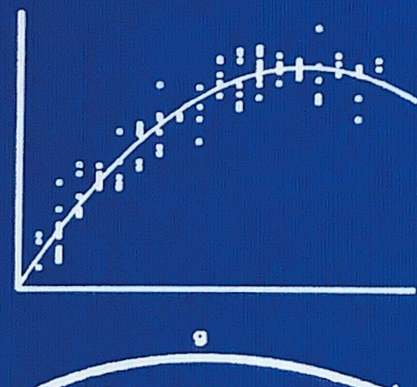
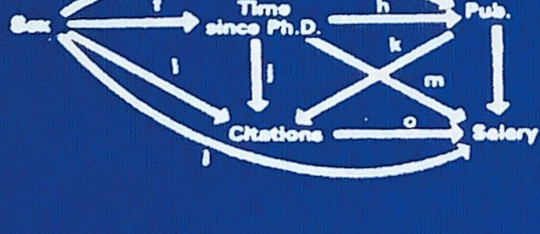
Applied Multiple Regression/ Correlation Analysis for the Behavioral Sciences

Third Edition





Jacob Cohen
Patricia Cohen
Stephen G. West
Leona S. Aiken

1

Introduction

1.1 MULTIPLE REGRESSION/CORRELATION AS A GENERAL DATA-ANALYTIC SYSTEM

1.1.1 Overview

Multiple regression/correlation analysis (MRC) is a highly general and therefore very flexible data analytic system. Basic MRC may be used whenever a quantitative variable, the dependent variable (Y), is to be studied as a function of, or in relationship to, any factors of interest, the independent variables (IVs).[1] The broad sweep of this statement is quite intentional.

1. The form of the relationship is not constrained: it may be simple or complex, for example, straight line or curvilinear, general or conditional, or combinations of these possibilities.

2. The nature of the research factors expressed as independent variables is also not constrained. They may be quantitative or qualitative, main effects or interactions in the analysis of variance (ANOVA) sense, or covariates in the analysis of covariance (ANCOVA) sense. They may be correlated with each other or uncorrelated as in balanced factorial designs in ANOVA commonly found in laboratory experiments. They may be naturally occurring ("organismic" variables) like sex, diagnosis, IQ, extroversion, or years of education, or they may be planned experimental manipulations (treatment conditions). In short, virtually any information whose bearing on the dependent variable is of interest may be expressed as research factors.

3. The nature of the dependent variable is also not constrained. Although MRC was originally developed for scaled dependent variables, extensions of the basic model now permit appropriate analysis of the full range of dependent variables including those that are of the form of categories (e.g., ill vs. not ill) or ordered categories.

4. Like all statistical analyses, the basic MRC model makes assumptions about the nature of the data that are being analyzed and is most confidently conducted with "well-behaved" data that meet the underlying assumptions of the basic model. Statistical and graphical methods now part of many statistical packages make it easy for the researcher to determine whether

[1] In this book we typically employ Y to indicate a dependent variable and IV to represent an independent variable to indicate their role in the statistical analysis without any *necessary* implication of the existence or direction of causal relationship between them.

estimates generated by the basic MRC model are likely to be misleading and to take appropriate actions. Extensions of the basic MRC model include appropriate techniques for handling "badly behaved" or missing data and other data problems encountered by researchers.

The MRC system presented in this book has other properties that make it a powerful analytic tool. It yields measures of the magnitude of the total effect of a factor on the dependent variable as well as of its partial (unique, net) relationship, that is, its relationship over and above that of other research factors. It also comes fully equipped with the necessary apparatus for statistical hypothesis testing, estimation, construction of confidence intervals, and power analysis. Graphical techniques allow clear depictions of the data and of the analytic results. Last, but certainly not least, MRC is a major tool in the methods of causal (path, structural equation) analysis. Thus, MRC is a versatile, all-purpose system of analyzing the data over a wide range of sciences and technologies.

1.1.2 Testing Hypotheses Using Multiple Regression/Correlation: Some Examples

Multiple regression analysis is broadly applicable to hypotheses generated by researchers in the behavioral sciences, health sciences, education, and business. These hypotheses may come from formal theory, previous research, or simply scientific hunches. Consider the following hypotheses chosen from a variety of research areas:

1. In health sciences, Rahe, Mahan, and Arthur (1970) hypothesized that the amount of major life stress experienced by an individual is positively related to the number of days of illness that person will experience during the following 6 months.
2. In sociology, England, Farkas, Kilbourne, and Dou (1988) predicted that the size of the positive relationship between the number of years of job experience and workers' salaries would depend on the percentage of female workers in the occupation. Occupations with a higher percentage of female workers were expected to have smaller increases in workers' salaries than occupations with a smaller percentage of female workers.
3. In educational policy, there is strong interest in comparing the achievement of students who attend public vs. private schools (Coleman, Hoffer, & Kilgore, 1982; Lee & Bryk, 1989). In comparing these two "treatments" it is important to control statistically for a number of background characteristics of the students such as prior academic achievement, IQ, race, and family income.
4. In experimental psychology, Yerkes and Dodson (1908) proposed a classic "law" that performance has an inverted U-shaped relationship to physiological arousal. The point at which maximum performance occurs is determined by the difficulty of the task.
5. In health sciences, Aiken, West, Woodward, and Reno (1994) developed a predictive model of women's compliance versus noncompliance (a binary outcome) with recommendations for screening mammography. They were interested in the ability of a set of health beliefs (perceived severity of breast cancer, perceived susceptibility to breast cancer, perceived benefits of mammography, perceived barriers to mammography) to predict compliance over and above several other sets of variables: demographics, family medical history, medical input, and prior knowledge.

Each of these hypotheses proposes some form of relationship between one or more factors of interest (independent variables) and an outcome (dependent) variable. There are usually other variables whose effects also need to be considered, for reasons we will be discussing in this text. This book strongly emphasizes the critical role of theory in planning MRC analyses. The researcher's task is to develop a statistical model that will accurately estimate the relationships among the variables. Then the power of MRC analysis can be brought to bear to test the hypotheses and provide estimations of the size of the effects. However, this task cannot be carried out well if the actual data are not evaluated with regard to the assumptions of the statistical model.

