, for example, during transhumance.
	V	Compact and relatively permanent settlements, that is, nucleated villages or towns.
	W	Compact but impermanent settlements, that is, villages whose location is shifted every few years.
	X	Complex settlements consisting of a nucleated village or town with outlying homesteads or satellite hamlets. Urban aggregations of population are not separately indicated, since column 31 deals with community size.
"Male genital mutilations" (col. 37, *Ethnographic Atlas*)	0	Absent or not generally practiced.
	1	Performed shortly after birth, that is, within the first two months.
	2	Performed during infancy, that is, from two months to two years of age.
	3	Performed during early childhood, that is, from two to five years of age.
	4	Performed during late childhood, that is, from six to ten years of age.
	5	Performed during adolescence, that is, from 11 to 15 years of age.
	6	Performed during early adulthood, that is, from 16 to 25 years of age.
	7	Performed during maturity, that is, from 25 to 50 years of age.
	8	Performed in old age, that is, after 50 years of age.

Table 2.3 (*cont'd*)

Variable	Symbol	Common Operational Definition
"Human hair form" (after Kelso 1970:253)		
"Human nasal flatness" (after Kelso 1970:245)		



"Community patterning" (Beardsley 1956: 135-146)	Free wandering	Remains should represent initial occupation of region; wide distribution of artifact types that are simple in nature and limited in variety; camps which moved frequently, evidence of occupation undetectable or at best scanty; consisting of ash or charcoal; split or burned animals bones and a few artifacts.
	Restricted wandering	More limited distribution of characteristic artifact types; more specialized tools; food-grinding implements and stone-lined hearths; awls and needles (reflect basketry making and skin working).
	Central-base wandering	Indications of seasonal occupation; thicker midden accumulations than above; shell middens were used; cave occupations common; occasional graves unusually rich (shaman?).
	Semipermanent sedentary	Refuse deposits thin; potsherds common; house structures patterned in relation to one another; isolated graves in village; or burial may be in cemetery; evidence of ceremonial activity discernable; storage or cache pits common.
	Simple nuclear centered	Ceremonial structures (earthworks) long-term habitation refuse; pottery differentiation into utilitarian and ceremonial wares; cemeteries have a few graves of very high status; stone sculpture; mosaics; engraved shell common; house structures of adobe or stone have patterned arrangement; ceremonial areas show planning.
	Advanced nuclear centered	Contrast between remains at "capital" and relative simplicity of satellite communities; public buildings show planned construction and uniformity of architecture; crafts indicate occupational specialization; head deformation and tooth mutilation common in upper class.
	Supranuclear integrated	Presence of roads between centers; extensive irrigation works; evidence of conquest; construction of forts in architectural style of conquerer at strategic places.

man who describes the daily temperature in New York City as "cold, cool, mild, or hot" is operating on an ordinal scale, since the distances between the categories are not specified. But if these identical temperatures were expressed in degrees Fahrenheit, the scale would be interval because each symbol (degree) represents exactly the same temperature interval as every other degree. We could say, for instance, that the difference in temperature between 70 and 75 F is exactly equal to the difference between 95 and 100 F. In the strict sense, interval scales are the first truly *quantitative* measures we have discussed—ordinal and nominal scales are generally considered *qualitative*—and most statistical and arithmetic manipulations are applicable to interval variates. Interval scale variates may be added or subtracted, and the addition of a constant does not change the internal relationships of such scales.

Few interval level measures are more intrinsically interesting than the Maya calendrical system. The Maya calendar crops up several times in this text; not only does the system have chronological significance but it also has arithmetic properties that are common to all interval scales. The basic element of the Maya calendar is the day (*kin*), and days can be grouped into several larger temporal divisions:

20 *kins* = 1 *uinal*, or 20 days
18 *uinals* = 1 *tun*, or 360 days
20 *tuns* = 1 *katun*, or 7200 days
20 *katuns* = 1 *baktun*, or 144,000 days

Groupings such as these are impossible on both the nominal and the ordinal scales because exact intervals are required between each basic category (see Coe 1966:chapter 3 for more details of the Maya calendar).

