Chapter 1 Introduction

1.1 THREE BASIC PROCEDURES OF FUNDAMENTAL MEASUREMENT

When measuring some attribute of a class of objects or events, we associate numbers (or other familiar mathematical entities, such as vectors) with the objects in such a way that the properties of the attribute are faithfully represented as numerical properties. In this book we investigate various systems of formal properties of attributes that lead to measurement in this sense. Some commentators on physical measurement have claimed that an attribute must exhibit a more or less unique set of formal properties in order for it to have a "fundamental" measure—one that does not require the prior measurement of other quantities. For example, length can be measured fundamentally (see below), whereas density = mass/volume depends on the prior measurement of mass and volume. Not only is the intuitive distinction between fundamental and other sorts of measurement an elusive one (see Chapter 10), but whatever it may be, it is surely wrong to think that there is only one fundamental system of properties adequate to lead to numerical measurement. We present many quite different systems that are all fundamental by the intuitive criterion of independence of other measurement.

Despite the variety of systems that may lead to measurement, only a few basic procedures are known for assigning numbers to objects or events on the basis of qualitative observations of attributes. In the rest of this section we sketch the three procedures that underlie the systems in Chapters 2–13. To make the discussion concrete, we formulate the ideas in terms of length measurement.

Suppose that we have a set of straight, rigid rods whose lengths are to be measured. If we place the rods a and b side by side and adjust them so that one is entirely beside the other and they coincide at one end, then either a extends beyond b at the other end, or b beyond a, or they appear to coincide at that end also. We say, respectively, that a is longer than b, b is longer than a, or that a and b are equivalent in length. For brevity, we write, respectively, $a \succ b$, $b \succ a$, or $a \sim b$. Two or more rods can be *concatenated* by laying them end to end in a straight line, and so we can compare the qualitative length of one set of concatenated rods with that of another by placing them side by side, just as with single rods. The concatenation of a and b is denoted $a \circ b$, and the observation that c is longer than $a \circ b$ is denoted $c \succ a \circ b$, etc. Many empirical properties of length comparison and of concatenation of rods can be formulated and listed, e.g., $\succ$ is transitive; $\circ$ is associative; if $a \succ b$, then $a \circ c \succ b$; etc. Although we do not systematically list these properties now, we will use them freely in the intuitive discussion that follows.

1.1.1 Ordinal Measurement

In measuring length ordinally, we confine our observations to comparisons between simple, unconcatenated rods, and we are concerned only with assigning numbers $\phi(a)$, $\phi(b)$, etc. to rods a, b, etc. so as to reflect the results of these comparisons. That is to say, *we require the numbers be assigned so that* $a \succ b$ *if and only if* $\phi(a) > \phi(b)$. Information arising from concatenation is not used at all.

A natural procedure for assigning numbers is this. We assign to the first rod selected any number whatsoever. If the second rod chosen exceeds the first, we assign it any larger number, whereas if the first exceeds the second, we assign the second rod any smaller number. We deal with the third rod similarly except if it is between the first two. Then we assign it any number between the numbers selected for the first two. This procedure can be continued indefinitely. The assignments of numbers already made need not be affected by subsequent observations because between any two distinct numbers others always exist.

The only difficulty that can arise in carrying out the above procedure is the existence of rods a and b such that neither $a \succ b$ nor $b \succ a$, i.e., $a \sim b$. If the comparison does not establish an order, we may be tempted to conclude that the lengths are actually equal and so we assign the same number to a and b. However, if the comparison process for establishing relations is insensitive to very small disparities in length we may then find that $b \sim c$, $c \sim d$, and $b \succ d$. But to represent these relations by numbers, we require $\phi(b) = \phi(c)$, $\phi(c) = \phi(d)$, and $\phi(b) > \phi(d)$, which is impossible. Thus, this procedure for ordinal measurement is suitable only when the sensitivity of the comparison process exceeds the disparities among the rods under consideration. In the ideal case either $a \succ b$ or $b \succ a$ for any two rods a and b, except for carefully prepared "perfect copies" (e.g., meter sticks or intervals on a meter stick); for the latter, $\sim$ holds between any two copies and thus is transitive.

Of course, really perfect copies cannot be prepared. Whenever physical differences become sufficiently small, any method for observing them ultimately deteriorates. In some cases, and perhaps in all, observations of two sufficiently similar entities are inconsistent when the same comparison is repeated several times. And when inconsistencies can occur, violations of transitivity may arise. But the ideal case described above can still be achieved if copies are prepared by "standard" methods that are much more sensitive than the "working" methods used to establish $\succ$ and if the set of rods is restricted so that either $a \succ b$ or $b \succ a$ except for standardized copies. Chapter 15 discusses the case where $\sim$ is not transitive, but $\succ$ is transitive.

Clearly, the above procedure of ordinal measurement can be applied to any attribute of objects—not just length of rods—provided that a suitable comparison process leads to relations $\succ$ and $\sim$ with the requisite properties and that the set of objects is finite. The same procedure can be used for an inductive definition of a scale φ on a countable ordered set (see Section 2.1). What is less obvious is how to construct the scale when there are pairs of objects that cannot be compared directly. An interesting case, in which an ordering $\succ$ is inferred from "revealed preference" observations, has been treated in the economic literature (Arrow, 1959; Hansson, 1968b; Houthakker, 1950, 1965; Richter, 1966; Samuelson, 1938, 1947; and Uzawa, 1960). We do not deal with this problem; instead, this book is mainly about the ramifications and extensions of another procedure, to which we now turn.

1.1.2 Counting of Units

If we take into account the concatenation of rods as well as their ordering, then further constraints on the numerical assignments arise quite naturally. Suppose, for example, that $a', a'', a''', \ldots$ are perfect copies of the rod a (see Section 1.1.1). If $a \circ a' \succ b$ and $b \succ a$, then we not only want to assign numbers such that $\phi(a \circ a') > \phi(b) > \phi(a) = \phi(a')$, but we also want to represent that $a \circ a'$ is twice as long as a, i.e., $\phi(a \circ a') = 2\phi(a)$. Hence we make the assignment $\phi(b)$ so that it is between $\phi(a)$ and $2\phi(a)$. Similarly, if $a \circ a' \circ a'' \circ a''' \succ b$ and $b \succ a \circ a' \circ a''$, we place $\phi(b)$ between $3\phi(a)$ and $4\phi(a)$. And so on.

The sequence a, $2a = a \circ a'$, $3a = (2a) \circ a''$, $4a$, $5a$,... is called a *standard sequence* based on a. A meter stick graded in millimeters provides, in convenient form, the first 1000 members of a standard sequence constructed from a one-millimeter rod. If we observe that rod b falls between na and $(n+1)a$, say, between 480 and 481 mm, then we assign it a length between $n\phi(a)$ and $(n+1)\,\phi(a)$ (in the present example, between $480\phi(a)$ and $481\phi(a)$, where $\phi(a)$ is the number assigned to a one-millimeter rod and its copies). The value of $\phi(a)$ depends on the selection of a particular rod (say, e) to have unit length. If $e \sim ma$, then $\phi(a) = 1/m$. Thus, if e is the meter stick, then $m = 1000$ and the length assigned to b must be between 0.480 and 0.481 meters; if e is a centimeter rod, then $m = 10$ and $\phi(b)$ must be between 48.0 and 48.1 cm.

By choosing finer and finer standard sequences, keeping the unit of measurement fixed, of course, the value of $\phi(b)$ can be placed within an interval as small as we like. In Section 2.2 we prove a theorem establishing the convergence of these estimates as the standard sequence becomes arbitrarily fine.

Note that for purposes of ordinal measurement we could have disposed of the problem of transitivity of $\sim$ by restricting the set of rods so that either $a \succ b$ or $b \succ a$ for all rods a and b, but for standard sequences copies are essential, and the discussion in Section 1.1.1 of perfect copies applies here as well.

Three remarks should be made about the procedure of counting using standard sequences.

1. *The numbers obtained form a satisfactory ordinal measure* (Section 1.1.1): If $b \succ c$, then for some sufficiently fine-grained standard sequence based on some a, we have $b \succ na$ and $na \succ c$, so $\phi(b) > n\phi(a) > \phi(c)$. It follows that $b \succ c$ if and only if $\phi(b) > \phi(c)$.

2. *The numbers assigned are additive with respect to concatenation*: $\phi(b \circ c) = \phi(b) + \phi(c)$. The reason is that if n copies of a must be concatenated to approximate b and n' copies to approximate c, then the concatenation of $n + n'$ copies of a will approximate the concatenation of b with c. The additivity equation holds only approximately for coarse standard sequences and approaches exactness for finer and finer sequences.

3. *Regardless of the choice of unit, ratios of numerical assignments are uniquely determined by the procedure:* For if n copies of a must be concatenated to approximate b and n' copies to approximate c, then n/n' approximates $\phi(b)/\phi(c)$ more closely the finer the standard sequence.

The technique we have sketched is obviously applicable to any similar situation where the relation $\succ$ and the concatenation $\circ$ are both defined empirically. Whenever it is applied, we say that the attribute in question has been *extensively* measured (see Chapter 3). It turns out, however, that the same basic technique actually can be applied in many other, much less obvious ways (see Section 1.3.2 for an illustration and Chapters 4–8, 12, and 13 for the detailed exploitation of this idea).

1.1.3 Solving Inequalities

Suppose that five rods denoted $a_1, a_2, \ldots, a_5$, are found to satisfy

$$a_1 \circ a_5 \succ a_3 \circ a_4 \succ a_1 \circ a_2 \succ a_5 \succ a_4 \succ a_3 \succ a_2 \succ a_1 . \tag{1}$$

Data such as these can arise whenever a limited set of preselected objects and concatenations are compared and where it is impractical to go through the elaborate process of constructing standard sequences. Denote by x_i the unknown value of the length of a_i, i.e., $x_i = \phi(a_i)$, $i = 1, \ldots, 5$. From the above observations, the unknown lengths x_i must satisfy the following system of simultaneous linear inequalities.

$$\begin{aligned} x_1 + x_5 - x_3 - x_4 &> 0, \\ x_3 + x_4 - x_1 - x_2 &> 0, \\ x_1 + x_2 - x_5 &> 0, \\ x_5 - x_4 &> 0, \\ x_4 - x_3 &> 0, \\ x_3 - x_2 &> 0, \\ x_2 - x_1 &> 0. \end{aligned} \tag{2}$$

Any solution to this set of seven inequalities in five unknowns gives a possible set of values for the lengths of $a_1, \ldots, a_5$. One can thus measure the five rods by finding a solution, if one exists. Alternatively, one can obtain bounds on certain ratios of numerical assignments from the inequalities; for example, it can be shown (with some manipulation) that the ratio x_3/x_2, or $\phi(a_3)/\phi(a_2)$, is between 1 and $\frac{3}{2}$ for any solution to the above inequalities (Exercise 1).