1.1.3 Multiple Regression/Correlation in Prediction Models

Other applications of MRC exist as well. MRC can be used in practical prediction problems where the goal is to forecast an outcome based on data that were collected earlier. For example, a college admissions committee might be interested in predicting college GPA based on high school grades, college entrance examination (SAT or ACT) scores, and ratings of students by high school teachers. In the absence of prior research or theory, MRC can be used in a purely exploratory fashion to identify a collection of variables that strongly predict an outcome variable. For example, coding of the court records for a large city could identify a number of characteristics of felony court cases (e.g., crime characteristics, defendant demographics, drug involvement, crime location, nature of legal representation) that might predict the length of sentence. MRC can be used to identify a minimum set of variables that yield the best prediction of the criterion for the data that have been collected (A. J. Miller, 1990). Of course, because this method will inevitably capitalize on chance relationships in the original data set, replication in a new sample will be critical. Although we will address purely predictive applications of MRC in this book, our focus will be on the MRC techniques that are most useful in the testing of scientific hypotheses.

In this chapter, we initially consider several issues that are associated with the application of MRC in the behavioral sciences. Some disciplines within the behavioral sciences (e.g., experimental psychology) have had a misperception that MRC is only suitable for nonexperimental research. We consider how this misperception arose historically, note that MRC yields identical statistical tests to those provided by ANOVA yet additionally provides several useful measures of the size of the effect. We also note some of the persisting differences in data-analytic philosophy that are associated with researchers using MRC rather than ANOVA. We then consider how the MRC model nicely matches the complexity and variety of relationships commonly observed in the behavioral sciences. Several independent variables may be expected to influence the dependent variable, the independent variables themselves may be related, the independent variables may take different forms (e.g., rating scales or categorical judgments), and the form of the relationship between the independent and dependent variables may also be complex. Each of these complexities is nicely addressed by the MRC model. Finally, we consider the meaning of causality in the behavioral sciences and the meanings of control. Included in this section is a discussion of how MRC and related techniques can help rule out at least some explanations of the observed relationships. We encourage readers to consider these issues at the beginning of their study of the MRC approach and then to reconsider them at the end.

We then describe the orientation and contents of the book. It is oriented toward practical data analysis problems and so is generally nonmathematical and applied. We strongly encourage readers to work through the solved problems, to take full advantage of the programs for three major computer packages and data sets included with the book, and, most important, to learn MRC by applying these techniques to their own data. Finally, we provide a brief overview of the content of the book, outlining the central questions that are the focus of each chapter.

1.2 A COMPARISON OF MULTIPLE REGRESSION/CORRELATION AND ANALYSIS OF VARIANCE APPROACHES

MRC, ANOVA, and ANCOVA are each special cases of the *general linear model* in mathematical statistics.[2] The description of MRC in this book includes extensions of conventional MRC analysis to the point where it is essentially equivalent to the general linear model. It thus follows that any data analyzable by ANOVA/ANCOVA may be analyzed by MRC, whereas the reverse is not the case. For example, research designs that study how a scaled characteristic of participants (e.g., IQ) and an experimental manipulation (e.g., structured vs. unstructured tasks) jointly influence the subjects' responses (e.g., task performance) cannot readily be fit into the ANOVA framework. Even experiments with factorial designs with unequal cell sample sizes present complexities for ANOVA approaches because of the nonindependence of the factors, and standard computer programs now use a regression approach to estimate effects in such cases. The latter chapters of the book will extend the basic MRC model still further to include alternative statistical methods of estimating relationships.

1.2.1 Historical Background

Historically, MRC arose in the biological and behavioral sciences around 1900 in the study of the natural covariation of observed characteristics of samples of subjects, including Galton's studies of the relationship between the heights of fathers and sons and Pearson's and Yule's work on educational issues (Yule, 1911). Somewhat later, ANOVA/ANCOVA grew out of the analysis of agricultural data produced by the controlled variation of treatment conditions in manipulative experiments. It is noteworthy that Fisher's initial statistical work in this area emphasized the multiple regression framework because of its generality (see Tatsuoka, 1993). However, multiple regression was often computationally intractable in the precomputer era: computations that take milliseconds by computer required weeks or even months to do by hand. This led Fisher to develop the computationally simpler, equal (or proportional) sample size ANOVA/ANCOVA model, which is particularly applicable to planned experiments. Thus multiple regression and ANOVA/ANCOVA approaches developed in parallel and, from the perspective of the substantive researchers who used them, largely independently. Indeed, in certain disciplines such as psychology and education, the association of MRC with nonexperimental, observational, and survey research led some scientists to perceive MRC to be less scientifically respectable than ANOVA/ANCOVA, which was associated with experiments.

Close examination suggests that this guilt (or virtue) by association is unwarranted—the result of the confusion of data-analytic method with the logical considerations that govern the inference of causality. Experiments in which different treatments are applied to randomly assigned groups of subjects and there is no loss (attrition) of subjects permit unambiguous inference of causality; the observation of associations among variables in a group of randomly selected subjects does not. Thus, interpretation of a finding of superior early school achievement of children who participate in Head Start programs compared to nonparticipating children depends on the design of the investigation (Shadish, Cook, & Campbell, 2002; West, Biesanz, & Pitts, 2000). For the investigator who randomly assigns children to Head Start versus Control programs, attribution of the effect to program content is straightforward. For the investigator who simply observes whether children whose parents select Head Start programs have higher school achievement than those who do not, causal inference becomes less certain. Many other possible differences (e.g., child IQ; parent education) may exist between

[2] For the technically minded, our primary focus will be on the "fixed" version of these models, representing the most common usage of the general linear model in the behavioral sciences.

the two groups of children that could potentially account for any findings. But each of the investigative teams may analyze their data using either ANOVA (or equivalently a t test of the mean difference in school achievement) or MRC (a simple one-predictor regression analysis of school achievement as a function of Head Start attendance with its identical t test). The logical status of causal inference is a function of how the data were produced, not how they were analyzed (see further discussion in several chapters, especially in Chapter 12).

1.2.2 Hypothesis Testing and Effect Sizes

Any relationship we observe, whether between independent variables (treatments) and an outcome in an experiment or between independent variables and a "dependent" variable in an observational study, can be characterized in terms of the strength of the relationship or its effect size (ES). We can ask how much of the total variation in the dependent variable is produced by or associated with the independent variables we are studying. One of the most attractive features of MRC is its automatic provision of regression coefficients, proportion of variance, and correlational measures of various kinds, all of which are kinds of ES measures. We venture the assertion that, despite the preoccupation of the behavioral sciences, the health sciences, education, and business with quantitative methods, the level of consciousness in many areas about strength of observed relationships is at a surprisingly low level. This is because concern about the statistical significance of effects has tended to pre-empt attention to their magnitude (Harlow, Mulaik, & Steiger, 1997). Statistical significance only provides information about whether the relationship exists at all, often a question of trivial scientific interest, as has been pointed out in several commentaries (e.g., J. Cohen, 1994; Meehl, 1967). The level of statistical significance reflects the sample size, incidental features of the design, the sampling of cases, and the nature of the measurement of the dependent variable; it provides only a very pale reflection of the effect size. Yet many research reports, at least implicitly, confuse the issues of effect size and level of statistical significance, using the latter as if it meant the former (Gigerenzer, 1993).