This additive property enabled the Maya to construct the systems of "Long Counts," which placed an event within the span of historic time. Long Count dates were commonly inscribed on monuments to commemorate great political events, to dedicate temples, or to celebrate a particularly important military episode. Long Count dates are obviously invaluable to archaeologists studying Maya cultural history.

When translated, all Maya Long Count dates read "*X* days since the end of the last Great Cycle," somewhat like the Christian calendar which records time as "*X* years since the birth of Christ." A system of numerical notation has been devised by Mayan epigraphers to express the Maya dates in more comprehensible form. The Long Count date of 9.10.19.5.11, for example, can be interpreted as follows (example from Coe 1966:58):

9 *baktuns*	=	1,296,000 days	
10 *katuns*	=	72,000 days	
19 *tuns*	=	6,840 days	
5 *uinals*	=	100 days	
11 *kins*	=	11 days	
		1,374,951 days	(since the end of the last Great Cycle)

In other words, the Maya calendar works because the constant intervals (the *kins*) can be grouped *through addition* into precisely equivalent larger units. This property does not hold for ordinal or nominal scales.

Interval scales can also be meaningfully manipulated by subtraction. Suppose, for example, that two monuments were found at a particular Maya site: Monument A had the above date (9.10.19.5.11 = 1,374,951 days since the end of the last Great Cycle) and stela B had a date of 9.10.12.2.3 (1,372,363 days since the end of the last Great Cycle). Since both counts date from the same zero point, and since we know that interval variates can be subtracted from one another, it becomes an easy matter to determine that the two stelae were erected exactly 1,374,951 − 1,372,363 = 2,588 days apart.

But note that the zero point in this, and in all, interval scales is arbitrary. Maya Long Count dates are expressed in "days since the end of the last Great Cycle." In fact, the zero point of the Long Count is so arbitrary that epigraphers and archaeologists labored for years trying to pin down that elusive date. The search was narrowed to a series of discrete time intervals, however, since any given *katun* can recur only once every 260 Maya years. As a result, it was possible for scholars to correlate any given Long Count date with only a few intervals along the Christian calendar, depending upon which *katun* was selected for the zero point. According to a correlation by George Spinden, the zero year corresponded to 3373 B.C. on the Christian calendar, while a second hypothesis, the Goodman–Thompson–Martinez correlation, set the magic year at 3113 B.C. Any Maya date could be converted to Christian years by either correlation, and the Spinden correlation is always 260 years older than the dating by the Goodman–Thompson–Martinez correlation.[1] Incidentally, radiocarbon evidence now strongly supports the Goodman–Thompson–Martinez correlation, by methods discussed in Chapter 7.

The lack of a true zero point restricts the utility of interval scales in some respects. Since zero is arbitrarily assigned, the *ratios* of two interval variates cannot be meaningfully compared. Consider the statement "80°F is twice as hot as 40°F." In effect, this statement implies the following ratio:

$$80°\text{F}:40°\text{F} = 2:1$$

Because the zero point on the Fahrenheit scale has been assigned by convention only, the ratio of these two variates is a logical absurdity. So, for that matter, is the assertion that A.D. 975 is twice as old as A.D. 1950, or that today is twice as hot (half as cold?) as yesterday. Only when zero points are dictated by the phenomena themselves—as is the case of ratio scale variates—can comparisons such as "twice as," "half as," or "three times as large as" be meaningful.

2.4.4 Ratio Scale

The ratio scale is the most advanced counting system considered here, and ratio variates are still rare, unfortunately, in most of anthropological research. What sets the ratio variates above those of the interval measurement scale is that the

[1] Note that correcting from the Spinden to the Goodman–Thompson–Martinez date is precisely the same logical operation as converting daily temperature from degrees Celsius to degrees Fahrenheit by the formula

$$F = (9/5)C + 32$$

There is nothing intrinsic about the zero point in any interval scale, whether it be the Long Count, an IQ score, height above sea level, or years before present (B.P.) which archaeologists arbitrarily take to be A.D. 1950.

zero point of a ratio scale is fixed rather than arbitrarily assigned. The most common ratio scales in anthropology involve the quantitative expression of physical properties such as length, width, size, weight, and so forth: Body stature, projectile point width, cranial capacity, basion–nasion diameter, mean distance travelled per year, and average weight are all ratio variables.