In setting up the inequalities in Equation (2) on the basis of the observations that are given by Equation (1), the concatenation operation $\circ$ is translated into addition $+$ of real numbers, and the observational order $\succ$ is translated into the order $>$ of real numbers. Thus, for example, $a_1 \circ a_2 \succ a_5$ is represented as $x_1 + x_2 > x_5$. This translation uses properties 1 and 2 of the previous section. In other words, measurement of length by counting units in a standard sequence is a procedure which assigns the sum $\phi(b) + \phi(c)$ to the concatenation $b \circ c$, and which assigns numbers whose numerical order preserves the observational order. To measure by solving inequalities, one *assumes* that these two properties are to be satisfied by the numerical assignment; this assumption allows one to set up the inequalities to be solved.

Solving inequalities to obtain numerical assignments has many applications other than the measurement of length. These are discussed in Chapter 9. In some cases more complicated numerical operations than simple addition are assumed to correspond to empirically defined notions, and this results in nonlinear inequalities.

1.2 THE PROBLEM OF FOUNDATIONS

1.2.1 Qualitative Assumptions: Axioms

When measuring the lengths of ordinary (as opposed to atomic-sized or interstellar) objects, the standard-sequence procedure of Section 1.1.2 is obviously the one of most interest. In analyzing its foundations, one is led to the following question: *What basic assumptions must be satisfied by* $\succ$ *and* $\circ$ *in order that the standard-sequence procedure can be carried through in a self-consistent manner?* In Section 1.1 some basic properties of $\circ$ and $\succ$, e.g., if $a \succ b$, then $a \circ c \succ b$, were cited, and in describing the counting-of-units procedure we tacitly invoked others. For example, we used the one just mentioned as follows: In finding elements na and $(n + 1)a$, in the standard sequence based on a, such that $(n + 1)\, a \succ b$ and $b \succ na$, we need to be assured that the integer n is unique. We infer this, because if $(n + 1)\, a \succ b$, then $(n + 2)\, a \circ a$ is also longer than b, etc., so that all subsequent members of the sequence are longer than b. The critical step, inferring that $(n + 1)\, a \circ a \succ b$ on the basis of $(n + 1)\, a \succ b$, depends on the truth of the above property (and on the transitivity of $\succ$). Similarly, another property tacitly assumed is the existence of some integer n such that $(n + 1)\, a \succ b$; i.e., we need to know that a is not "infinitesimally small" compared to b. For no really good reason, except mathematical tradition, this is called an Archimedean assumption.

To put our understanding of the counting-of-units procedure in good order, we must make all of these assumptions explicit. A measurement procedure certainly is not adequately understood if it depends on properties that are not explicitly recognized. Once they are explicit, deciding whether or not the same measurement procedure is applicable in a new domain reduces to testing whether or not the requisite properties are satisfied.

Still more orderly foundations are obtained if we can deduce most properties as theorems from a few basic properties. Applicability to a new situation is then judged according to whether those few properties are satisfied. This logical step is called the *axiomatization* of the measurement procedure. The few explicit properties from which all others are deduced are called the *axioms*. Axiomatization of any body of propositions can always be achieved in more than one way; some criteria for "good" axiomatizations are discussed later.

Geometry is a beautiful and far-reaching example of a foundational treatment of measurement. The science of geometry (i.e., earth measurement) was probably first developed as a set of practical procedures, either for the direct measurement of lengths and areas on the earth's surface or in connection with the astronomy devised to serve astrology. Eventually, the tacit assumptions of practice were formulated explicitly as theorems of geometry, and these were systematically organized and deduced from a few axioms and postulates by Euclid. Certain additional tacit assumptions, unrecognized by Euclid, were discovered later, and an orderly axiomatization of Euclidean geometry was finally devised by Hilbert (1899) and others. These axiomatic studies are known as the foundations of geometry. For further discussion of geometry as measurement, see Chapters 11–13.

Although it seems quite natural to view the task of the foundations of measurement to be the explication and systematization of the assumptions required by particular interesting procedures of measurement, doing so has actually led to some serious misunderstandings. These stem from the easy supposition that an empirical concatenation operation is *sine qua non* for the standard-sequence procedure. Campbell (1920, 1928), in his influential books on measurement, and some later philosophers (e.g., Cohen & Nagel, 1934; Ellis, 1966) treated fundamental measurement as practically synonymous with procedures involving empirically defined concatenation operations. (One of Campbell's remarks, p. 327 of the 1957 edition of 1920, makes clear that he was aware of a potential distinction: "Of course there may possibly be some other way of assigning numerals to represent properties differing in first principles from that described in Chapter X [which is devoted to the procedure of Section 1.1.2]; but until somebody suggests such a way, it is hardly worthwhile to discuss the possibility; it is certainly not employed in the actual physics of today.") The absence of appropriate, empirically defined, concatenation operations in psychology has even led some serious students of measurement to conclude that fundamental measurement is not possible there in the same sense that it is possible in physics (Guild, 1938). Reese (1943) discussed this in detail and attempted to provide psychological examples of scales based on concatenation. Many examples given in this book show that Campbell's viewpoint is untenable.

1.2.2 Homomorphisms of Relational Structures: Representation Theorems

We may view the foundations of measurement in a slightly different way by focusing on properties of the numerical assignment, rather than on the procedure for making the assignment. Specifically, for the standard-sequence procedure, we may pose the following question: Given a set of rods, a comparison relation $\succ$, and a concatenation $\circ$, what assumptions concerning $\succ$ and $\circ$ are necessary and/or sufficient to construct a real-valued function ϕ that is order preserving and additive—i.e., that satisfies properties 1 and 2 of Section 1.1.2? This question still asks for an axiomatization—the listing of certain properties of $\succ$ and $\circ$—however, the conclusion aimed for is not that a certain procedure is possible, but rather that a numerical function ϕ satisfying certain properties exists. The procedure to be used in assigning numbers (constructing ϕ) is not specified in posing the problem; thus, quite distinct axiomatizations, which operate through different procedures—say, counting units and solving inequalities—may be expected.

The next step, a small but important one, recognizes that the numerical assignment ϕ, satisfying properties 1 and 2 of Section 1.1.2 (order preserving and additive), is a homomorphism of an empirical relational structure into a numerical relational structure. To make our meaning clear we must first say what we mean by a relational structure and then what we mean by a homomorphism between two relational structures.

A *relational structure* is a set together with one or more relations on that set. If we denote the set of all the rods and all the finite concatenations of rods under consideration by A, then the empirical relational structure for the procedures of Sections 1.1.2 and 1.1.3 is denoted $\langle A, \succ, \circ \rangle$. (The concatenation operation is a ternary relation on A, holding among a, b, and $c = a \circ b$, whereas $\succ$ is a binary relation on A.) An appropriate numerical relational structure is $\langle \mathrm{Re}, >, + \rangle$, where Re is the set of real numbers, $>$ is the usual greater than relation, and $+$ is the ordinary operation of addition. (Angle brackets $\langle\ \rangle$ rather than parentheses are used in giving an explicit listing of a relational structure.) The numerical assignment ϕ is a *homomorphism*[1] in the sense that it sends A into Re, $\succ$ into $>$, and $\circ$ into $+$ in such a way that $>$ preserves the properties of $\succ$ (property 1, Section 1.1.2) and $+$ the properties of $\circ$ (property 2, Section 1.1.2).

This formulation generalizes naturally to other relational structures. Given an empirical relation R on a set A and a numerical relation S on Re, a function ϕ from A into Re takes R into S provided that the elements $a, b,...$ in A stand in relation R if and only if the corresponding numbers $\phi(a)$, $\phi(b)$,... stand in relation S. More generally, if $\langle A, R_1, ..., R_m \rangle$ is an empirical relational structure and $\langle \mathrm{Re}, S_1, ..., S_m \rangle$ is a numerical relational structure, a real-valued function ϕ on A is a *homomorphism* if it takes each R_i into S_i, $i = 1,..., m$. Still more generally, we may have n sets $A_1, ..., A_n$, m relations $R_1, ..., R_m$ on $A_1 \times \cdots \times A_n$, and a vector-valued homomorphism ϕ, whose components consist of n real-valued functions $\phi_1, ..., \phi_n$ with ϕ_j defined on A_j, such that ϕ takes each R_i into relation S_i on Re^n. A *representation theorem* asserts that if a given relational structure satisfies certain axioms, then a homomorphism into a certain numerical relational structure can be constructed. A homomorphism into the real numbers is often referred to as a scale in the psychological measurement literature.

[1] We speak of a homomorphism (rather than an isomorphism) because ϕ is not usually one to one; in general $\phi(a) = \phi(b)$ does not mean that the rods a and b are identical, but merely of equal length. A one-to-one homomorphism is called an *isomorphism*.

From this standpoint, measurement may be regarded as the construction of homomorphisms (scales) from empirical relational structures of interest into numerical relational structures that are useful. Foundational analysis consists, in part, of clarifying (in the sense of axiomatizing) assumptions of such constructions.

This view of measurement would be entirely too abstract were it not for the fact that interesting examples of measurement exist that involve relational structures quite different from $\langle A, \succ, \circ \rangle$. The development of such additional examples of measurement spurred the formulation of this abstract viewpoint. Among the key examples were the axiomatizations of utility measurement by von Neumann and Morgenstern (1947), Savage (1954), Suppes and Winet (1955), and Davidson, Suppes, and Siegel (1957), and the axiomatization of semiorders by Luce (1956). An explicit statement, of the relational structure viewpoint was first given by Scott and Suppes (1958); also see Ducamp and Falmagne (1969) and Suppes and Zinnes (1963).

Despite the proliferation of measurement axiomatizations, the procedures for constructing numerical assignments remain about the same; the most important ones are those described in Section 1.1. One of our goals in this book is to show clearly just how the assignments of numbers in what are quite disparate representation theorems all reduce to one or another of these three basic procedures.

1.2.3 Uniqueness Theorems

In discussing the measurement of length based on the counting of units (Section 1.1.2), we pointed out that the number $\phi(a)$ assigned to rod a depends on which rod e is chosen as unit, i.e., $\phi(e) = 1$. This choice is entirely arbitrary. Moreover, as we noted, the ratio, $\phi(a)/\phi(e)$ is uniquely determined independent of whether e or some other rod e' is chosen as the unit. Thus, if ϕ is the numerical function constructed with e as unit and if ϕ' is constructed with e' as unit, we have

$$\phi(a)/\phi(e) = \phi'(a)/\phi'(e),$$

or, substituting $\phi(e) = 1$ and $\phi'(e) = \alpha$,

$$\phi'(a) = \alpha\phi(a). \tag{3}$$

Conversely, starting with ϕ based on e as unit, we can select e' with $\phi(e') = 1/\alpha$, and obtain a new scale (homomorphism) ϕ' satisfying Equation (3). These facts are usually expressed by saying that the *similarity* transformations

$$\phi \to \alpha\phi \doteq \phi', \quad \alpha > 0, \tag{4}$$

are *permissible* transformations of the scale ϕ. A scale whose permissible transformations are only those of Equation (4) is called a *ratio scale.*

The term "ratio scale" comes from the fact that if $\phi \to \alpha\phi$ are the only permissible transformations, then the ratios of scale values are determined uniquely.