Part of the reason for this unfortunate tendency is that traditional ANOVA/ANCOVA yields readily interpretable F and t ratios for significance testing and differences between cell means for interpretation of the direction of the effect, but no standardized index of effect size. When the dependent measure is in commonly understood units, such as yield of cotton per acre in agricultural research or dollars of income in economic research, the difference in means provides an informative measure. In the social sciences mean differences may also be informative, providing that some method of establishing meaningful measurement units has been accomplished. However, such unit establishment is often not the case, a problem discussed further in Section 5.2. In such a case standardized measures of effect size provided by the MRC analysis often permit more straightforward interpretation. Indeed, researchers in the ANOVA/ANCOVA tradition have become aware of standardized measures of effect size because of the rise of meta-analytic approaches that provide quantitative summaries of entire research literatures (e.g., Rosenthal, 1991). Some journal editors have also begun to encourage or even require inclusion of standardized effect size measures in articles published in their journals.

In addition to effect size measures in original (raw) and standardized units, the MRC system routinely provides several measures of the proportion of variance accounted for (the squares of simple, multiple, partial, and semipartial correlation coefficients). These measures of effect size are unit free and are easily understood and communicated. Each of the measures comes with its significance test value for the null hypothesis (F or t) so that no confusion between the two issues of *whether* and *how much* need arise.

1.3 MULTIPLE REGRESSION/CORRELATION AND THE COMPLEXITY OF BEHAVIORAL SCIENCE

The greatest virtue of the MRC system is its capacity to represent, with high fidelity, the types and the complexity of relationships that characterize the behavioral sciences. The word *complexity* is itself used here in a complex sense to cover several issues.

1.3.1 Multiplicity of Influences

The behavioral sciences inherited from older branches of empirical inquiry the simple experimental paradigm: Vary a single presumed causal factor (C) and observe its effects on the dependent variable (Y) while holding constant other potential factors. Thus, $Y = f(C)$; that is, to some degree, variation in Y is a function of controlled variation in C. This model has been, and continues to be, an effective tool of inquiry in the physical sciences, engineering, and in some areas of the behavioral sciences. A number of areas within the physical sciences and engineering typically deal with a few distinct causal factors, each measured in a clear-cut way, and each in principle independent of others.

However, as one moves to the broad spectrum of the basic and applied behavioral sciences ranging from physiological psychology to cultural anthropology to evaluation of educational programs, the number of potential causal factors increases, their representation in measures becomes increasingly uncertain, and weak theories abound and compete. Consider the following set of dependent variables from selected areas of the behavioral sciences, health sciences, education, and business: number of presidential vetoes (political science), extent of women's participation in the labor force (sociology), distance from home to work (geography), reaction time (experimental psychology), migration rate (demography), depression (clinical psychology), kinship system (anthropology), new business startups (economics), compliance with medical regime (health sciences), school achievement (education), and personnel turnover (business). A few moment's reflection about the context in which each of these is embedded suggests the multiplicity of both the potential causal factors and the forms of their relationships to the dependent variables. Given several research factors, C, D, E, etc., to be studied, one might use the single-factor paradigm repeatedly in a program of research: $Y = f(C)$, then $Y = f(D)$, then $Y = f(E)$, etc. But MRC permits the far more efficient simultaneous examination of the influences of multiple factors; that is, $Y = f(C, D, E, \text{etc.})$. Moreover, techniques such as structural equation analysis use interlocking regression equations to estimate formal models of causal processes derived from complex substantive theories.

1.3.2 Correlation Among Research Factors and Partialing

A far more important type of complexity than the sheer multiplicity of research factors lies in the effect of relationships among them. The simplest condition is that in which the factors $C, D, E, \ldots$ are statistically unrelated (orthogonal) to each other, as is the case in experiments in which the subject's level on each factor is under the experimenter's control and equal (or proportional) numbers of subjects are represented at each combination of factors. The overall importance of each factor in the experiment can be unambiguously determined because its independence of the other factors assures that its effects on Y cannot overlap with the effects of the others. Consider an experiment in which the apparent age (30 vs. 40) and sex (male, female) of a communicator are manipulated and their separate and joint effects on attitude change of male subjects is observed. The orthogonality of the factors is assured by having equal numbers of subjects in each of the four cells defined by the possible combinations of gender and age of the communicator (30-year-old male, 30-year-old female, 40-year-old male, 40-year-old female). No part of the difference in overall Y means for the two communicator ages can be attributed to their gender, nor can any part of the difference in the overall Y means for the two sexes be attributed to their ages.

Complexity arises when one departs from equal or proportional numbers of subjects in different conditions, because the independent variables are no longer independent. If in an experiment, the majority of the 40-year-olds were male and the majority of the 30-year-olds were female, then any difference between male and female communicators in the overall Y means would be confounded with (correlated with) communicator age. The age and sex effects would no longer be additive. Many issues in the behavioral sciences are simply inaccessible to true experiments and can only be addressed by the systematic observation of phenomena as they occur in their natural context. In nature, factors that influence Y are generally correlated with one another. Thus, if attitudes toward abortion (Y) are studied in a sample of survey respondents as a function of political party (C), religious background (D), and socioeconomic status (E), it is likely that C, D, and E will be correlated with each other. Relationships with Y, taken singly, will not accurately represent their separate influences, because of correlations among the factors (see Section 3.4). This is the familiar phenomenon of redundancy among correlated independent variables with regard to what they explain. The Y relationship with each of the independent variables overlaps to some degree with their relationships with other variables in the statistical model. This, in turn, requires a concept of the unique ("partialed") relationship of each variable with Y, in the context of the other variables in the model. This picture is often sharply different from that provided by looking at each factor singly. For example, it might be argued that the apparent influence of political party on attitudes toward abortion is entirely attributable to the relationship of party affiliation to religious preference or socioeconomic status. Such a pattern of results suggests that the apparent influence of political party on attitudes when appraised by itself may be "spurious"; that is, within subgroups that are homogeneous with regard to religious background and socioeconomic status, there is no difference on the average between members of one party and members of the other. Detailed attention to the relationships *among* potentially causal independent variables and how these bear on Y is the hallmark of causal analysis, and may be accomplished by MRC.