Anthropology also makes use of ratio scales in the enumeration of cases: the minimum settlement size, the number of cervical vertebrae, the number of same-sexed siblings, the number of storage cists in a habitation cave. While the counts themselves are discrete variates, they form a sophisticated set of measurements; not only are the counts exact—rather than approximations, as with continuous variates—but also a zero point is always implicit in the enumeration process.

Ratio variates can also be derived from other primary variates. Population density, for instance, is generally defined as

$$\text{population density} = \frac{\text{number of people}}{\text{unit area}}$$

An estimate of 500 people per square mile is one such derived ratio variable, in which the numerator and denominator are themselves ratio levels. Many common demographic indices—mortality, fecundity, rate of immigration, intrinsic rate of increase—are derived ratio variates, as are the ratios common to physical anthropology (cephalic index, nasal index, metabolic rate) and also to archaeology (the room-to-kiva ratio, the number of beta emissions per 1000 minutes in radiocarbon dating, the ratio of domesticated to nondomesticated foodstuffs).

Ratio variates admit a wide range of mathematical properties because ratio scales are totally *isomorphic* to arithmetic (meaning that, since the structures are identical, all arithmetic operations can be performed upon ratio variates without destroying the relationships among the variates). As the name implies, ratio scales can also be meaningfully compared as ratios: A population density of 50 people per square mile is exactly twice that of a city with only 25 people per square mile; a 5.0 gram projectile point has only one-third the mass of a 15.0 gram point; a nuclear family of four individuals is exactly half the size of a family of eight. In addition, unlike interval scaling, ratio level variates can be transformed through multiplication by a constant. Inches are readily converted by multiplication to feet, for example, by using the correction factor of $1/12 = 0.083$.

Table 2.4 summarizes the arithmetic operations permissible for each level of measurement.

TABLE 2.4 Permissible Operations for Measurement Scales.

Scale	$=, \neq$	$>, <$	$+, -$	$\times, \div$
Nominal	Yes	No	No	No
Ordinal	Yes	Yes	No	No
Interval	Yes	Yes	Yes	No
Ratio	Yes	Yes	Yes	Yes

2.4.5 Refining Levels of Measurement

The preceding sections have characterized measurement scales into four basic (ordinal) categories. At this point, one might justifiably question: So what? Was this merely another academic exercise in number games and classification, or do the levels of measurement have some practical worth?

Two reasons justify our consideration of the various levels of measurement. First of all, these levels of observations dictate to some degree the statistical operations which can be applied to particular sets of data. Nominal measures are often termed "weak" because few arithmetic operations are applicable to simple unordered categories. Ratio scales, on the other hand, are "strong" because these measures are applicable to all mathematical operations. Every statistical test makes certain explicit assumptions about the underlying structure of the data and generally requires some minimal level of measurement. All statistical procedures to be considered here can be legitimately applied when strong measurement is available, but weak measurements severely restrict the potential avenues of analysis. The relationship between scale of measurement and statistical test is a complex topic and not so straightforward as social scientists once thought. It is sufficient to recognize at this point that the kinds of measurements available will influence the mode of analysis.

Levels of measurement can also be taken as a rough gauge of scientific maturity within a given discipline. Physics, for instance, is generally considered to be a most sophisticated science, and nearly all physical measurements are "strong." Social science measurements, on the other hand, are often only nominal and ordinal, indicative of a more primitive state of investigation. As Kemeny (1959:143) has pointed out, many sciences (and especially the social sciences) have yet to pass beyond the stage of preliminary classification. Students in introductory anthropology courses, for example, often complain about the strange names and categories which they are expected to assimilate: Paleolithic, Mesolithic, cross-cousin marriage, couvade, morpheme, australopithecine, cognatic, animistic, moiety. While, admittedly, learning these elementary categories can be tedious for the novice, the procedure is essential. Science usually begins with classification in one form or another, and the fact that a young science can characterize two objects as sharing a single variable state is a significant milestone. That is, when two objects are judged to be the "same" with respect to a variable, the conclusion is a primitive form of generalization, and a unifying thread throughout science is the consistent search for more encompassing generalizations. To say, for example, that the Nisenan Indians of central California have the "same" cousin kinship terminology (the Hawaiian system) as the Blackfoot of the Plains is an important step toward understanding kinship systems in general.