Other families of measurement procedures are related by different sets of admissible transformations. For example, Celsius temperature is related to Fahrenheit by C = (5/9)(F − 32). Observe that in any ordinary temperature measurement, two arbitrary choices are made, the zero point and the unit. Varying these leads to the *affine* transformations which are of the form

$$\phi \to \alpha\phi + \beta, \quad \alpha > 0. \tag{5}$$

A scale whose permissible transformations are only those of Equation (5) is called an *interval scale* because ratios of intervals are invariant:

$$\frac{\phi'(a) - \phi'(b)}{\phi'(c) - \phi'(d)} = \frac{[\alpha\phi(a) + \beta] - [\alpha\phi(b) + \beta]}{[\alpha\phi(c) + \beta] - [\alpha\phi(d) + \beta]} = \frac{\phi(a) - \phi(b)}{\phi(c) - \phi(d)}.$$

Two other classes of transformations play a key role in measurement. The *power* transformations are of the form

$$\phi \to \alpha\phi^{\beta}, \quad \alpha > 0, \quad \beta > 0, \tag{6}$$

and a scale whose permissible transformations are those of Equation (6) is called a *log-interval scale* because a logarithmic transformation of such a scale results in an interval scale. As we shall argue in Chapter 10, many of

the common physical scales that are usually said to be ratio scales are, in fact, log-interval scales. Density is an example.

Finally, the *monotonic increasing* transformations are of the form

$$\phi \to f(\phi), \tag{7}$$

where f is any strictly increasing real-valued function of a real variable, and a scale whose permissible transformations are those of Equation (7) is called an *ordinal scale.* The reason is that only order is preserved under these transformations.

Stevens (1946, 1951) was the first to recognize and emphasize the importance of the type of uniqueness exhibited by a measurement homomorphism and he isolated these four types—ratio, interval, log-interval, and ordinal.[2]

A classification of measurement in terms of permissible transformations is clear cut only so long as it is certain which transformations are permissible. Little ambiguity exists for length: there is a family of closely related procedures, described in Section 1.1.2, which differ from one another only in the rather trivial and arbitrary matter of which size rod is chosen as unit. Permissible transformations are precisely those produced by variations in this matter of procedure. On closer examination, however, it becomes less clear which other choices in a measurement procedure are arbitrary and which are not. For example, in measuring length, we choose not only the unit, but we choose to count and record the number of copies of a (say, n) that are needed to approximate b, rather than to record, for example, the square or exponential of that number, n^2 or e^n. Is this also an arbitrary choice? Would a procedure that recorded n^2 be closely related to one that recorded n, with the consequence that the transformation $\phi \to \phi^2$ is also permissible? If not, why not?

[2] Surprisingly, he later (Stevens, 1957, 1968) generated an ambiguity in the use of these terms by describing his magnitude estimation scale as a "ratio scale" of measurement. In this experimental procedure, observers are asked to assign numbers to stimuli "in proportion to the sensations evoked," and the resulting numbers are taken to be scale values. In the sense that subjects are asked to produce numbers that preserve subjective "ratios," one sees why this scale might be described as a ratio scale—except for the fact that he earlier introduced the term to refer to those theories in which any two homomorphisms are related by a similarity transformation. Stevens has not provided any argument showing that the procedure of magnitude estimation can be axiomatized so as to result in a ratio-scale representation; he has neither described the empirical relational structure, the numerical relational structure, nor the axioms which permit the construction of a homomorphism. In Chapter 4, we provide a set of plausible axioms for families of matching experiments (which generalize magnitude estimation) and, if the axioms are empirically valid, we have nearly justified Stevens' claim.

The reason for rejecting the n^2 or e^n procedures is easily given: the resulting scales, related to the normal one by ϕ^2 or e^ϕ, are not additive. Instead, they satisfy other rules, namely,

$$\phi^2(a \circ b) = \phi^2(a) + 2[\phi^2(a)\,\phi^2(b)]^{1/2} + \phi^2(b),$$
$$e^\phi(a \circ b) = e^\phi(a)\,e^\phi(b).$$

Hence, they do not yield homomorphisms of $\langle A, \succsim, \circ \rangle$ into $\langle \mathrm{Re}, \geq, + \rangle$; instead, they yield homomorphisms of $\langle A, \succsim, \circ \rangle$ into numerical structures $\langle \mathrm{Re}, \geq, * \rangle$, where $*$ is a binary operation differing from addition. In the above examples $*$ is defined, respectively, by $x * y = x + 2(xy)^{1/2} + y$ and $x * y = xy$.

The concept of permissible transformations is much clearer from the standpoint of homomorphisms between relational structures than from the standpoint of arbitrary choices in measurement procedures. A transformation $\phi \to \phi'$ is permissible if and only if ϕ and ϕ' are both homomorphisms of $\langle A, R_1, ..., R_n \rangle$ into the *same* numerical structure $\langle \mathrm{Re}, S_1, ..., S_n \rangle$. Thus, if ϕ is order preserving and additive—a homomorphism of $\langle A, \succsim, \circ \rangle$ into $\langle \mathrm{Re}, \geq, + \rangle$—the same is true for $\alpha\phi$, provided that $\alpha > 0$; moreover, if ϕ' is *any* homomorphism of $\langle A, \succsim, \circ \rangle$ into $\langle \mathrm{Re}, \geq, + \rangle$, then $\phi' = \alpha\phi$ for some $\alpha > 0$. The latter result, which is the substance of what is proved in a *uniqueness theorem*, is not obvious.

We conclude, then, that an analysis into the foundations of measurement involves, for any particular empirical relational structure, the formulation of a set of axioms that is sufficient to establish two types of theorems: a representation theorem, which asserts the existence of a homomorphism ϕ into a particular numerical relational structure, and a uniqueness theorem, which sets forth the permissible transformations $\phi \to \phi'$ that also yield homomorphisms into the same numerical relational structure. A measurement procedure corresponds to the construction of a ϕ in the representation theorem.

It is important to note that every pair of representation and uniqueness theorems involves a choice of a numerical relational structure. This choice is essentially a matter of convention, although the conventions are strongly affected by considerations of computational convenience. For example, either of the $*$ operations proposed above is clumsy compared with addition. The subject of alternative numerical structures is considered in more detail in Sections 3.9, 4.4.2, 6.5.2, 7.2, 7.4.2, and in Chapter 19.

For an interesting attempt to treat the mathematics of measurement in an elementary way by focusing mostly on the uniqueness properties of various representations without, however, going deeply into the question of the existence of representations, see Blakers (1967).

1.2.4 Measurement Axioms as Empirical Laws

We have just emphasized that the numerical scales of measurement are subject to arbitrary conventions. There are permissible transformations, corresponding to arbitrary choices of unit, and the very representation and uniqueness theorems themselves depend on the conventional choice of a numerical relational structure. What is invariant, and so is not a matter of convention, is the empirical relational structure and its empirical properties, some of which are formulated as axioms. A set of axioms leading to representation and uniqueness theorems of fundamental measurement may be regarded as a set of qualitative (that is, nonnumerical) empirical laws. In some cases, as in the measurement of length, these laws are rather trivial, i.e., not intrinsically very interesting. In other empirical contexts, the axioms can be quite interesting and nonobvious. In such cases, the development of measurement scales is closely linked to the formulation and testing of appropriate qualitative laws. This viewpoint has been discussed by Krantz (1971).

We shall make an effort to point out the status of various axioms or classes of axioms as empirical laws. The type of consideration that arises is illustrated in Sections 1.3, 1.4.5, and 1.5.

1.2.5 Other Aspects of the Problem of Foundations

In analyzing the foundations of measurement, one of the main concerns is formalization: the choice of an empirical relational structure as an abstraction from the available data, the choice of an appropriate numerical relational structure, the discovery of suitable axioms, and the construction of numerical homomorphisms, i.e., proving the representation theorem and uniqueness theorem. However, this formalization process does not exhaust the problem of foundations by any means. The most important omission is an analysis of error of measurement. This involves difficult conceptual problems concerning the relation between detailed, inconsistent data and the abstraction derived from them, the empirical relational structure. We discuss these aspects of the foundations of measurement, which are poorly understood, as well as we can in Chapters 15–17.

1.3 ILLUSTRATIONS OF MEASUREMENT STRUCTURES

In the previous section, we presented a general statement of what is involved in a foundational analysis of measurement. Nevertheless, we refrained from trying to hammer out an acceptable definition of the concept of a formal theory of measurement. Experience suggests that after some exposure to paradigmatic theories of measurement, students have little difficulty in recognizing other examples. It is similar to recognizing grammaticalness of English sentences: with some borderline exceptions, sequences of words are readily identified as grammatical or not grammatical, even though no general definition is yet available. In this section, then, we try to make our previous generalities about relational structures, axioms, etc. more comprehensible by presenting two examples. These examples also motivate the discussion of axiom types in Section 1.4 and provide the material for a set of exercises.

1.3.1 Finite Weak Orders

It is useful to begin with a consideration of ordinal measurement. Most of the relational structures we shall consider involve an ordering relation, and the concepts needed to handle the ordering relation in the more complicated structures can be developed and presented in isolation here. Moreover, weak orders provide the simplest illustration of the ideas presented above.

DEFINITION 1. *Let A be a set and $\succsim$ be a binary relation on A, i.e., $\succsim$ is a subset of $A \times A$. The relational structure $\langle A, \succsim \rangle$ is a* weak order *iff,*[3] *for all $a, b, c \in A$, the following two axioms are satisfied:*

1. *Connectedness: Either $a \succsim b$ or $b \succsim a$.*
2. *Transitivity: If $a \succsim b$ and $b \succsim c$, then $a \succsim c$.*

Definition 1 is typical of our format: we single out and name a class of relational structures that satisfy a particular set of axioms. In this case, the name "weak order" is well established, although "pre-order" is sometimes used.

A weak order is always *reflexive* since Axiom 1 implies $a \succsim a$ for all a.

Note that the numerical relational structure $\langle \text{Re}, \geqslant \rangle$, where Re denotes the set of all real numbers and $\geqslant$ is the usual ordering (i.e., $x \geqslant y$ if and only if x is greater than or equal to y), is a weak order. However, this weak order is special in that it is *antisymmetric*: if both $x \geqslant y$ and $y \geqslant x$, then $x = y$. Such an anti-symmetric weak order is called a *simple* or *total order*. In general, a weak order is distinguished from a simple order because it is possible that $a \succsim b$ *and* $b \succsim a$, for distinct elements a, b of A. As we shall see, every weak order is associated in a natural way with a simple order. Most empirical ordering operations yield weak orders, in the sense that there exist distinct elements that are equivalent, i.e., $a \succsim b$ and $b \succsim a$ hold.

[3] In all formal definitions, theorems, and proofs, we use "iff" to stand for "if and only if."

THEOREM 1. *Suppose that A is a finite nonempty set. If $\langle A, \succsim \rangle$ is a weak order, then there exists a real-valued function ϕ on A such that for all $a, b \in A$,*

$$a \succsim b \quad \text{iff} \quad \phi(a) \geqslant \phi(b).$$

Moreover, ϕ' is another real-valued function on A with the same property iff there is a strictly increasing function f, with domain and range equal to Re, *such that for all $a \in A$*

$$\phi'(a) = f[\phi(a)],$$

i.e., ϕ is an ordinal scale.