MRC's capability for assessing unique or partial relationships is perhaps its most important feature. Even a small number of research factors define many alternative causal systems. Some of these causal systems will be implausible because of considerations of prior research findings, logic, or research design (e.g., in a longitudinal design variables that occur later in time may be ruled out as potential causes of earlier variables). However, selection among the remaining causal systems is greatly facilitated by the ability, using MRC, of assessing the unique effect of a research factor, statistically controlling for (partialing) the effects of any desired set of other factors. Correlation does not prove causation; however, the absence of correlation implies the *absence* of the existence of a causal relationship. Thus, the skillful use of MRC can invalidate causal alternatives, assist researchers in choosing between competing theories, and help disentangle multiple influences through its partialing feature.

1.3.3 Form of Information

Variables employed in MRC may represent several different levels of measurement, of which it is often useful to distinguish the following (S. S. Stevens, 1951, 1958):

1. ***Ratio scales.*** These are equal interval scales with a true zero point, a point at which there is none of whatever the scale is measuring. Only such scales make statements such as "John weighs twice as much as Jim" or "Mary earns two-thirds as much as Jane" sensible.

Some examples of ratio scale measures include inches, pounds, seconds, size of group, dollars, distance from hospital, years in prison, and literacy rate.

2. *Interval scales.* These scales have equal intervals but are measured from an arbitrary point. For example, the Fahrenheit temperature scale uses the temperature at which a certain concentration of salt water freezes to represent 0. Values on the scale of less than 0 can and do occur. Many psychological and sociological indices are at this level, for example, scores on tests of intelligence, special abilities, achievement, personality, temperament, vocational interest, and social attitude. Such scales may not have a meaningful zero value at all.

3. *Ordinal scales.* Only the relative positions within a specific collection are signified by the values of ordinal scales. These scales do not have either equal intervals or a true zero point. Examples of ordinal scales include simple rankings of subjects in a sample as well as re-expressions of such rankings into percentiles, deciles, and quartiles.

4. *Nominal scales.* Nominal (categorical) scales involve simple classification of subjects into categories. The categories of nominal scales represent distinguishable qualities without a natural order or other quantitative properties. Examples include ethnic group, experimental treatment condition, place of birth, religion, marital status, psychiatric diagnosis, type of family structure, political party, public versus private sector, and gender. The set of categories are usually mutually exclusive and exhaustive. Thus, nominal scales are sets of groups that differ on some qualitative attribute.

This classification scheme is not exhaustive of quantitative scales, and others have been proposed. For example, psychological test scores are unlikely to measure with exactly equal intervals and it may be argued that they fall between interval and ordinal scales. Also, some rating scales frequently used in psychological research are not covered by Stevens' conception of levels of measurement. For example, scales like "0 = never, 1 = seldom, 2 = sometimes, 3 = often, and 4 = always" have a defined zero point, but intervals of dubious equality, although for most purposes they are treated as if they are approximately equal.

Basic MRC analysis can potentially consider information at any single level or any mixture of these levels of measurement. Ratio- and interval-level independent variables can be directly included in MRC models. Nominal variables can be expressed as coded variables (e.g., male = 0; female = 1), as will be discussed in Chapters 2, 8, and 9. Ordinal IVs may be treated as if they were interval variables in MRC models, and the results of the analyses may often be satisfactory. However, such an employment of these variables requires special caution, as is discussed further in Chapter 4. On the dependent variable side, Y may be measured at any of the levels of measurement, but the basic MRC model will usually work best if the data are interval or ratio. Some types of dependent variables may lead to violations of basic assumptions of the MRC model. In such cases, generalizations of the basic MRC model (the generalized linear model) can lead to improvements in the accuracy of the results over the basic MRC model (discussed in Chapter 13). This capacity of MRC and its generalizations to use information in almost any form, and to mix forms as necessary, is an important part of its adaptive flexibility.

1.3.4 Shape of Relationship

Consider the relationship $Y = f(C)$, where Y is a measure of poor health such as number of days of illness per year. For some factors the relationship may be well described by a straight line on the usual graph, for example, if C is daily cigarette consumption. Or, adequate description may require that the line be curved; for example, if C is age in years, the very young and the elderly are more often sick than young and middle-aged adults. Or, the shape may not be definable, as when C is a nominal variable like sex, ethnic background, or religion. When multiple research factors are being studied simultaneously, each may relate to Y (and each other) in any of these ways. Thus, when we write $Y = f(C, D, E, \ldots)$, f (as a function of) potentially covers a variety of complex functions that are readily brought under the sway of MRC.

How so? Many readers will know that MRC is often (and properly) referred to as *linear* MRC and may well be under the impression that correlation and regression are restricted to the study of straight-line relationships. This mistaken impression is abetted by the common usage of *linear* to mean "rectilinear" (straight line) and *nonlinear* to mean "curvilinear" (curved line). What is meant by *linear* in the MRC framework is any relationship of the form

$$Y = a + bU + cV + dW + eX + \cdots \tag{1.1.1}$$

where the lowercase letters are constants (either positive or negative) and the capital letters are variables. Y is said to be "linear in the variables U, V, etc." because it may be estimated by taking certain amounts (b, c, etc.) of each variable, and the constant a, and simply adding them together. In the fixed regression model framework in which we operate, there is no constraint on the nature of the IVs.[3] To illustrate this, consider substituting other variables for specific variables in the equation. For example, we could replace U and V in Eq. (1.1.1) with U and V^2, resulting in $Y = a + bU + cV^2$. Or, we could replace W with the logarithm of Z, resulting in $Y = a + d\log(Z)$. Or, we could replace X with a code variable representing sex (S, which takes values 0 = male and 1 = female), $Y = a + eS$. As our substitutions illustrate, the variables may be chosen to define relationships of *any* shape, rectilinear or curvilinear, or of no shape at all for unordered nominal independent variables, as well as all the complex combinations of these which multiple factors can produce.

Multiple regression equations are, indeed, linear; they are exactly of the form of Eq. (1.1.1). Yet they can be used to describe such a complex relationship as the length of psychiatric hospital stay as a function of ratings of patient symptoms on admission, diagnosis, age, sex, and average length of prior hospitalizations. This complex relationship is patently not rectilinear (straight line), yet it is readily described by a linear multiple regression equation.

To be sure, most relationships studied in the behavioral sciences are not of this order of complexity. But, the critical point is the capacity of MRC to represent any degree or type of shape—complexity is yet another of the important features which make it truly a *general* data-analytic system.