A second important signpost of scientific maturity is the introduction of orders of magnitude. Not only may two objects be judged "equal" with respect to variable X, but ordinal scales allow the additional judgment that some categories contain *more* of variable X than do other categories. A simple order is sufficient for many purposes, but a higher level of measurement is generally preferred so that the more advanced mathematical techniques are available. The more sophisticated the measurement system, the more precise will be the theories which result.

But the real problem in developing adequate systems of measurement lies outside the scope of mathematics and statistics. Statistics deals only with symbols, and statistical methods work equally well regardless of the referents which these symbols represent (see Hays 1973:89). Progress in measuring systems requires a deeper understanding of anthropological phenomena rather than more sophisticated mathematical statistics. Once the variables have been defined on solid anthropological grounds, then appropriate manipulative techniques are readily available from established mathematics. Some examples should illustrate how measurements progress in anthropology.

Anthropologists have faced few more truculent problems than in their efforts to measure *cultural evolution*. In fact, the effectiveness of measuring cultural development reflects in large part the maturation of anthropological inquiry. Most primitive sciences commence substantive research with a lengthy period of name giving, and these early classifications are generally nominal scales. But the scientific study of cultural evolution has progressed somewhat differently because the name-giving phase began directly with the ordinal level and bypassed the nominal phase altogether. By its very nature, the study of cultural evolution is concerned with the sequences of events. Ordering was implied right from the start. Even when discussion involved simplistic dichotomies—such as civilized versus noncivilized or hunter versus farmer—an ordering was always implied. The categories of cultural evolution were never nominal.[2]

Literally dozens of ordinal scales have been suggested to measure the cultural progress of man. One notable effort was Condorcet's ten-stage scheme in the *Outline of Intellectual Progress of Mankind*, originally published in 1795:

1st Tribal society.
2d Pastoral society.
3d Agricultural society to the invention of the alphabet.
4th The progress of the human mind in Greece up to the division of the sciences about the time of Alexander the Great.
5th The progress of the sciences from their division to their decline.
6th The decadence of knowledge to the restoration about the time of the Crusades.
7th The early progress of science from its revival in the West to the invention of printing.
8th From the invention of printing to the time when philosophy and the sciences shook off the yoke of authority.
9th From Descartes to the foundation of the French Republic.
10th The French Republic.

Not only are Condorcet's categories impressionistic, but they are also ethnocentric, dealing largely with European history and ignoring cultural evolution throughout the rest of the world. Marvin Harris (1968:35) has succinctly characterized Condorcet's perspective as "the more remote the age, the duller

[2]Although the approaches to the study of cultural evolution have varied enormously over the past century, the basic definition of cultural evolution has changed surprisingly little. Robert Carneiro (1973:90) recently defined evolution as a "change from a relatively indefinite, incoherent homogeneity to a relatively definite, coherent heterogeneity, through successive differentiations and integrations." The essence of this statement is little modified from the 1862 definition by Herbert Spencer.

the mind, the less enlightened is man's social life." But regardless of the shortcomings, Condorcet's scheme illustrates the principles of ordinal scaling.