Theorem 1 formulates the representation and uniqueness results for ordinal measurement in finite sets. It asserts that if a relational structure $\langle A, \succsim \rangle$ satisfies Axioms 1 and 2 of Definition 1, then there is a homomorphism ϕ into the numerical structure $\langle \text{Re}, \geqslant \rangle$, and the permissible transformations consist of all strictly increasing functions from Re onto Re. We shall usually state representation and uniqueness theorems in the format of Theorem 1; the appropriate numerical structure and the properties corresponding to a homomorphism (in this case, ϕ carries $\succsim$ into $\geqslant$) will be apparent.

The proof of Theorem 1 is well worth presenting in detail here; it introduces concepts and notations that are essential in dealing with weak orders throughout the book (Definition 2 below) and it illustrates the relation between axioms and measurement procedures. A proof of Theorem 1 should provide two things: a definite method for constructing the order-preserving function ϕ, and a method for constructing the function f, given ϕ and ϕ'. The method of constructing ϕ is precisely the measurement procedure. This will be true in all our representation theorems.

DEFINITION 2. *If $\succsim$ is a binary relation on A, two new relations $\sim$ and $\succ$ are defined on A as follows:*

$$a \sim b \quad \text{iff} \quad a \succsim b \text{ and } b \succsim a,$$
$$a \succ b \quad \text{iff} \quad a \succsim b \text{ and not } (b \succsim a).$$

These are referred to as the symmetric *and* asymmetric *parts of $\succsim$, respectively. If $\langle A, \succsim \rangle$ is a weak order, it is easy to prove (Exercise 4) that $\sim$ is an equivalence relation on A (i.e., it is reflexive, symmetric, and transitive) and that $\succ$ is transitive and asymmetric [i.e., if $a \succ b$, then not $(b \succ a)$]. The set*

$$\mathbf{a} = \{b \mid b \in A, \;\; b \sim a\}$$

is called the equivalence class *determined by a. It is well known that* $\mathbf{a} \cap \mathbf{b}$ *is nonnull iff* $b \in \mathbf{a}$, *in which case* $\mathbf{a} = \mathbf{b}$ *(i.e.,* $c \sim a$ *iff* $c \sim b$*); hence, the distinct equivalence classes form a* partition *of A (i.e., they form a family of pairwise disjoint subsets whose union is A). The set of equivalence classes is denoted* $A/\!\sim$. *The weak order* $\succsim$ *induces a new ordering relation* $\succsim$ *on* $A/\!\sim$,

$$\mathbf{a} \succsim \mathbf{b} \quad \textit{iff} \quad a \succsim b.$$

It is easy to show that $\succsim$ *is a simple order (Exercise 4).*

Definition 2 gives a brief review of several basic concepts used throughout the book. If you are unfamiliar with these ideas, you should study some elementary theory of sets and relations found in, e.g., Kershner and Wilcox (1950) or Suppes (1957).

The concepts of Definition 2 are important for the following reason. If ϕ has been constructed as specified in Theorem 1 and if $a \sim b$, then both $\phi(a) \geqslant \phi(b)$ and $\phi(b) \geqslant \phi(a)$ must hold. By the antisymmetric property of $\langle \text{Re}, \geqslant \rangle$, we have $\phi(a) = \phi(b)$. In short, two elements of A that lie in the same equivalence class must have the same scale value; clearly, the converse is also true. The proof of Theorem 1 thus reduces to constructing scale values that preserve the order between different equivalence classes. That is, it suffices to construct a real-valued function ϕ on $\langle A/\!\sim, \succsim \rangle$, such that

$$\mathbf{a} \succsim \mathbf{b} \quad \text{iff} \quad \phi(\mathbf{a}) \geqslant \phi(\mathbf{b}).$$

We then obtain ϕ on A by setting $\phi(a) = \phi(\mathbf{a})$.

In short, we have reduced the conditions of the representation theorem to the case where the weak order is a simple order. The uniqueness theorem also reduces to this case, because ϕ and ϕ have the same range in Re; therefore, the function f will be exactly the same: $\phi'(a) = f[\phi(a)]$ iff $\phi'(\mathbf{a}) = f[\phi(\mathbf{a})]$.

We now complete the proof of the representation theorem. For each $\mathbf{a} \in A/\!\sim$, let $\phi(\mathbf{a})$ be the number of distinct equivalence classes $\mathbf{b}$ such that $\mathbf{a} \succsim \mathbf{b}$. (Note that this counting process assigns the number 1 to the lowest equivalence class, 2 to the next lowest, etc.) If $\mathbf{a} \succsim \mathbf{b}$ then for every $\mathbf{c}$, if $\mathbf{b} \succsim \mathbf{c}$, then $\mathbf{a} \succsim \mathbf{c}$ (transitivity), so if $\mathbf{c}$ is counted for $\phi(\mathbf{b})$, it is also counted for $\phi(\mathbf{a})$. Thus $\phi(\mathbf{a}) \geqslant \phi(\mathbf{b})$. Conversely, if not ($\mathbf{a} \succsim \mathbf{b}$), then $\mathbf{b} \succsim \mathbf{a}$ (connectedness), and also $b \succ a$. Thus, there is at least one $\mathbf{c}$ (namely, $\mathbf{b}$) counted in $\phi(\mathbf{b})$ but not in $\phi(\mathbf{a})$, and we have $\phi(\mathbf{b}) > \phi(\mathbf{a})$. This completes the proof of the representation theorem.

The proof of the uniqueness theorem is easy and not particularly instructive; so we omit it (but see Exercise 7).

Note that the procedure used to construct ϕ—counting the number of equivalence classes below a given one—only works for finite sets. A different proof of the representation theorem can be given using the procedure outlined in Section 1.1.1 (see also Exercise 6) which applies to *countable* sets (those in one-to-one correspondence with the positive integers). See Theorem 2.1 for the details.

Finally, we remark that the axioms of transitivity and connectedness were essential in the above construction: their uses were noted in the course of the proof. Both of these qualitative laws have been challenged on empirical grounds in social-science applications. For example, consider a binary relation $\succsim$ that reflects an individual's preferences for various objects. He may decide $a \succ b$ and $b \succ c$ by concentrating on one dimension (say, quality) and ignoring small differences in another (say, price); but the price difference between a and c may be more salient, leading to $c \succsim a$. Similarly, there may be some pairs for which there is neither strict preference ($a \succ b$ or $b \succ a$) nor indifference ($a \sim b$). Thus, these laws can be nontrivial from an empirical standpoint.

1.3.2 Finite, Equally Spaced, Additive Conjoint Structures

In the simplest relational structure, $\langle A, \succsim \rangle$, considered above, we cannot count units because there is no way of identifying which element of A is the "sum" of two others and, hence, no way of deciding what constitutes two units. In order to count units, the structure must have some additional features. One of the simplest possibilities is for the set A to be a cartesian product, $A = A_1 \times A_2$. Empirically, this amounts simply to saying that two factors determine the ordering $\succsim$. A given object a corresponds to a level a_1 of the A_1-factor and a level a_2 of the A_2-factor. Such objects are denoted $a = (a_1, a_2)$, $b = (b_1, b_2)$, etc.

As an example, let the A_1-factor be temperature, the A_2-factor humidity, and the relation $\succsim$ be discomfort. Thus, $(a_1, a_2) \succ (b_1, b_2)$ if temperature a_1 together with humidity a_2 is less comfortable than temperature b_1 together with humidity b_2. This example suggests an attribute such as discomfort induces an ordering on each component separately—in this case, the ordering by temperature and by humidity. Formally, we mean the ordering obtained when the value of the other component is held fixed. Thus, we have $\succsim_1$ on A_1 defined by

$$a_1 \succsim_1 b_1 \text{ iff } (a_1, c_2) \succsim (b_1, c_2) \text{ for all } c_2 \text{ in } A_2;$$

and $\succsim_2$ on A_2 defined by

$$a_2 \succsim_2 b_2 \text{ iff } (c_1, a_2) \succsim (c_1, b_2) \text{ for all } c_1 \text{ in } A_1.$$

Observe that, at the moment, we cannot assert that either $\succsim_1$ or $\succsim_2$ is a weak order because we cannot be sure that they are connected; it may happen, for all we know, that for some c_2 in A_2 we have $(a_1, c_2) \succ (b_1, c_2)$ and for some d_2 in A_2 $(b_1, d_2) \succ (a_1, d_2)$. We suppose, however, in the following discussion that both $\succsim_1$ and $\succsim_2$ are weak orders.

With a product structure of this sort, the entities that can be concatenated are *intervals within one factor*. By an interval in A_1, we simply mean the formal entity denoted a_1b_1, where a_1, b_1 are in A_1 and are called the "end points" of the interval. It will prove convenient to adhere to the convention that when $a_1 \succsim_1 b_1$ we write the interval as a_1b_1, not as b_1a_1. Since the intervals a_1b_1 and b_1c_1 are adjacent, a_1c_1 can be regarded as their "sum." Two intervals a_1b_1 and c_1d_1 can be regarded as equal if they are matched by the *same* interval a_2b_2 on the second factor, where by a_1b_1 matching a_2b_2 we simply mean that $(a_1, b_2) \sim (b_1, a_2)$. This assumes *additivity* in the effects of the two factors: if the sum of a_1 and b_2 effects equals the sum of b_1 and a_2 effects, then the difference between a_1 and b_1 effects must equal the difference between a_2 and b_2 effects. The method of forming equal intervals and concatenating them to obtain a standard sequence (Section 1.1.2) on the A_1-factor is illustrated in Figure 1. In short, the presence of a second

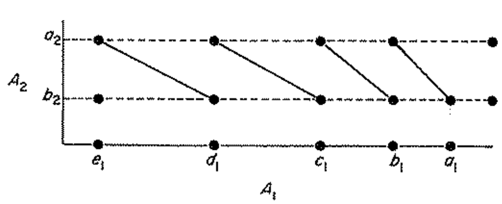

FIGURE 1. Use of an interval a_2b_2 in A_2 to lay off equal adjacent intervals a_1b_1, b_1c_1, ..., d_1e_1 of a standard sequence on A_1. The points joined by straight lines are observed equivalences in the ordering $\succsim$ on $A_1 \times A_2$. Thus, $(a_1, b_2) \sim (b_1, a_2)$ implies a_1b_1 matches a_2b_2; ...; and $(d_1, b_2) \sim (e_1, a_2)$ implies d_1e_1 matches a_2b_2.

factor together with the assumption of additivity of the effects of the two factors allows us both to calibrate equal units on the first factor and to combine adjacent equal units to form a standard sequence. This permits us to use a counting-of-units process formally the same as that of Section 1.1.2.

DEFINITION 3. *Let A_1 and A_2 be nonempty sets and let $\succsim$ be a binary relation on $A = A_1 \times A_2$. The relational structure $\langle A_1 \times A_2, \succsim\rangle$ is called*

an independent conjoint structure iff, for all $a, b, c \in A$ and all $a_i, b_i, c_i, d_i \in A_i$, $i = 1, 2$, the following four axioms are satisfied:

1. *Either $a \succsim b$ or $b \succsim a$.*
2. *If $a \succsim b$ and $b \succsim c$, then $a \succsim c$.*
3. *If $(a_1, c_2) \succsim (b_1, c_2)$, then $(a_1, d_2) \succsim (b_1, d_2)$.*
4. *If $(c_1, a_2) \succsim (c_1, b_2)$, then $(d_1, a_2) \succsim (d_1, b_2)$.*

Independent Conjoint Structure

Note that Axioms 1 and 2 can be combined into the statement that $\langle A, \succsim\rangle$ is a weak order.