1.3.5 General and Conditional Relationships

Some relationships between Y and some factor C remain the same in regard to both degree and form despite variation in other factors D, E, F. In the MRC context, we will call such relationships *general* or *unconditional*: Readers familiar with ANOVA will know them as *main effects*. For example, suppose Y is a measure of visual acuity and C is age. In our example, both the form and degree of the relationship between visual acuity and age may remain the same under varying conditions of education level (D), ethnic group (E), and sex (F). The relationship between Y and C can then be said to be general insofar as the other specific factors are concerned. Note that this generality holds regardless of the form and degree of relationship between Y (visual acuity) and D, E, and F, between C (age) and D, E, and F, or among

[3] As we will note in Section 3.3, the "fixed" model we use throughout much of this book implies that we have generated or preselected the values of the IVs to which we wish to generalize.

D, E, and F. The Y–C relationship can thus be considered unconditional with regard to, or independent of, D, E, and F.

Now consider the same research factors, but with Y as a measure of attitudes toward abortion. The form and/or degree of relationship of age to Y is now almost certain to vary as a function of one or more of the other factors: it may be stronger or shaped differently at lower educational levels than higher (D), and/or in one ethnic group or another (E), and/or for men compared to women (F). The relationship of Y to C is now said to be conditional on D and/or E and/or F. In ANOVA contexts, such relationships are called *interactions*. For example, if the C–Y relationship is not constant over different values of D, there is said to be a $C \times D$ (age by educational level) interaction. Greater complexity is also possible: The C–Y relationship may be constant over levels of D taken by themselves, and over levels of E taken by themselves, yet may be conditional on the *combination* of D and E levels. Such a circumstance would define a "three-way" interaction, represented as $C \times D \times E$. Interactions of even higher order, and thus even more complex forms of conditionality, are theoretically possible, although rarely reliably found because of the very large sample size typically required to detect them.

Some behavioral science disciplines have found it useful to discriminate two types of conditional relationships.[4] *Moderation* indicates that the strength of the relationship between C and Y is *reduced* as the value of D increases. For example, researchers interested in the relationship between stress and illness report that social support moderates (weakens or buffers) this relationship. In contrast, *augmentation* or synergy means that the strength of the relationship between C and Y is *increased* as the value of D increases. Thus, moderation and augmentation describe particular forms of conditional relationships.

One facet of the complexity of the behavioral sciences is the frequency with which such conditional relationships are encountered. Relationships among variables often change with changes in experimental conditions (treatments, instructions, even experimental assistants), age, sex, social class, ethnicity, diagnosis, religion, personality traits, geographic area, etc. As essential as is the scientific task of estimating relationships between independent and dependent variables, it is also necessary to identify the conditions under which these estimates hold or change.

In summary, the generality of the MRC system of data analysis appropriately complements the complexity of the behavioral sciences, which complexity includes multiplicity and correlation among potential causal influences, a variety of forms in which information is couched, and variations in the shape and conditionality of relationships. Multiple regression/correlation also provides a full yield of measures of effect size with which to quantify various aspects of the strength of relationships (proportions of variance and correlation and regression coefficients). Finally, these measures are subject to statistical hypothesis testing, estimation, construction of confidence intervals, and power-analytic procedures.

1.4 ORIENTATION OF THE BOOK

This book was written to serve as a textbook and manual in the application of the MRC system for data analysis by students and practitioners in diverse areas of inquiry in the behavioral sciences, health sciences, education, and business. As its authors, we had to make many

---

[4]Elsewhere moderation may be used to describe both forms of conditional relationship. Whether a relationship may be considered to be moderated or augmented in the sense used here is entirely dependent on the (often arbitrary) direction of scoring of the IVs involved.

decisions about its level, breadth, emphasis, tone, and style of exposition. Readers may find it useful, at the outset, to have our orientation and the basis for these decisions set forth.

1.4.1 Nonmathematical

Our presentation of MRC is generally as conceptually oriented and nonmathematical as we could make it. Of course, MRC is itself a product of mathematical statistics, based on matrix algebra, calculus, and probability theory. There is little question that such a background makes possible a level of insight otherwise difficult to achieve. However, it is also our experience that some mathematically sophisticated scientists may lack the conceptual frame that links the mathematical procedures to the substantive scientific task in a particular case. When new mathematical procedures are introduced, we attempt to convey an intuitive conceptual rather than a rigorous mathematical understanding of the procedure. We have included a glossary at the end of the book in which the technical terms employed repeatedly in the book are given a brief conceptual definition. We hope that this aid will enable readers who have forgotten the meaning of a term introduced earlier to refresh their memories. Of course, most of these same terms also appear in the index with notation on the many times they may have been used. A separate table at the end of the book reviews the abbreviations used for the statistical terms in the book.

We thus abjure mathematical proofs, as well as unnecessary offhand references to mathematical concepts and methods not likely to be understood by the bulk of our audience. In their place, we heavily emphasize detailed and deliberately redundant verbal exposition of concrete examples. Our experience in teaching and consulting convinces us that our audience is richly endowed in the verbal, logical, intuitive kind of intelligence that makes it possible to understand how the MRC system works, and thus use it effectively (Dorothy Parker said, "Flattery will get you anywhere.") This kind of understanding is eminently satisfactory (as well as satisfying), because it makes possible effective use of the system. We note that to drive a car, one does not need to be a physicist, nor an automotive engineer, nor even a highly skilled auto mechanic, although some of the latter's skills are useful when one is stuck on the highway. That is the level we aim for.

We seek to make up for the absence of mathematical proofs by providing demonstrations instead. For example, the regression coefficient for a dichotomous or binary (e.g., male–female) independent variable that is scored 0–1 equals the difference between the two groups' Y means. Instead of offering the six or seven lines of algebra that would constitute a mathematical proof, we demonstrate that it holds using a small set of data. True, this proves nothing, because the result may be accidental, but curious readers can check it out using their own or our data (and we urge that such checks be made throughout). Whether it is checked or not, we believe that most of our audience will profit more from the demonstration than the proof. If the absence of formal proof bothers some readers from Missouri (the "show me" state), all we can do is pledge our good faith.

1.4.2 Applied

The first word in this book's title is *applied*. Our heavy stress on illustrations serves not only the function of clarifying and demonstrating the abstract principles being taught, but also that of exemplifying the kinds of applications possible. We attend to statistical theory only insofar as sound application makes it necessary. The emphasis is on "how to do it." This opens us to the charge of writing a "cookbook," a charge we deny because we do not neglect the whys and wherefores. If the charge is nevertheless pressed, we can only add the observation that in the kitchen, cookbooks are likely to be more useful than textbooks in organic chemistry.