Refined ordinal classification was later framed by Lewis Henry Morgan in *Ancient Society* (1877). Morgan divided the progress of human achievement into three major "Ethnical Periods": savagery, barbarism, and civilization, which were scaled to seven categories according to status as follows:

I. *Lower Status of Savagery:* commenced with the infancy of the human race in restricted habitats, subsistence upon fruits and nuts. No such tribes remained into the historical period.
II. *Middle Status of Savagery:* commenced with acquisition of fish and use of fire. Mankind spread over greater portion of earth's surface. Exemplified by Australians and Polynesians.
III. *Upper Status of Savagery:* commenced with invention of the bow and arrow. Exemplified by Athapascan tribes of Hudson's Bay Territory.
IV. *Lower Status of Barbarism:* commenced with invention or practice of pottery. Exemplified by the Indian tribes of the United States east of Missouri River.
V. *Middle Status of Barbarism:* commenced with domestication of animals in the Eastern hemisphere, and in the Western with cultivation by irrigation and use of adobe brick and stone in architecture. Exemplified by villages of New Mexico and Mexico.
VI. *Upper Status of Barbarism:* commenced with manufacture of iron. Exemplified by Grecian tribes of the Homeric Age and Germanic tribes of the time of Caesar.
VII. *Status of Civilization:* commenced with use of a phonetic alphabet and production of literary records; divided into *Ancient* and *Modern*.

The ordinal stages of Condorcet and Morgan can, of course, be faulted on several scores, but these early classifications were important steps in sharpening the perception of cultural evolution.

After decades of strong disfavor, the definition and study of evolutionary stages has again become the subject of legitimate anthropological research. Service (1962), for example, defined three stages of primitive human social organization—bands, tribes, and chiefdoms—and some archaeologists, such as Sanders and Marino (1970), have directly adopted Service's sociocultural stages in interpreting the archaeological record. Other archaeologists have attempted to apply schemes of stages based upon technological criteria (such as the Beardsley classification presented in Table 2.3). In fact, most archaeological research is from time to time synthesized using similar ordinal frameworks, as in the Willey and Phillips' (1958) scheme of "historical-developmental stages" of New World prehistory: Lithic, Archaic, Formative, Classic, and Postclassic. Yet, regardless of sophistication (or lack of it), all such scales remain subject to the limits of ordinal level measurements.

In 1948, Carleton Coon prepared a six-part classification of cultural development, basing his stages upon four quantitative measures: number of specialists, amount of trade, number of institutions to which an individual may belong, and the complexity of those institutions (Coon 1948:612f). In other words, Coon bridged the gap between ordinal and ratio level variates. Somewhat later, Raoul

Naroll (1956) furthered this line of investigation by devising a single index of cultural evolution. Naroll's "*Preliminary Index of Social Development*" is based upon three equally weighted and operationally defined indicators: craft specialization, organization ramification, and urbanization. The 1960s saw an important series of related studies in which literally hundreds of traits were considered for use as measures of cultural complexity (see Tatje and Naroll, 1970, for a comprehensive discussion of these indices). The more recent studies indicate that several of these indices are highly interrelated and often produce nearly identical results. The single best indicator of cultural evolution seems to be the *maximum settlement size* variable, defined earlier (Section 2.2).

The point is that when substantive research began on cultural evolution, the level of measurement consisted of relatively crude—and usually ethnocentric—ordinal scales such as Morgan's "Ethnical Periods." Stages of evolution were redefined on several occasions, sometimes for rather specialized objectives and at other times to provide more precision, but they remained on the ordinal scale and were therefore subject to the restrictions which apply to all rank-order variables. Finally, a second line of metric investigation led to more sophisticated measurements: Naroll, Carneiro, and others were able to derive interval and ratio scale indices to characterize cultural evolution. It was ultimately discovered—largely as a result of attempts to redefine such measures—that a single ratio scale indicator, the maximum settlement size, is sufficient to adequately characterize cultural complexity throughout the world. This discovery is of particular significance to archaeologists, since settlement size can often be estimated from the maximum floor areas of archaeological sites by using Naroll's constant; thus, much of the prehistoric record can be applied to the study of cultural evolution. The scientific investigation of cultural evolution progressed hand in hand with the progressive definition of more adequate operational indices, which in turn raised the overall levels of measurement.