DEFINITION 4. *Let $\langle A_1 \times A_2, \succsim\rangle$ be an independent conjoint structure. Define relations $\succsim_1$ on A_1 and $\succsim_2$ on A_2 by:*

$a_1 \succsim_1 b_1$ iff there exists c_2 in A_2 with $(a_1, c_2) \succsim (b_1, c_2)$;
$a_2 \succsim_2 b_2$ iff there exists c_1 in A_1 with $(c_1, a_2) \succsim (c_1, b_2)$.

It is easy to show (using Axioms 1–4) that $\langle A_i, \succsim_i\rangle$ is a weak order, $i = 1, 2$ (Exercise 11). Axiom 3 is precisely what is needed in order to prove that $\succsim_1$ on A_1 is a weak order, and Axiom 4 plays the parallel role in proving $\succsim_2$ on A_2 is a weak order. The term "independent" in Definition 3 refers to the fact that the induced order $\succsim_i$ on A_i does not depend on the choice of c_j or d_j in the other factor A_j.

DEFINITION 5. *Define a relation J_i on A_i, $i = 1, 2$, by $a_i J_i b_i$ iff, for all $c_i \in A_i$, exactly one of the following holds: $c_i \succsim_i a_i$ or $b_i \succsim_i c_i$. The structure $\langle A_1 \times A_2, \succsim\rangle$ is called an* equally spaced, additive conjoint structure, *if in addition to Axioms 1–4, the following axiom holds for all $a_i, b_i \in A_i$, $i = 1, 2$:*

5. *If $a_1 J_1 b_1$ and $b_2 J_2 a_2$, then $(a_1, a_2) \sim (b_1, b_2)$.*

The J_i relation means that a_i is strictly larger than b_i (with respect to $\succsim_i$) and nothing lies between the two elements; thus, any c_i is either $\succsim_i a_i$ or $\precsim_i b_i$, but not both. Axiom 5 asserts that objects are equally spaced, meaning that any two J-intervals are equal in the intuitive sense of calibration of A_1-intervals against A_2-intervals as discussed above (Figure 1). The representation and uniqueness theorems are formulated as follows:

THEOREM 2. *Suppose that A_1 and A_2 are finite nonempty sets. If $\langle A_1 \times A_2, \succsim\rangle$ is an equally spaced, additive conjoint structure, then there exist real-valued functions ϕ_i on A_i, $i = 1, 2$, such that, for all $a = (a_1, a_2)$, $b = (b_1, b_2) \in A$,*

$$a \succsim b \quad \text{iff} \quad \phi_1(a_1) + \phi_2(a_2) \geqslant \phi_1(b_1) + \phi_2(b_2).$$

Moreover, assuming that each $A_i/\sim_i$ contains at least two equivalence classes, then if ϕ_1', ϕ_2' are any other pair of real-valued functions with the above property there exist constants α, β_1, β_2, with $\alpha > 0$, such that $\phi_i' = \alpha\phi_i + \beta_i$, $i = 1, 2$.

The proof of this theorem is left as Exercises 12–14. All of the essential ideas for the proof were sketched above: the intervals in each factor form a finite standard sequence. Note that the representation theorem consists of a two-component vector homomorphism, namely (ϕ_1, ϕ_2), between $\langle A_1 \times A_2, \succsim \rangle$ and $\langle \text{Re} \times \text{Re}, \succsim' \rangle$, where $\succsim'$ is defined on Re × Re by $(x, y) \succsim' (u, v)$ iff $x + y \geqslant u + v$. The uniqueness theorem asserts that the ϕ_i are interval scales with a common unit (constant α) but independent zero points (constants β_1, β_2). This is not surprising. Since the counting-of-units process is applied to intervals on each factor, the ordinary ideas of extensive measurement lead one to expect invariance of ratios of intervals. But since only intervals in each factor are determined, the origins of the two scales are arbitrary.

In the previous section, we briefly discussed the empirical status of the weak order assumption (Axioms 1 and 2). Little needs to be added here except to note that the two factor aspect of the objects is more likely to lead to violations of transitivity. These will occur in any situation where attention is sometimes exclusively focused on one factor, sometimes on the other.

Axioms 3 and 4 are empirical laws of a very interesting type. We call them *independence laws.* Axiom 3 asserts that the ordering of A_1-effects is *independent of* the choice of a fixed level in A_2, which we abbreviate by saying that A_1 is independent of A_2. Axiom 4 asserts that A_2 is independent of A_1. Intuitively, independence is a qualitative, ordinal version of noninteraction between two variables. Of course, additivity (the conclusion of Theorem 2) asserts a quantitative noninteraction that is much stronger. Independence laws play a very prominent role in the discussion of additive and polynomial conjoint measurement (Chapters 6 and 7), in utility measurement (Chapter 8), and in multidimensional proximity measurement (Chapter 13).

The final axiom, 5, is of a very different nature. It is hardly ever satisfied by accident. If A_1 represents a finite set of levels of some factor and A_2 represents a different factor, there is no reason whatsoever to suppose that when we move from (b_1, b_2) to the next higher level of A_1, say (a_1, b_2), the effect is *exactly* the same as when we move to the next higher level of A_2, say (b_1, a_2). What one might try to do, in practice, is to *select* subsets of levels of the two factors so as to satisfy this property, much as we select a standard sequence of weights and lengths. Thus, if we start by choosing high levels a_1 and a_2 and if we choose as the next highest level b_1 in A_1, then we are constrained to choose the next highest level b_2 in A_2 so that $(a_1, b_2) \sim (b_1, a_2)$. We are then forced to choose the next A_1 level c_1 to be such that $(c_1, a_2) \sim (b_1, b_2)$, since $b_1 J_1 c_1$ and $a_2 J_2 b_2$. Similarly, c_2 has to be chosen so that $(b_1, b_2) \sim (a_1, c_2)$. But now, with all degrees of freedom gone, we are forced to have $(b_1, c_2) \sim (c_1, b_2)$ which, empirically, could be false. Thus, it is not possible, in general, even if Axioms 1–4 are satisfied, to select elements satisfying Axiom 5. At the very least, the following law must hold (as was just shown):

5′. *If* $(a_1, b_2) \sim (b_1, a_2)$ *and* $(c_1, a_2) \sim (b_1, b_2) \sim (a_1, c_2)$, *then* $(c_1, b_2) \sim (b_1, c_2)$.

If we continued the selection process, we would soon discover yet more laws of the same type that must hold in order to satisfy Axiom 5.

So it is not satisfactory to propose as empirical laws either Axiom 5 or the simple statement that the sets A_1 and A_2 can be selected from some larger sets so that Axiom 5 holds. What is required for a satisfactory analysis is that one or more laws like 5′, as simple as possible, be found that guarantee the internal consistency of constructed equally spaced sequences. This problem is solved in Chapter 6 in a surprisingly satisfying way.

As is probably obvious, the equal-spacing notion is identical to that of a standard sequence; both terminologies exist in the literature.

1.4 CHOOSING AN AXIOM SYSTEM

The following discussion of axiomatization neither exhibits the spirit of nor uses the highly developed technical apparatus of mathematical logic. Chapter 18 is devoted to such a formal treatment. Here we touch upon only a few of the best-known logical features of axiom systems (Section 1.4.5) and otherwise we describe in nontechnical terms some of the types of axioms typically found in measurement systems.

1.4.1 Necessary Axioms

The previous discussion suggests that much of the effort in analyzing measurement goes into finding a good axiom system. It should be clear by now that at least one axiom is required to construct a representation. Specifically, if $\succsim$ is an arbitrary binary relation on A, then there need not be any homomorphism of $\langle A, \succsim \rangle$ into $\langle \text{Re}, \geqslant \rangle$. Indeed, if we suppose that such a homomorphism ϕ exists, then it follows that $\succsim$ is transitive (for if $a \succsim b$ and $b \succsim c$, then $\phi(a) \geqslant \phi(b) \geqslant \phi(c)$; hence $\phi(a) \geqslant \phi(c)$, which implies $a \succsim c$). Thus $\succsim$ is not arbitrary at all. We express this by saying that transitivity is a *necessary* axiom. "Necessary" here means mathematical, not practical, necessity. An axiom is necessary if it is a consequence of the existence of the homomorphism which we are trying to establish. Reflexivity is also necessary in the sense of being a consequence of the representation of $\langle A, \succsim \rangle$ in $\langle \text{Re}, \geqslant \rangle$, but we did not need to assume it as an axiom to prove Theorem 1 because it follows from the other axioms. Connectedness is also a necessary axiom for Theorem 1, and it was needed in the proof of Theorem 1.

For the representation of additive conjoint structures (Theorem 2) Axioms 1–4 are all necessary. Axioms 1 and 2 are necessary for the same reason as in Theorem 1 since the mapping $\phi(a) = \phi_1(a_1) + \phi_2(a_2)$ is a homomorphism of $\langle A, \succsim \rangle$ into $\langle \text{Re}, \geqslant \rangle$. To show that Axiom 3 is necessary, suppose that $(a_1, c_2) \succsim (b_1, c_2)$ and that Theorem 2 holds. Then $\phi_1(a_1) + \phi_2(c_2) \geqslant \phi_1(b_1) + \phi_2(c_2)$. Adding $\phi_2(d_2) - \phi_2(c_2)$ to both sides of the above inequality in Re yields $\phi_1(a_1) + \phi_2(d_2) \geqslant \phi_1(b_1) + \phi_2(d_2)$, and this implies $(a_1, d_2) \succsim (b_1, d_2)$. Similarly, Axiom 4 is necessary.

In contrast, Axiom 5 is not necessary. This is shown by the following example. Let $A_1 = A_2 = \{0, 1, 2, 4\}$, and define $\succsim$ on $A = A_1 \times A_2$ by: $(a_1, a_2) \succsim (b_1, b_2)$ iff $a_1 + a_2 \geqslant b_1 + b_2$. It is obvious that the representation part of Theorem 2 holds with ϕ_1 and ϕ_2 just the identity functions. Slightly less obvious is the fact that the uniqueness part of Theorem 2 is true. But clearly it is true for $\{0, 1, 2\}$, since that part is equally spaced; and since for any ϕ_1', ϕ_2' providing a representation,

$$\phi_1'(0) + \phi_2'(4) = \phi_1'(2) + \phi_2'(2) = \phi_1'(4) + \phi_2'(0),$$

we have $\phi_1'(4)$ and $\phi_2'(4)$ as linear combinations of lower values, and this extends the uniqueness theorem to them. Obviously, Axiom 5 is not true in this structure.

Each of the axiom systems in this book contains several fairly simple necessary axioms. We usually present these axioms first, sometimes discussing their intuitive meanings and their roles as empirical laws. Almost always, the proof that they are necessary is very simple and either is given at once or is omitted altogether. We have no rule for selecting the right set of necessary axioms; in general, it is a matter of trial and error or of insight.