1.4.3 Data-Analytic

Mathematical statisticians proceed from exactly specified premises such as independent random sampling, normality of distributions, and homogeneity of variance. Through the exercise of ingenuity and appropriate mathematical theory, they arrive at exact and necessary consequences (e.g., the F distribution, statistical power functions). They are, of course, fully aware that no set of real data will exactly conform to the formal premises from which they start, but this is not properly their responsibility. As all mathematicians do, they work with abstractions to produce formal models whose "truth" lies in their self-consistency. Borrowing their language, we might say that inequalities are symmetrical: Just as behavioral scientists are not mathematicians, mathematicians are not behavioral scientists.

The behavioral scientist relies very heavily on the fruits of the labors of theoretical statisticians. Taken together with contributions from substantive theory and previous empirical research, statistical models provide guides for teasing out meaning from data, setting limits on inference, and imposing discipline on speculation (Abelson, 1995). Unfortunately, in the textbooks addressed to behavioral scientists, statistical methods have often been presented more as harsh straightjackets or Procrustean beds than as benign reference frameworks. Typically, a method is presented with some emphasis on its formal assumptions. Readers are advised that the failure of a set of data to meet these assumptions renders the method invalid. Alternative analytic strategies may not be offered. Presumably, the offending data are to be thrown away.

Now, this is, of course, a perfectly ridiculous idea from the point of view of working scientists. Their task is to contrive situations that yield information about substantive scientific issues—*they must and will analyze their data*. In doing so, they will bring to bear, in addition to the tools of statistical analysis and graphical display of the data, their knowledge of theory, past experience with similar data, hunches, and good sense, both common and uncommon (Krantz, 1999). They attempt to apply the statistical model that best matches their data; however, they would rather risk analyzing their data using a less than perfect model than not at all. For them, data analysis is not an end in itself, but the next-to-last step in a scientific process that culminates in providing information about the phenomenon. This is by no means to say that they need not be painstaking in their efforts to generate and perform analyses of the data. They need to develop statistical models to test their preferred scientific hypothesis, to rule out as many competing explanations for the results as they can, and to detect new relationships that may be present in the data. But, at the end they must translate these efforts into substantive information.

Most happily, the distinction between data analysis and statistical analysis has been made and given both rationale and respectability by one of our foremost mathematical statisticians, John Tukey. In his seminal *The Future of Data Analysis* (1962), Tukey describes data analysis as the special province of scientists with substantial interest in methodology. Data analysts employ statistical analysis as the most important tool in their craft, but they employ it together with other tools, and in a spirit quite different from that which has come to be associated with it from its origins in mathematical statistics. Data analysis accepts "inadequate" data, and is thus prepared to settle for "indications" rather than conclusions. It risks a greater frequency of errors in the interest of a greater frequency of occasions when the right answer is "suggested." It compensates for cutting some statistical corners by using scientific as well as mathematical judgment, and by relying upon self-consistency and repetition of results. Data analysis operates like a detective searching for clues that implicate or exonerate likely suspects (plausible hypotheses) rather than seeking to prove out a balance. In describing data analysis, Tukey has provided insight and rationale into the way good scientists have always related to data.

The spirit of this book is strongly data-analytic, in exactly this sense. We offer a variety of statistical models and graphical tools that are appropriate for common research questions in the behavioral sciences. We offer straightforward methods of examining whether the assumptions of the basic fixed-model MRC are met, and provide introductions to alternative analytic approaches that may be more appropriate when they are not. At the same time, we are aware that some data sets will fail to satisfy the assumptions of any standard statistical model, and that even when identified there may be little that the data analyst can do to bring the data "into line." We recognize the limits on inference in such cases but are disposed to treat the limits as broad rather than narrow. We justify this by mustering whatever technical evidence there is in the statistical literature (especially evidence of the "robustness" of statistical tests), and by drawing upon our own and others' practical experience, even upon our intuition, all in the interest of getting on with the task of making data yield their meaning. If we risk error, we are more than compensated by having a system of data analysis that is general, sensitive, and fully capable of reflecting the complexity of the behavioral sciences and thus of meeting the needs of behavioral scientists. And we will reiterate the injunction that no conclusions from a given set of data can be considered definitive: Replication is essential to scientific progress.

1.4.4 Inference Orientation and Specification Error

As noted earlier, perhaps the single most important reason for the broad adoption of MRC as a data-analytic tool is the possibility that it provides for taking into account—"controlling statistically or partialing"—variables that may get in the way of inferences about the influence of other variables on our dependent variable Y. These operations allow us to do statistically what we often cannot do in real life—separate the influences of variables that often, or even usually, occur together. This is often critically important in circumstances in which it is impossible or unethical to actually control one or more of these related variables. However, the centrality of this operation makes it critically important that users of these techniques have a basic, sound understanding of what partialing influences does and does not entail.

In emphasizing the extraction of meaning from data we will typically focus primarily on potential problems of "specification error" in the estimates produced in our analyses. Specification errors are errors of inference that we make because of the way we analyze our data. They include the assumption that the relationship between the dependent variable Y and each of the independent variables (IVs) is linear (constant over the range of the independent variables) when it is not, and that the relationships of some IVs to Y do not vary as a function of other IVs, when they do. When we attempt to make causal inferences on the basis of the relationships expressed in our MRC analyses, we may also make other kinds of specification errors, including assuming that Y is dependent on the IVs when some of the IVs are dependent on Y, or that the relationship between Y and certain IVs is causal when these relationships reflect the influence of common causes or confounders. Or assuming that the estimated relationship reflects the relationship between Y and the theoretically implicated ("true") IV when it only reflects the relationship between Y and an imperfectly measured representative of the theoretically implicated IV. More technically, specification errors may include the conclusion that some relationship we seem to have uncovered in our sample data generalizes to the population, when our statistical analyses are biased by distributional or nonindependence problems in the data.

1.5 COMPUTATION, THE COMPUTER, AND NUMERICAL RESULTS

1.5.1 Computation

Like all mathematical procedures, MRC makes computational demands. The amount of computation increases with the size of the problem. Indeed, Darlington and Boyce (1982) estimate that computation time increases roughly with k^5, where k is the number of IVs. Early in the book, in our exposition of bivariate correlation and regression and MRC with two independent variables, we give the necessary details with small worked examples for calculation by hand calculator. This is done because the intimate association with the arithmetic details makes plain to the reader the nature of the process: *exactly* what is being done, with what purpose, and with what result. With one to three independent variables, where the computation is easy, not only can one see the fundamentals, but a basis is laid down for generalization to many variables.