Similar refinements in measuring technique have occurred in physical anthropology. One prime example is the measurement of the PTC taste-deficiency trait. In 1931, A. L. Fox observed quite by chance that while some individuals were unable to taste the synthetic compound phenylthiocarbamide (PTC), others reported the taste as quite bitter, somewhat like quinine. Subsequent investigation revealed that the ability to taste PTC is inherited as a simple Mendelian dominant. The test was initially administered by instructing informants to directly ingest crystals of PTC or to touch the tongue with a PTC-impregnated paper strip. The results of this procedure were based upon informant response, and subjects were characterized as "taster" or "nontaster," two dichotomous classes on the nominal scale. The test was later refined by administering to informants a series of diluted PTC solutions, and then asking them to describe their sensations. While this test was largely subjective—determined by the informant's ability to articulate his sensations—the refined test indicated a range of sensitivity among tasters which could be expressed in terms of crude rank orders of tasters along a scale of sensitivity.

The test was even further improved when informants were given a series of unlabelled tumblers, half of which contained a PTC solution and half of which contained just water. Subjects were then asked to discriminate the bottles containing PTC. If the groups were successfully sorted, then the next lowest concentration was used. The lowest concentration at which a completely correct answer was given can be operationally defined as the tasting threshold. In order to standardize the results, the following standardized solutions were applied (after Harris and Kalmus 1950:table 1):

Solution No.	PTC, mg/liter
1	1300.00
2	650.00
3	325.00
4	162.50
5	81.25
6	40.63
7	20.31
8	10.16
9	5.08
10	2.54
11	1.27
12	0.63
13	0.32
14	0.16

In this manner, taster threshold is characterized by the solution number, a ratio-level variable. This test is repeated over a large series of subjects in the same biological population, producing a bar graph which characterizes a population's thresholds (see Fig. 2.3) in percentage. The horizontal scale refers to solution number. Such graphs frequently exhibit two peaks (or *modes*) distinguishing tasters from nontasters.

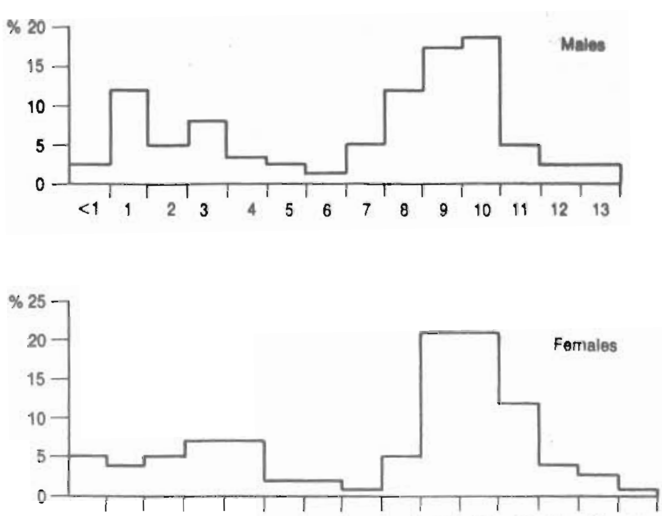


Fig. 2.3 PTC thresholds for 100 females and 114 males between the ages of 20 and 59 (after Harris and Kalmus 1950:fig. 3).

The case of PTC tasting is an excellent illustration of the principle that progress in science is generally paralleled by an increasing refinement in the level of measurement. Similar progress has occurred in the study of skin and hair pigmentation, color blindness, and nutritional ecology, to mention but a few areas of ongoing research in physical anthropology.

2.5 POPULATION AND SAMPLES

● *It is like the North Carolina mule which a tourist once stopped to admire. He asked the mule's owner what the animal's name was. The farmer replied, "I don't know, but we call it Bill."*
—S. Ervin

In Winick's *Dictionary of Anthropology*, the word "population" is defined twice:

"*population, amphimictic*. A population which has freely crossing fertile descendants."
"*population, inbreeding*. A group of persons among whom mating takes place." (Winick 1966:428)