We do not try to keep the number of axioms used to a bare minimum. The number of axioms is a rather misleading quantity anyway, since they can always be reduced to one axiom by stringing them all together by conjunctions. More realistically, it is often possible to hide a rather complex property within apparently simple axioms. If that property can be seen to be wrong—i.e., contrary to empirical fact—then we see no point in burying it in an apparently more innocent system. If the proposed representation is wrong, then it needs to be altered. We try to state our axioms so that they are conceptually distinct, even at the cost of increasing their number.

1.4.2 Nonnecessary Axioms

Nonnecessary axioms are frequently referred to as *structural* because they limit the set of structures satisfying the axiom system to something less than the set determined by the representation theorem. Three main types of structural axioms occur. First, some demand that the system be nontrivial in one sense or another—that a certain set be nonempty, that there be at least two nonequivalent elements, etc. These do not really limit the applicability of the theory because the structures excluded are of no empirical interest. Second, we occasionally assume that certain sets are finite or countable. Both Theorems 1 and 2 above are of this character. The fact that in neither case did we list finiteness as a separate axiom, but included it instead as a hypothesis of the representation–uniqueness theorem, is purely a matter of style; in a fully formalized theory, all of the assumptions would be listed as axioms. This stylistic device is used occasionally throughout the book. Finiteness is a real limitation. In each case, however, we present alternative theorems which replace finiteness with other axioms (see Section 1.4.3).

Structural axioms of the third type assert that solutions exist to certain classes of equations or inequalities; these are known as *solvability* axioms. For example, in systems for length measurement, two different solvability axioms sometimes are used. One postulates that the set of rods A is so "dense" that whenever $a \succ b$, then some c in A exists such that $a \succsim b \circ c$ (i.e., c solves an inequality) or even that $a \sim b \circ c$ (i.e., c solves an equation). It is easy to find examples of sets of rods where this does not hold, even though the representation and uniqueness theorems for extensive measurement of length are valid. Another type of solvability axiom asserts that the concatenation $a \circ b$ exists for, at least, certain pairs a, b in A. This type of axiom is sometimes well concealed. For example, one of the primitive relations may be taken to be a binary operation, in which case by definition of an operation $a \circ b$ exists for every a, b. Nevertheless, in a formalization that makes explicit all existential assumptions, it would appear as follows: there is a primitive ternary relation, which can be written $a \circ b = c$, with the property that for every a, b in A exactly one c exists in A such that $a \circ b = c$ holds. In other cases, where $a \circ b$ is not defined for all a and b, this kind of solvability is less well concealed. Nevertheless, it should always be considered as a specific assumption, for it plays much the same role in extensive measurement as does the other kind of solvability axiom in other measurement systems.

Axiom 5 for equally spaced, additive conjoint structures falls in none of the three classes, but in effect it is of the solvability type. If one were trying to *construct* a structure $\langle A_1 \times A_2, \succsim \rangle$ in which Axiom 5 held, one would proceed as indicated in Section 1.3.2, by choosing a_1, a_2, b_1, then selecting b_2 with $(a_1, b_2) \sim (b_1, a_2)$. To do so requires the assumption that $(a_1, b_2) \sim (b_1, a_2)$ can be solved for b_2. In Chapter 6, we replace Axiom 5 by solvability axioms of precisely that type together with necessary axioms similar to Axiom 5′.

We try to select structural properties that are as nonrestrictive as possible. Sometimes, alternative sets can be offered, covering different classes of structures. In particular, sometimes a trade can be effected in which the class of structures for which the representation is provable is enlarged (i.e., the structural axioms are weakened) at the expense of explicitly introducing additional necessary conditions which, in the presence of the former, stronger nonnecessary ones, had been deducible from the total axiom system. Such an exchange is deemed desirable when both an appreciable gain in applicability is effected and the added necessary conditions are neither too numerous nor too complex.

In many cases—especially in Chapters 3, 4, 5, 6, and 13, where we are dealing with additive representations of one kind or another—we have succeeded in limiting the structural properties to such an extent that we feel the remaining restrictions are of little practical import, i.e., they are quite likely to be acceptable empirically in many of the potential applications. In other cases—especially where nonlinear numerical structures are involved, as in Chapter 7—the structural restrictions are unsatisfactory, and much work needs to be done to weaken them. But even if strong structural axioms must be invoked, it is important to obtain axioms that are logically sufficient for the representation. If we only have necessary axioms, we remain unsure how to perform a thorough test of the representation.

1.4.3 Necessary and Sufficient Axiom Systems

The ubiquity of these nonnecessary restrictions may well seem puzzling. It seems far more desirable to find axiom systems composed entirely of necessary axioms that are also sufficient to prove the desired representation and uniqueness theorems. Such an axiom system is said to be *necessary and sufficient* for the representation. The advantage lies in the exclusion of examples such as on p. 22 where the additive conjoint representation holds, but Axiom 5 is violated.

As it happens, there are very few examples of what we consider satisfactory necessary and sufficient axiomatizations. Why is this? Loosely speaking, the reason is that the total set of structures admitting homomorphisms into a particular numerical structure is very heterogeneous and may include rather unusual and difficult-to-describe or pathological instances as well as more regular ones. Thus, the conditions which completely characterize such a set of structures are probably too complicated to be useful; in any event, they are not known. More systematic results, clarifying the above informal statement, are found in Chapter 9, Measurement Inequalities, and Chapter 18, Axiomatizability.

The requirement that the axiomatization be "satisfactory" is important (even though informal) because an unsatisfactory necessary and sufficient axiomatization always exists: take the representation and uniqueness theorems themselves as axioms. What criteria, then, do we impose on an axiomatization for it to be satisfactory? One demand is for the axioms to have a direct and easily understood meaning in terms of empirical operations, so simple that either they are evidently empirically true on intuitive grounds or it is evident how systematically to test them. In part, simplicity and clarity of meaning lie in the eye of the beholder. By the time you finish this book, some axioms may be clear which now might leave you aghast. Axiomatization is partly a search for simplicity and partly a restructuring of the axiomatizer's cognitive processes so that more things seem simple.

1.4.4 Archimedean Axioms

In addition to the types of axioms just described, a rather odd axiom is usually stated as part of each system. It is called *Archimedean* because it corresponds to the Archimedean property of real numbers: for any positive number x, no matter how small, and for any number y, no matter how large, there exists an integer n such that $nx \geqslant y$. This simply means that any two positive numbers are comparable, i.e., their ratio is not infinite. Another way to say this, one which generalizes more readily to qualitative structures, is that the set of integers n for which $y > nx$ is a finite set. For example, in extensive measurement, let a, $a \circ a = 2a$, $3a, \ldots$, be a standard sequence. Then the Archimedean axiom says that for any b, the set of integers n for which $b \succ na$ is finite. More generally, whenever we have defined a standard sequence, namely, entities having nonzero, equal spacing in the intended numerical representation, then we may always formulate the Archimedean property as: *every strictly bounded standard sequence is finite.*

It is evident that since the Archimedean property is true of the real numbers, it must also be true within the empirical relational system; it is a necessary axiom. What is surprising is that it is a needed axiom. In the few cases where the independence of axioms has been studied, the Archimedean axiom has been found to be independent of others; and no one seems to have suggested a more satisfactory substitute. It can be deleted if quite strong structural assumptions are made (see Section 6.11.1 and most of the systems in Pfanzagl, 1968), but with our relatively weak structural assumptions, we do not know how to eliminate it in favor of more desirable necessary axioms.

The objection to it as a necessary axiom is that either it is trivially true in a finite structure (that is why it was not stated in either Theorem 1 or 2) or it is unclear what constitutes empirical evidence against it since it may not be possible to exhibit an infinite standard sequence (see Section 1.5). Nonetheless, we can produce examples where it is violated, and these reveal something of the role it plays. Suppose that $\langle A_1 \times A_2, \succsim \rangle$ is a weak order in which any difference whatsoever on the first factor is decisive; the second factor matters only when a tie exists on the first. In such a case, any A_1-interval is infinitely large in comparison with any A_2-interval. If sufficiently large A_2 differences can compensate for small A_1 differences and vice versa, then the Archimedean axiom (for additive conjoint structures) seems reasonable. The difficulty in testing lies in deciding what evidence would be sufficient to conclude that a weak order on $A_1 \times A_2$ had the non-Archimedean character just described (see Section 1.5 and Chapters 17 and 18).

1.4.5 Consistency, Categoricalness, and Independence

All of the axiom systems we shall present have several nonisomorphic models in the real numbers. Therefore, the systems are consistent (i.e., something satisfies the axioms) and not categorical (i.e., two or more inherently different things satisfy them). The issue of independence of the axioms is more difficult. We have not knowingly included any axioms that are entirely derivable from the others in a system, but we may very well have done so inadvertently. In a few cases we establish independence formally (Chapters 3 and 6); but in many others, we are not sure that independence is met.

Axiom systems differ also in their general logical form and in the types of models they admit (cardinality, closure under submodels, etc.). There has been considerable work on such matters, including some recent results on the equivalence of different systems for finite models (Adams, Fagot, & Robinson, 1970); this work is discussed in Chapter 18.

1.5 EMPIRICAL TESTING OF A THEORY OF MEASUREMENT

Formal systems of measurement, although axiomatic, are not wholly or even largely evaluated on mathematical grounds. To be sure, we attempt to be explicit, precise, and consistent, and our proofs meet reasonable contemporary standards of rigor for informal set theory, but elegance and esthetics must give way, to a degree, to empirical criteria. The axioms purport to describe relations, perhaps idealized in some fashion, among certain potential observations, and adequacy of description is a more telling arbiter than beauty or simplicity. To carry out a satisfactory empirical evaluation of an axiom system is rather more difficult than it might first seem. Because some problems are very nearly universal, we sketch them briefly here.

1.5.1 Error of Measurement

The most pervasive problem is error—not human failure of one sort or another, but inherent features of the observational situation that cause us to fail to observe exactly what we wish to observe. As an example, suppose that we are judging qualitative weight by deflections of an equal-arm pan balance. When objects are placed in the two pans and do not cause a deflection of the arm from the horizontal, do we know that they have the same weight? In an operational sense relative to that balance, we do, but in some idealized sense we may doubt that we do. The contact between the knife edge and the arm exhibits some friction that makes the balance less than perfectly sensitive to what is placed in its pans. Moreover, the condition of the point of contact may vary over time as the result of movements and electrochemical effects, and so the amount of friction may fluctuate in some irregular way from observation to observation. Thus, we suspect that when two weights differ by an amount just at the edge of sensitivity of the balance, repeated observations may not yield the same results. Even when we avoid this boundary region of random error, we may still find evidence for systematic errors. For example, we may find that a sequence of weights has the property that each successive pair is judged equivalent in weight, but the first and last ones of the sequence are definitely not equivalent. Clearly then, the observed relation is not a weak order—in particular, the indifference relation $\sim$ is not transitive—and so the order cannot be represented in terms of $\geqslant$ in as simple a way as in Theorem 1. On the other hand, experience has shown that when we have such a sequence and when the observational conditions are improved by constructing a more sensitive balance, then at least one of the original equivalences is converted into a nonequivalence and, to a better approximation, the weak-order properties are satisfied. Of course, we can select a new set of objects that exhibits a refined version of the same phenomenon on the new balance. Nevertheless, the pattern of improved approximations is such that we elect to retain the assumption of an underlying weak ordering of weight and to say that, in any particular set of observations, there are systematic errors due to imperfections in the observational situation.