With most real problems, MRC requires the use of a computer. An important reason for the rapid increase in the use of MRC during the past three decades is the computer revolution. Widely available computers conduct analyses in milliseconds that would have taken months or even years in Fisher's time. Statistical software has become increasingly user friendly, with versions that allow either simple programming or "point and click" analysis. Graphical routines that permit insightful displays of the data and the results of statistical analyses have become increasingly available. These advances have had the beneficial effect of making the use of MRC analysis far faster and easier than in the past.

We have deliberately placed the extensive calculational details of the early chapters outside the body of the text to keep them from distracting attention from our central emphasis: understanding how the MRC system works. We strongly encourage readers to work through the details of the many worked illustrations using both a hand calculator and a statistical package. These can help provide a basic understanding of the MRC system and the statistical package.

But readers should then apply the methods of each chapter to *data of their own* or data with which they are otherwise familiar. The highest order of understanding is achieved from the powerful synergism of the application of unfamiliar methods to familiar data.

Finally, we caution readers about an unintended by-product of the ease of use of current statistical packages: Users can now easily produce misleading results. Some simple commonsense checks can often help avoid errors. Careful initial examination of simple statistics (means; correlations; number of cases) and graphical displays can often provide a good sense of the data, providing a baseline against which the results of more complicated analyses can be compared. We encourage readers using new software to try out the analysis first on a previously analyzed data set, and we include such data sets for the worked examples in the book, for which analyses have been carried out on the large SAS, SPSS, and SYSTAT statistical programs. Achieving a basic understanding of the MRC system and the statistical packages as well as careful checking of one's results is an important prerequisite to publication. There is no guarantee that the peer review process in journals will detect incorrect analyses.

1.5.2 Numerical Results: Reporting and Rounding

Statistical packages print out numerical results to several decimal places. For comparison purposes, we follow the general practice in this book of reporting computed correlation and regression coefficients rounded to two places (or significant digits) and squared coefficients rounded to three. When working with a hand calculator, the reader should be aware that small rounding errors will occur. Checks that agree within a few points in the third decimal may thus be taken as correct.

Following Ehrenberg (1977), we encourage readers to be conservative in the number of significant digits that are reported in their research articles. Despite the many digits of accuracy that characterize modern statistical programs, this level of accuracy only applies to the sample data. Estimates of population parameters are far less accurate because of sampling error. For the sample correlation (r) to provide an estimate of the population correlation (ρ) that is accurate to *two* decimal places would require as many as 34,000 cases (J. Cohen, 1990).

1.5.3 Significance Tests, Confidence Intervals, and Appendix Tables

Most behavioral scientists employ a hybrid of classical Fisherian and Neyman-Pearson null hypothesis testing (see Gigerenzer, 1993; Harlow, Mulaik, & Steiger, 1997), in which the probability of the sample result given that the null hypothesis is true, p, is compared to a prespecified significance criterion, α. If $p <$ (is less than) α, the null hypothesis is rejected and the sample result is deemed statistically significant at the α level of significance. The null hypothesis as typically specified is that the value of the parameter corresponding to the sample result is 0; other values can be specified based on prior research.

A more informative way of testing hypotheses in many applications is through the use of confidence intervals. Here an interval is developed around the sample result that would theoretically include the population value $(1 - \alpha)\%$ of the time in repeated samples. Used in conjunction with MRC procedures, the center of the confidence interval provides an estimate of the strength of the relationship and the width of the confidence interval provides information about the accuracy of that estimate. The lower and upper limits of the confidence interval show explicitly just how small and how large the effect size in the population (be it a regression coefficient, multiple R^2,or partial r) might be. Incidentally, if the population value specified by the null hypothesis is not contained in the confidence interval, the null hypothesis is rejected.

The probability of the sample result given that the null hypothesis is true, p, is based on either the t or F distribution in basic MRC. Nearly all statistical packages now routinely compute exact values of p for each significance test. We also provide tables of F and t for $\alpha = .05$ and $\alpha = .01$. These values are useful for the construction of confidence intervals and for simple problems which can be solved with a hand calculator. The $\alpha = .05$ criterion is widely used as a standard in the behavioral sciences. The $\alpha = .01$ criterion is sometimes used by researchers as a matter of taste or tradition in their research area. We support this tradition when there are large costs of falsely rejecting the null hypothesis; however, all too frequently researchers adopt the $\alpha = .01$ level because they erroneously believe that this decision will necessarily make their findings stronger and more meaningful. The $\alpha = .01$ level is often used as a partial control on the incidence of spuriously significant results when a large number of hypothesis tests are being conducted. The choice of α also depends importantly on considerations of statistical power (the probability of rejecting the null hypothesis), which is discussed in several places, particularly in Section 4.5. We present tables for statistical power analysis in the Appendix; several programs are commercially available for conducting statistical power analyses on personal computers (e.g., Borenstein, Cohen, & Rothstein, 2001).

The statistical tables in the Appendix were largely abridged from Owen (1962) and from J. Cohen (1988). The entry values were selected so as to be optimally useful over a wide range of MRC applications. In rare cases in which the needed values are not provided, linear interpolation is sufficiently accurate for almost all purposes. Should more extensive tables be required, Owen (1962) and Pearson and Hartley (1970) are recommended. Some statistical packages will also compute exact p values for any specified df for common statistical distributions such as t, F, and χ^2.

1.6 THE SPECTRUM OF BEHAVIORAL SCIENCE

When we address behavioral scientists, we are faced with an exceedingly heterogeneous audience. They range in level from student to experienced investigator and possess from modest to fairly advanced knowledge of statistical methods. With this in mind, we assume a minimum background for the basic exposition of the MRC system. When we must make assumptions about background that may not hold for some of our readers, we try hard to keep everyone on board. In some cases we use boxes in the text to present more technical information, which provides a greater understanding of the material. The boxes can be skipped on first reading without loss of continuity.

But it is with regard to substantive interests and investigative methods and materials that our audience is of truly mind boggling diversity. Behavioral science itself covers areas of "social", "human", and even "life" sciences—everything from the physiology of behavior to cultural anthropology, in both their "basic science" and "applied science" aspects. Add in health sciences, education, and business, and the substantive range becomes immense. Were it not for the fact that the methodology of science is inherently more general than its substance, a book of this kind would not be possible. This permits us to address substantive researchers whose primary interests lie in a bewildering variety of fields.