Ecological populations are groups of living organisms of a single species found in a circumscribed area at a given time. This definition of population is frequently used by physical anthropologists with reference to both living and extinct animals. In this sense, one can study the chimpanzee population of the Gombe Stream reserve in Tanzania, the hunting behavior of *Australopithecus* populations in South Africa, or the gene frequency of the sickle-cell trait in Afro-American populations. Cultural anthropologists also use "population" in this sense. Human *society*, for instance, has been defined by Marvin Harris as a "population that has an organized way of life" (1971:654); societies are groups of people (populations) who depend upon each other for survival. Similarly, archaeologists often discuss the "prehistoric Pueblo III populations" or the "nomadic population of Clovis hunters," and a linguist might refer to his informants collectively as "the Shoshoni-speaking population." In general, anthropologists tend to use the term "population" in a phenomenological sense, denoting a concrete set of living (or once-living) individuals. All populations of this sort could be physically delimited under ideal conditions, enumerated, and assembled in one place.

The word "population" takes on a rather different meaning in statistics, but this difference is subtle and could easily be overlooked. A statistical population does not consist of physical objects at all (people, lemurs, or microbes), but rather of *variates* measured upon those objects. Arikara Indians could comprise a biological or sociocultural population, but they could never be a statistical population. Only a set of related variates—such as the stature of Arikara Indians—could be a statistical population. Body weight, daily caloric intake, marriage to cross-cousins, presence of the Rh blood factor, frequency of reflexive verb usage, and cranial capacity are all variates comprising statistical populations characteristic of the aggregate of Arikara Indians.

A *statistical population* is a set of variates (counts, measurements, or characteristics) about which relevant inquiries are to be made.

Unlike the phenomenological populations of biology and anthropology, statistical populations are *ideational*. Statistical populations include only observations; and observations have in themselves no objective reality. Statistical populations are comprised of variates, not objects.

Populations in this sense are arbitrary and must be carefully (that is, operationally) defined. Some populations may consist of a finite number of variates, such as the stature of all living Arikara Indians. Populations could also be defined to include not only living Arikara, but also all Arikara who lived in the past, and also those who will ever live in the future. So populations could also be infinite. Clearly, there is nothing "natural" about a statistical population because it is defined to satisfy specific research objectives. All these definitions, however, must be based upon *anthropological* rather than *statistical* considerations. Statistical methodology comes into play only after the relevant populations have been defined to suit the research objectives at hand.

It would be onerous indeed for an anthropologist to attempt to interrogate, measure, observe, or photograph the entire physical population of all living Arikara. And, if the statistical population had been defined to include Arikara of all times and all places, complete observation would be patently impossible. One characteristic of most statistical populations is that they are *incompletely observable*. Physical anthropologists can never hope to measure the entire population of *Australopithecus robustus* crania, for example, and in fact, primatologists have trouble dealing with an entire biological population of living nonhuman primates.

Thus it is that populations of variates must usually be estimated from a small subset of the actual statistical population.

A *sample* is defined as any subset of a statistical population. While many samples are (or should be) formed through random selection of variates, *any* subset of a statistical population can be termed a sample. Some methods of sampling are more efficient than others, while other sampling procedures produce more reliable results. Regardless of the sampling procedure, however, samples are bound to reflect the character of the parent population at least to some degree.

Neither statistical populations nor their samples are thus comprised of objects per se, but rather are variates measured upon the objects. These sets of variates can be characterized by certain fundamental measures.

A *parameter* is any quantitative measure which seems to characterize a population.

Chapter 4 considers several indices (the mean, mode, standard deviation, and so forth) which estimate certain characteristics of populations. Since parameters, by definition, must always refer to populations, the phrase "population parameter" is redundant.

A *statistic* is any quantitative measure which seems to characterize a sample.

Note that this usage of the term "statistic" differs from that heard in ordinary parlance.

A parameter is a constant, fixed for the referent population. Parameters are generally unknown because most populations are incompletely observable. If a population is defined as the cranial capacity of all living mountain gorilla, then clearly the mean of that population (the mean is a parameter in this case) will never be known. For this reason, parameters are usually *estimated* from statistics which have been derived from samples. Statistics are never constant for a population, since several possible samples could have been drawn from the same population. Greek letters such as μ, σ, and ρ are conventionally assigned to represent populations, whereas Latin letters such as S, X, and r are traditionally reserved for statistics.