The existence of error also has implications for the construction of measurement scales. As we have indicated in Sections 1.1.2 and 1.3.2, such constructions most often involve a counting-of-units procedure based on some appropriate definition of standard sequence. But any definition of standard sequence involves finding exact replicas of a given object or interval. In the presence of error, we are confronted with two choices. The first, which often is the solution adopted, is to equate the replicas by a method that is much more expensive, but much more precise, than the method that is used to make comparisons in the field. For example, a good meter stick, calibrated in millimeters, should have the property that, if a comparison object falls clearly between the marks x and $x + 1$ millimeters, then with high probability its true length actually lies between x and $x + 1$ millimeters. This will happen if the procedure by which the millimeter steps are equated yields a standard deviation not exceeding 10^{-3} mm or one micron. The standard deviation of a sequence of x steps, $x < 10^3$, will be about $(10^3)^{1/2} \cdot 10^{-3} = 10^{-3/2}$ mm. Thus, an error of more than 10^{-1} mm will be extremely improbable (three standard deviations); and so an object that is clearly between the marks x and $x + 1$, by visual comparison, is quite likely to be actually between x and $x + 1$ mm long. If an object falls "right on" the x mark, we will be quite uncertain whether its length is more or less than x mm, but with high probability it will lie in the interval $(x - \frac{1}{2}, x + \frac{1}{2})$.

The second solution to the problem is to dispense with exact standard sequences and to use only clear-cut observations of inequality to construct approximate standard sequences. A good example is found in Section 4.4.4 on difference measurement. Some of the inequality observations may be inferred rather than observed directly. For example, if a is clearly greater than c, but b seems to be indifferent to c, we infer that a is also greater than b. In order that such inferences yield a weak order, the clear-cut observations must satisfy the axioms of a semiorder (Chapter 15).

Obviously, a subtle interplay obtains among observations, theory, and refined observations, in which the theory is both tested and used in a normative fashion to define the existence and nature of error. It is probably not possible at present to formulate generally the exact conditions that lead us to attribute a discrepancy between theory and observation to error rather than to an inadequacy in the theory. As explicit error theories are developed to accompany measurement theories, it should become easier to make these decisions more routine. Today, however, few error theories exist; what we know about them is described in Chapters 15–17.

1.5.2 Selection of Objects in Tests of Axioms

A second ubiquitous problem of testing is that most theories are stated for large, often infinite sets of elements, whereas empirical tests usually involve small finite subsets. The general problem of inductive generalization

from limited data arises in testing all scientific theories, but there are some special problems connected with most measurement axiom systems.

First, some axioms may be easier to disconfirm than others. For example, transitivity can be disconfirmed if there is a single intransitive triple, $a \succsim b$, $b \succsim c$, $c \succ a$. Usually the empirical interpretation of $\succsim$ is such that it is possible to decide, for given a, b, whether $a \succsim b$, or at least, whether $a \succsim b$ is extremely probable. With such an empirical interpretation, transitivity can be unequivocally rejected or, at least, assigned a very low probability. On the other hand, the Archimedean axiom cannot be rejected merely because $b \succ a$, $b \succ 2a, \ldots, b \succ ma$. There may be some $n > m$, for which $na \succsim b$. One may eventually consider it improbable that a large enough n will ever be found and, thus, reject the Archimedean axiom—partly because it may seem not to make scientific sense to go on searching for n large enough and partly because a more attractive and manageable theory may result from a non-Archimedean representation. Alternatively, there may be other empirical interpretations of $\succsim$, e.g., where underlying rules for generating $\succsim$ are directly observable, which enable us to reject the Archimedean axiom. One must keep in mind the fact that the refutability of axioms depends both on their mathematical form and on their empirical interpretation.

Nonnecessary axioms are usually not tested. If the elements are thought to exhibit some kind of fine grainedness, then one's belief in the existence of solutions to inequalities may be extremely strong. For example, if rod b is longer than rod c, one assumes without much question that a small rod d can be found such that $b \succsim c \circ d$. Frequently one also accepts, as an idealization, the "continuity" (mathematically, connectedness in the order topology) of the domain of objects, in which case solvability of equations is also accepted as a consequent idealization. However, if there is real doubt about the matter, solvability axioms can offer a difficulty similar to that sometimes encountered with Archimedean axioms: the mere fact that a solution has not yet been found may or may not lead one to believe that one never will be found, no matter what objects are tested.

Second, some axioms are more difficult to confirm than others. If we fail to disconfirm an axiom, we need to ask a question akin to that of the statistical power of the test: Did we select the elements in such a way that the data had some chance of showing the axiom wrong if, in fact, it is wrong? For example, if a is much greater than b, and b is much greater than c, it does not surprise us that a is greater than c. A more convincing test of transitivity is obtained by selecting triples a, b, c that are likely to violate it, if it is indeed false. Sometimes, an alternative theory can be a guide (see Tversky, 1969, for the use of an alternative theory to locate violations of transitivity). If we choose a slightly greater than b and b slightly greater than c, we may be more gratified to find $a \succ c$; but then, there is more danger of falsely rejecting transitivity, as a result of errors of measurement.

Another manifestation of the problem of selection of objects in confirming axioms is the fact that, even though none of the necessary axioms of a system are disconfirmed by a given set of data, and even though the structural axioms of the system are assumed *a priori*, nevertheless, other necessary consequences of the representation, which were not needed to prove the representation theorem, may be disconfirmed by those same data. Because this seems almost contradictory, we amplify the point.

In most theories of measurement, several necessary axioms along with a few nonnecessary ones are shown to be sufficient for the numerical representation (homomorphism) to exist. From the representation, other necessary properties follow which, of course, do not need to be listed among the axioms, since they are deducible from them. Now, a particular set of data may not disconfirm any axiom of the theory; however, an axiom such as solvability may be false if attention is restricted just to that subset of objects tested: the solution to some inequality or equation may lie outside that subset. In fact, we may have accepted solvability to begin with because of the fine grainedness of the entire object set. Since the axiom system as a whole (including solvability) does not hold for the particular subset sampled, there is no mathematical reason why the unneeded necessary property must hold for that subset, even if none of the needed necessary axioms is disconfirmed for that subset. The disconfirmation of the necessary but unneeded property points, in fact, to an inadequate selection of objects for testing the necessary axioms. Since solvability presumably holds in the large structure, one of the tested necessary axioms would in fact be disconfirmed elsewhere, if suitable objects were selected. (For an example, see Section 9.1.)

These remarks are intended to point up the need for thoroughness in sampling objects before accepting any particular set of axioms as probably being satisfied. Another conclusion, which is tempting but overhasty, is that one should carry out all possible indirect tests of the axiom system by testing as many consequences of the system as possible in a given set of data. Since the ultimate consequence is the representation theorem itself, why not test whether a representation can be constructed for the sample at hand? There are two reasons why such a conclusion is unwarranted. The first has to do with fallibility of data. The more tests that are performed, the greater the chance that one of them will fail due to sampling error. This must be compensated for by relaxing the criterion for disconfirmation; but then, there may be an excellent chance of failing to reject the axiom system when it is in fact systematically wrong.

The existence of a numerical representation, for a fixed sample, generally

corresponds to the existence of a solution to a large set of simultaneous inequalities (see Section 1.1.3 or Chapter 9). The sets of inequalities that arise in this way, in practice, rarely have solutions; but if one relaxes the criterion, accepting a "solution" that solves most of the inequalities, a "solution" may very well exist even when there is some systematic failure of one of the axioms.

The second danger in testing all necessary consequences of the representation by trying to construct the representation for a sample is that failure tends not to be instructive. Direct tests of particular axioms, on the other hand, are often very informative, since they can easily fail in a systematic way. Consider, for example, the independence axioms of additive conjoint measurement (Section 1.3.2). If both $(a_1, c_2) \succ (b_1, c_2)$ and $(a_1, d_2) \prec (b_1, d_2)$, we may be able to subclassify or order the A_2-factor so that the first inequality holds for values c_2 in certain classes or at one end of the dimension, and the reverse inequality holds in other classes or at the end of the dimension near d_2. This kind of systematic rejection greatly alleviates statistical problems in rejecting the axiom: we can be surer that the rejection is not due to sampling error if an alternative hypothesis is shown to fit the data very well. Moreover, such systematic rejection tells us a good deal about what is wrong, and it may suggest either other measurement schemes that will work or a different choice of basic factors.

Thus, one value of a satisfactory axiomatization is that it provides a set of relatively simple, conceptually distinct, empirically testable conditions to be tested. The problems of error and of selection of objects to be tested have no easy solutions, however; they must be tackled with whatever experimental and statistical tools are available.

1.6 ROLES OF THEORIES OF MEASUREMENT IN THE SCIENCES

As measurement surely plays an essential role in all science, one might anticipate great interest attaching to theories of measurement. This is not true, however, in much of contemporary physics. Some exists in applied physics—mechanics, thermodynamics, hydrodynamics, etc.—because the methods of dimensional analysis depend explicitly on properties of physical measures, and some also arises in connection with questions deep in the foundations of quantum theory and the theory of relativity. But for the most part, questions about physical measurement are regarded as being in the province of philosophy of physics, not in physics itself. Usually, the measurability of the variables of interest in physics is taken for granted and the actual measurements are reduced, via the elaborate superstructure of physical theory, to comparatively indirect observations. The construction and calibration of measuring devices is a major activity, but it lies rather far from the sorts of qualitative theories we examine here.

Other sciences, especially those having to do with human beings, approach measurement with considerably less confidence. In the behavioral and social sciences we are not entirely certain which variables can be measured nor which theories really apply to those we believe to be measurable; and we do not have a superstructure of well-established theory that can be used to devise practical schemes of measurement. For these reasons, the analysis of measurement and the construction of new systems of measurement have been an active preoccupation of some behavioral scientists. Included is some work that is highly sophisticated, and some that is remarkably naive.

A recurrent temptation when we need to measure an attribute of interest is to try to avoid the difficult theoretical and empirical issues posed by fundamental measurement by substituting some easily measured physical quantity that is believed to be strongly correlated with the attribute in question: hours of deprivation in lieu of hunger; skin resistance in lieu of anxiety; milliamperes of current in lieu of aversiveness, etc. Doubtless this is a sensible thing to do when no deep analysis is available, and in all likelihood some such indirect measures will one day serve very effectively when the basic attributes are well understood, but to treat them now as objective definitions of unanalyzed concepts is a form of misplaced operationalism.

Little seems possible in the way of a careful analysis of an attribute until means are devised to say which of two objects or events exhibits more of the attribute. Once we are able to order the objects in an acceptable way, we need to examine them for additional structure, for example, by selecting two or more factors that affect the ordering. Then begins the search for qualitative laws satisfied by the ordering and the additional structure. In contrast to fundamental physical measurement, which is typically one-dimensional (see, however, Chapter 10), many of the theories of measurement that appear applicable to behavioral problems are inherently multidimensional, and so the measurement theories deal simultaneously with several measures and the laws connecting them. These theories suggest new qualitative laws to be tested, and even when they are found to be wrong, much may be learned if the violations are systematic. Moreover, these theories lead to selection among the many factors that might be relevant by focusing attention on those variables that enter into simple qualitative laws.