We have sought to accommodate to this diversity, even to capitalize upon it. Our illustrative examples are drawn from different areas, assuring the comfort of familiarity for most of our readers at least some of the time. Their content is presented at a level that makes them intellectually accessible to nonspecialists. We try to use the nontechnical discussion of the examples in a way that may promote some methodological cross-fertilization between fields of inquiry. Our hope is that this discussion may introduce better approaches to fields where data have been analyzed using traditional rather than more optimal procedures.

1.7 PLAN FOR THE BOOK

1.7.1 Content

Following this introductory chapter, we continue by introducing the origins and meanings of the coefficients that represent the relationship between two variables (Chapter 2). Chapter 3 extends these concepts and measures first to two independent variables and then to any larger number of independent variables. Chapter 4 expands on the graphical depiction of data, and particularly on the identification of data problems, and methods designed to improve the fit of the data to the assumptions of the statistical model. Chapter 5 describes the strategies that a researcher may use in applying MRC analyses to complex substantive questions, including selecting the appropriate statistical coefficients and significance tests. It continues by describing two widely useful techniques, hierarchical (sequential) analyses of data and the analysis of independent variables grouped into structural or functional sets.

Chapters 6 and 7 describe and illustrate the methods of identifying nonlinear and conditional relationships between independent variables and Y, beginning with methods for representing curvilinearity in linear equations. This chapter is followed by detailed presentations of the treatment and graphic display of interactions between scaled variables in their relationship with Y. Chapter 8 continues with the consideration of sets of independent variables representing mutually exclusive categories or groups. Relationships between scaled measures and Y may vary between sample subgroups; techniques for assessing and describing these interactions are reviewed in Chapter 9.

Chapter 10 presents the problem of multicollinearity among predictors and methods of controlling its extent. Chapter 11 details the full range of methods for coping with missing data in MRC, and the considerations appropriate for choosing among them.

Chapter 12 expands on the discussion of MRC applications to causal hypotheses that is found in earlier chapters and introduces the reader to some of the more complex methods of estimating such models and issues relevant to their employment.

Chapter 13 describes uses of the generalized linear model to analyze dependent variables that are dichotomous, ordered categories, or counts of rare phenomena. Chapter 14 introduces the reader to the multilevel analysis of data clusters arising from nonindependent sampling or treatment of participants.

Chapter 15 provides an introduction to a whole range of methods of analyzing data characterized by multiple observations of units over time. Beginning with simple repeated measure ANOVA and two time-point MRC, the chapter presents an overview of how the substantive questions and the structure of the data combine to suggest a choice among available sophisticated data analytic procedures.

The final chapter presents a multivariate method called set correlation that generalizes MRC to include sets (or partialed sets) of dependent variables and in so doing, generalizes multivariate methods and yields novel data-analytic forms.

For a more detailed synopsis of the book's contents, the reader is referred to the summaries at the ends of the chapters. The data for almost all examples in the book are also provided on the accompanying CD-ROM, along with the command codes for each of the major statistical packages that will yield the tabular and other findings presented in the chapters.

A note on notation. We have tried to keep the notation simultaneously consistent with the previous editions of this book and with accepted practice, insofar as possible. In general, we employ Greek letters for population estimates, but this convention falls down in two places. First, β is used conventionally both for the standardized regression coefficient and for the power: We have followed these conventions. Second, the maximum likelihood estimations methods discussed in Chapters 13 and 14 use a range of symbols, including Greek letters, designed to be distinct from those in use in OLS. We also use P and Q $(= P - 1.0)$ to indicate proportions of samples, to distinguish this symbol from $p =$ probability.

We have attempted to help the reader keep the major concepts in mind in two ways. We have included a glossary of technical terms at the end of the book, so that readers of later chapters may refresh their recall of terms introduced earlier in the book. We have also included a listing of the abbreviations of statistical terms, tests, and functions. In addition there are two technical appendices, as well as the appendix Tables.

One more difference between this edition and previous editions may be noted. In the introductory Chapter 2 we originally introduced equations using the sample standard deviation, with n in the denominator. This forced us into repeated explanations when later statistics required a shift to the sample-based population estimate with $n - 1$ in the denominator. The advantage was simplicity in the early equations. The serious disadvantage is that every statistical program determines sd with $n - 1$ in the denominator, and so students trying to check sds, z scores and other statistics against their computer output will be confused. In this edition we employ the population estimate sd consistently and adjust early equations as necessary.

1.7.2 Structure: Numbering of Sections, Tables, and Equations

Each chapter is divided into major sections, identified by the chapter and section numbers, for example, Section 5.4.3 ("Variance Proportions for Sets and the Ballantine Again") is the third subsection of Section 5.4. Further subdivisions are not numbered, but titled with an italicized heading.

Tables, figures, and equations within the body of the text are numbered consecutively within major sections. Thus, for example, Figure 5.4.1 is the first figure in Section 5.4, and Eq. (2.6.5) is the fifth equation in Section 2.6. We follow the usual convention of giving equation numbers in parentheses. A similar plan is followed in the two appendices. The reference statistical tables make up a separate appendix and are designated as Appendix Tables A through G.

On the accompanying data disk each chapter has a folder; within that folder each example for which we provide data and syntax/command files in SAS, SPSS, and SYSTAT has a folder.

1.8 SUMMARY

This introductory chapter begins with an overview of MRC as a data-analytic system, emphasizing its generality and superordinate relationship to the analysis of variance/covariance (Section 1.1). MRC is shown to be peculiarly appropriate for the behavioral sciences in its capacity to accommodate the various types of complexity that characterize them: the multiplicity and correlation among causal influences, the varieties of form of information and shape of relationship, and the frequent incidence of conditional (interactive) relationships. The special relevance of MRC to the formal analysis of causal models in described (Section 1.2).

The book's exposition of MRC is nonmathematical, and stresses informed application to scientific and technological problems in the behavioral sciences. Its orientation is "data analytic" rather than statistical analytic, an important distinction that is discussed. Concrete illustrative examples are heavily relied upon (Section 1.3).

The popularity of MRC in the analysis of nonexperimental data for which manipulation of variables is impossible or unethical hinges on the possibility of statistical control or partialing. The centrality of this procedure, and the various kinds of errors of inferences that can be made when the equations include specification error are discussed (Section 1.4).

The means of coping with the computational demands of MRC are briefly described and largely left to the computer, with details relegated to appendices so as not to distract the reader's attention from the conceptual issues (Section 1.5). We acknowledge the heterogeneity of background and substantive interests of our intended audience, and discuss how we try to accommodate to it and even exploit it to pedagogical advantage (Section 1.6).

The chapter ends with a brief outline of the book and the scheme by which sections, tables, figures, and equations are numbered.