2.6 WHAT ARE STATISTICS?

● *After all, the higher statistics are only common sense reduced to numerical appreciation*—K. Pearson

We are finally in a position to answer a question which I hope has been troubling you throughout this chapter: What exactly are statistics? We know that a *statistic* is a measure characterizing a sample, but is this all there is to statistics?

Statistical procedures assume basically two objectives when applied to anthropological data. The initial objective is to provide precise description of phenomena, and the branch of statistics which enables one to characterize diverse data sets is called *descriptive statistics*. The other major objective of statistical analysis is to provide a systematic procedure of predicting unknown parameters through the application of *inferential statistics*.

2.6.1 Descriptive Statistics

An archaeological student has just finished excavating a 19-room pueblo; a graduate student has returned from his initial ethnological fieldwork; a novice physical anthropologist has just administered a questionnaire on heredity. What do they do with their data?

Descriptive statistics consists of a battery of standardized procedures through which masses of data can be reduced to manageable proportions. Sometimes the variates are grouped into categories, which are then ordered into a frequency distribution (discussed in Section 3.2). Or perhaps a bar graph would better illustrate the relationships of interest (Section 3.3.1). If the data are in percentages, then perhaps a circle graph should be used instead (Section 3.3.3). But occasions arise when the data are too complex for simple schemes of graphing; then new measures must be applied in order to ferret out fundamental information within the data. Measures of *central tendency*, such as the mean, the mode, and the median are handy for finding whether the data tend to group about a single point or whether there are several clusters of variates on each variable. There are also measures of *dispersion*, which summarize the tendency of variates to disperse about the central tendencies. Chapter 4 examines the usefulness, applicability and computation of common descriptive statistics.

than the average for the temperate glacial variety, can we justifiably conclude that the interglacial population had a cranial capacity which is on the average 2 cc smaller than the glacial variety? Or is this 2 cc difference too small to worry about? Are the two varieties identical with respect to cranial capacity?

Anthropologists almost always work with samples and are continually faced with similar problems. How far can they generalize their findings? Researchers sometimes try to use their common sense for such generalizations, but psychological experiments have shown that few individuals are capable of assimilating large batches of data, mentally weighing each bit of relevant information and arriving at a good, unbiased generalization (Mendenhall 1971: 176). Fortunately, a powerful tool exists which can help in analyzing sample results, and, that tool is called *statistical inference*.

Statistical inference is the process of reasoning from a sample statistic to a population parameter using the principles of probability.

The main objective of statistical inference is to make a sensible conclusion about the whole when only a part is known. Because statistical inference is always based upon incomplete data, the conclusions are only tentative. The main objective of the theory of sampling is to determine the chances of error in making inferences from a sample to a population. The degree to which such a statistical decision is reliable is often expressed in the form of *probabilities*, or the odds of making a correct decision. Chapters 3 and 4 consider various descriptive statistical methods, Chapter 5 presents the basics of probability theory, and Chapter 6 combines potentiality with description to provide the rudiments of statistical inference.

● *Know then thyself, presume not God to scan!*
The proper study of mankind is man.—A. Pope

SUGGESTIONS FOR FURTHER READING

General

Clarke (1968:chapter 1). A systems perspective of archaeology and its data base.

Kemeny (1959:chapter 15). General look at the limits of social science data.

Kerlinger (1973:chapters 1, 2 and 3). A detailed and quite sophisticated treatment of the nature of science and its relationship to social phenomena; Kerlinger carefully defines the categories of variables encountered in social research.

Naroll (1962b:chapter 1). Consideration of the nature of errors within ethnological data.

Simpson, Roe, and Lewontin (1960:chapters 1 and 2). Readable introduction to measurements and data of the biological sciences; especially relevant to physical anthropologists.

Sokal and Rohlf (1969:chapter 2). A more advanced look at biological data.

Levels of Measurement

Blalock (1972:chapter 2). Detailed consideration of the specific problems involved with common measurements scales of social science.