The work on fundamental measurement representations, which is relatively recent in the behavioral sciences, contrasts with an older research field known as psychometrics and scaling theory. Most of the psychometric literature is based on numerical rather than qualitative relations (e.g., matrices of correlation coefficients, test profiles, choice probabilities), although there is a tradition, which has recently grown considerably, focusing on ordinal relations. This work aims to represent such relations by numerical relations, mostly of a geometric nature, that are more compact and more revealing than the input data. Among the unidimensional methods are Thurstonian scaling (Bock & Jones, 1968; Thurstone, 1959; Torgerson, 1958) and test theory (Lord and Novick, 1968); and among the multidimensional methods are the classic bilinear models of factor analysis (Harman, 1967; Spearman, 1927; Thurstone, 1947) and the ordinal procedures of Coombs (1964), Guttman (1944, 1968), and Shepard (1966). Most of these scaling procedures assume the validity of the proposed model and produce a best-fitting numerical representation of the data, whether or not the assumed model is really appropriate.

Here, by contrast, we are concerned almost exclusively with the qualitative conditions under which a particular representation holds. To some extent, therefore, theories of measurement may be regarded as complementary to the methods of scaling, with the former being concerned with empirical laws (axioms) that make a particular type of numerical representation appropriate and the latter with methods for finding a numerical representation of a particular type. This seeming complementarity is, however, somewhat illusory because the bulk of the scaling literature involves mapping one numerical structure into another one rather than a qualitative structure into a numerical one. For example, in scaling aptitude, intelligence, or social attitudes, test scores or numerical ratings are usually interpreted as measures of the attribute in question. But in the absence of a well-defined homomorphism between an empirical and a numerical relational structure, it is far from clear how to interpret such numbers. We return to this issue in Chapter 20.

The clearest complementarity exists with the ordinal scaling methods, which lately have become one of the main foci of scaling research (partly because the widespread availability of fast computers has made them practical). In fact, it was the earlier work on ordinal multidimensional scaling (e.g., Coombs, 1964; Shepard, 1966) that motivated the reworking of the foundations of geometry which is presented in Chapter 13. Such axiomatizations play the important role of showing how to test whether a particular scaling method is at all justified, and it invites the search for systematic departures from the axioms.

1.7 PLAN OF THE BOOK

Much of the book develops and demonstrates the theme stated in Sections 1.1 and 1.2 that, although many different empirical relational structures and many different axiom systems lead to measurement, the procedure for obtaining the numbers always reduces to one of three basic methods. Chapter 2 is the mathematical pivot which provides proofs, mainly using constructive methods, of a series of rigorous isomorphism theorems which amount to showing that, under suitable assumptions, the procedures outlined in Section 1.1 do give internally consistent numerical answers. In Chapters 3–9 the representation and uniqueness theorems are reduced to applications of the theorems in Chapter 2.

If you intend to skip proofs you need not read Chapter 2 in detail; however, you should go over Section 1.1 carefully to gain a good intuitive idea of how numerical scales are constructed. Then in reading a later chapter, you should try to see how the appropriate method of Section 1.1 is applied to the situation at hand. To understand this, you may find it useful to scan the statements of the theorems in Chapter 2.

Chapters 3–8 all depend on the counting-of-units procedure outlined in Section 1.1.2 and formulated more rigorously in Section 2.2. In Chapter 3, counting of units arises directly because the empirical relational structure contains a concatenation operation. This is extensive measurement. Some special variants arise in connection with problems of relativity and thermodynamics. The first half of the probability chapter (Chapter 5) may be considered as another variant of extensive measurement in which the union of disjoint events plays the role of concatenation. The latter half of Chapter 5 and Chapters 6 and 8 use the device discussed in Section 1.3.2, i.e., counting off equal units by laying equal intervals end to end, where equality of intervals is defined in terms of balancing by a single interval on another factor. Chapter 4 (Difference Measurement) studies this counting device in pure form; many results in later chapters reduce to those in Chapter 4. Chapter 6 deals with additive conjoint measurement, and Chapter 8 applies the results of Chapter 6 to expected-utility measurement. Still a third variant of the counting-of-units process is used for combined additive-multiplicative (polynomial) conjoint measurement studied in Chapter 7.

Some of the topics of Chapters 3–8 are reconsidered in Chapter 9 in terms of the solution-of-inequalities method (Sections 1.1.3 and 2.3) instead of the counting-of-units method. The results are primarily concerned with finite structures.

The final chapter of Volume I attempts to construct a bridge between fundamental measurement and dimensional analysis, and it includes formulations of the qualitative equivalents of numerical laws satisfied by fundamentally measured variables.

In contrast to the relative unity of Volume I, the ten chapters of Volume II are more diverse. The first four deal with geometric representations. Since geometry is by far the earliest and most far-reaching example of measurement, we have included in Chapter 11 and 12 a general discussion of geometric structures and an overview of classical foundations of geometry, seen as measurement theory. A self-contained treatment of this classical theory is, of course, beyond the scope of this book. Chapter 13 presents a new approach to the foundations of geometry which is based on the ordering of distances as a primitive notion. The representation theorems rely on the theories of extensive, difference, and additive conjoint measurement developed in Chapters 3, 4, and 6, respectively, of Volume I. Chapter 14 deals with one of the best-developed measurement systems outside of the physical sciences, namely color measurement. The representation is geometric, but it is unlike any other in the book.

The next three chapters approach the problem of error of measurement in two very different ways. Chapters 15 and 16 deal with empirical relational structures in which error is incorporated directly and is dealt with axiomatically. Chapter 17 presents some statistical methods for testing theories of measurement of the kind discussed earlier in the book, where error is treated as an extraneous phenomenon.

Next we deal with two philosophical issues: the (logical) problem of axiomatizability (Chapter 18) and the relationship between uniqueness theorems and the meaningfulness of statements involving numerical measurements (Chapter 19).

Finally, Chapter 20 sums up the approach to measurement embodied in the rest of the book and compares it with other approaches.

EXERCISES[4]

1. Show that the inequalities in Equation (2) imply $1 < x_3/x_2 < \frac{3}{2}$. (1.1.3)

2. Suppose that $P(x, y)$ denotes the proportion of times that x is chosen over y in a preference experiment. Define $x \gtrsim y$ iff $P(x, y) \geqslant \frac{1}{2}$. When does this yield a weak order (Definition 1)? (1.3.1)

3. Suppose that $\gtrsim$ is defined as in Exercise 2 and that $P(x, y)$ is given for all distinct pairs in the set $A = \{a, b, c, d\}$ by the values in the following matrix.

$x \backslash y$	a	b	c	d
a	—	.72	.65	.67
b	.28	—	.39	.32
c	.35	.61	—	.40
d	.33	.68	.60	—

[4] The directly relevant sections of Chapter 1 are listed in parentheses at the end of each exercise.

Verify that $\langle A, \succsim \rangle$ is a weak order (assuming $x \succsim x$ for all x). (1.3.1)

4. Let $\langle A, \succsim \rangle$ be a weak order. Show that the symmetric part $\sim$ (Definition 2) is an equivalence relation and that the asymmetric part $\succ$ is transitive and asymmetric. Show that $\succsim$ on $A/\sim$ is a simple order. (1.3.1)

5. Construct scale values $\phi(a)$, $\phi(b)$, $\phi(c)$, $\phi(d)$ successively, in that order, using the data in Exercise 3, by the method of Section 1.3.1.

6. Construct $\phi'(a)$, $\phi'(b)$, $\phi'(c)$, $\phi'(d)$ successively, in that order, using the data in Exercise 3, by the method sketched in Section 1.1.1.

7. Construct a strictly increasing real-valued function h from Re onto Re such that $h[\phi(x)] = \phi'(x)$ for all x in A, where ϕ is the scale of Exercise 5 and ϕ' is the scale of Exercise 6.

8. Suppose that the ordinal-scale values have been assigned to all elements of $A_1 \times A_2$, where $A_1 = \{a_1, b_1, c_1, d_1\}$, $A_2 = \{a_2, b_2, c_2, d_2\}$, as given in the following matrix.

$A_1 \backslash A_2$	a_2	b_2	c_2	d_2
a_1	5	1	13	29
b_1	29	13	61	125
c_1	61	29	125	253
d_1	13	5	29	61

Verify that Axioms 3 and 4 of independent conjoint structures (Definition 3) are satisfied. Determine the weak orderings $\succsim_1$ and $\succsim_2$ on A_1 and A_2, respectively. (1.3.2)

9. Verify that Axiom 5 of Definition 5 holds for the matrix of Exercise 8. (1.3.2)

10. Construct functions ϕ_1, ϕ_2 on A_1, A_2 that satisfy the requirements of Theorem 2, for the data of Exercise 8. What is the relationship between the sums $\phi_1 + \phi_2$ and the numbers in the matrix? (1.3.2)

11. Let $\langle A_1 \times A_2, \succsim \rangle$ be an independent conjoint structure. Show that $\langle A_i, \succsim_i \rangle$, $i = 1, 2$, is a weak order (Definition 4). (1.3.2)

12. Suppose that $\langle A_1 \times A_2, \succsim \rangle$ is a finite, equally spaced, additive conjoint structure (Definition 5). Assume that the weak orders $\succsim_1$ and $\succsim_2$ are simple orders, and label the elements of A_1 as $a_1^{(i)}$, $i = 1, \ldots, m$, with

$$a_1^{(m)} \succ_1 a_1^{(m-1)} \succ_1 \cdots \succ_1 a_1^{(1)}.$$

Similarly, let $A_2 = \{a_2^{(j)} \mid j = 1, \ldots, n\}$, with $a_2^{(j+1)} \succ_2 a_2^{(j)}$.

(a) Use Axiom 5, and mathematical induction, to prove that if $i + j = k + l$, then $(a_1^{(i)}, a_2^{(j)}) \sim (a_1^{(k)}, a_2^{(l)})$.

(b) Use the result of (a), plus Axioms 3 and 4, to show that if $i + j > k + l$, then $(a_1^{(i)}, a_2^{(j)}) \succ (a_1^{(k)}, a_2^{(l)})$.

(c) Show that $\phi_1(a_1^{(i)}) = i$, $\phi_2(a_2^{(j)}) = j$ satisfy the representation theorem (Theorem 2). (1.3.2)

13. Extend the result of Exercise 12 to the case where $\succsim_1, \succsim_2$ need not be antisymmetric by using equivalence classes (Definition 2) with respect to $\sim_1$ and $\sim_2$. (1.3.1, 1.3.2)

14. Prove the uniqueness theorem for finite, equally spaced, additive conjoint structures by showing that if $\phi_1'(a_1^{(1)}) = \sigma_1$, $\phi_2'(a_2^{(1)}) = \sigma_2$ and $\phi_1'(a_1^{(2)}) = \tau$, then

$$\phi_1' = (\tau - \sigma_1)(\phi_1 - 1) + \sigma_1,$$
$$\phi_2' = (\tau - \sigma_1)(\phi_2 - 1) + \sigma_2.$$

Do this by using the results of Exercise 12(a). (1.3.2)