

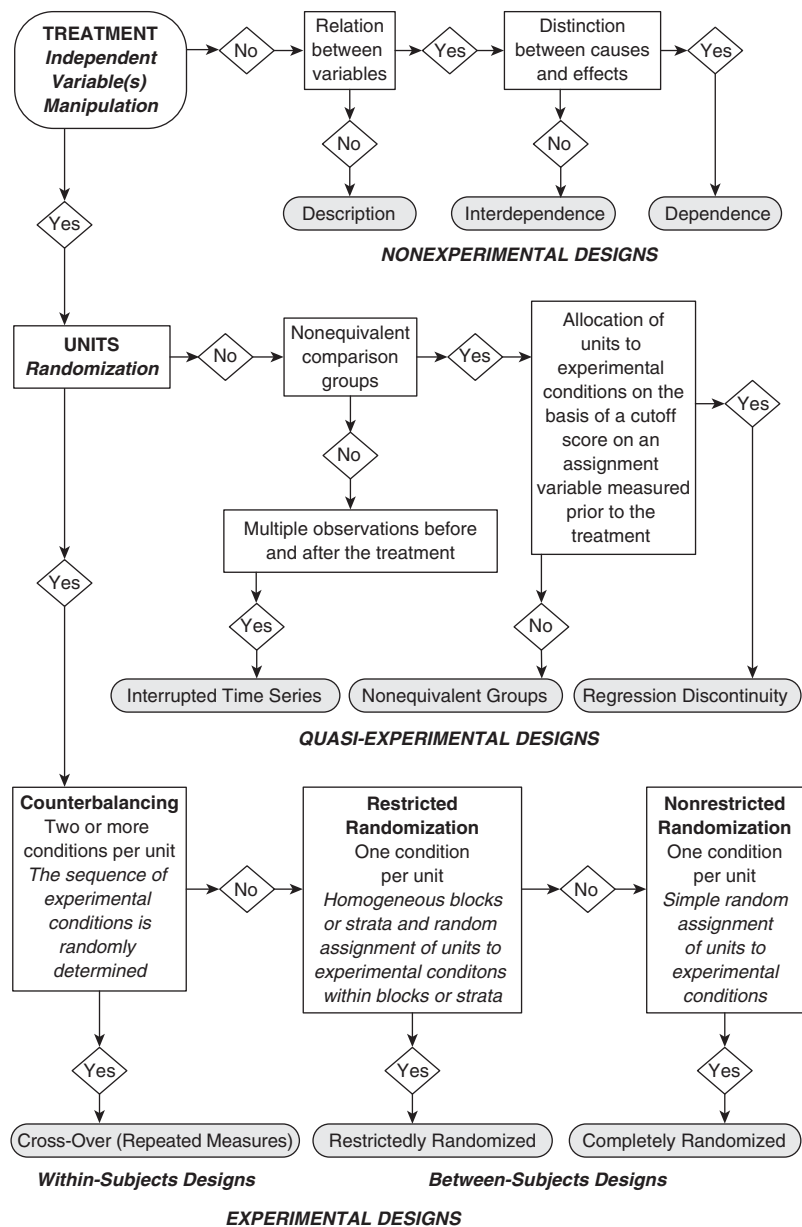
# CHAPTER 1. RANDOMIZED EXPERIMENTS

## 1.1 Nature and Structure of Randomized Experiments

In broad terms, methods are the linking procedures between theory and data. They embody the theoretical hypothesis in the research design, specifying the conditions and technical devices to collect, analyze, and interpret relevant basic information (raw data). In the nomothetic or quantitative approach, as opposed to the idiographic or qualitative one, methods of empirical investigation are usually classified on the basis of the structural properties of the underlying design and the degree to which they allow valid causal inferences.

Leaving out *theoretical innovation, research synthesis and evaluation* (including meta-analysis and bibliometric surveys), and *documental studies* focused on previously archived materials, all investigations based either on *self-reports* (via self-administrated questionnaires or interviews) or on *direct observation of research participants* fall into one of the three main categories: (1) *experimental*, (2) *quasi-experimental*, and (3) *nonexperimental* research (see Figure 1.1; see also Alferes, 2012, for a methodological classification of theoretical and empirical studies in psychology and related disciplines). As the reader can verify from the decision chart shown in Figure 1.1, this classification scheme—inspired by the Campbellian approach to the validity of causal inferences (Campbell, 1957; Shadish et al., 2002)—is organized around two key attributes of empirical studies: (1) *systematic variation of the presumed causes* (independent variables manipulation) and (2) *use of randomization procedures*. The presence of the first attribute separates *experimental* and *quasi-experimental* research from *nonexperimental* research; the presence of the second one separates *randomized experiments* (experimental designs) from *nonrandomized experiments* (quasi-experimental designs). The specific use of randomization procedures in experimental design depends on the manipulation strategy adopted by the researcher: (a) each unit is only exposed to *one* experimental condition and the randomization procedure is used to determine what exactly is that condition (*between-subjects designs*), or (b) each unit is exposed to *two or more* experimental conditions and the randomization procedure is used to determine the order in which the conditions will be presented (*within-subjects designs*).

As stated in the opening paragraph of the Preface, this book is about experimental designs and the devices required for an *unbiased and efficient* estimation of treatment causal effects. Formally speaking, a causal effect is



**Figure 1.1** Decision chart for the classification of experimental, quasi-experimental, and nonexperimental designs according to the Campbellian approach

the difference between what happens when an experimental unit is subjected to a treatment and what would happen if it was not subjected to the same treatment or, which is equivalent, the difference between the responses of an experimental unit when simultaneously subjected to two alternative treatments (differential causal effect). Stated in another way, the inference of a causal effect requires counterfactual evidence. Yet, regarding a concrete experimental unit, it is impossible to obtain such evidence: We are unable to apply and not apply a treatment (or apply two alternative treatments) to an experimental unit at the same time.

A tentative solution for this problem could be either the comparison of the responses of two units (one receiving and one not receiving the treatment or one receiving the treatment and the other the alternative treatment) or the comparison of the responses of the same unit observed in two successive periods. However, this time, the counterfactual evidence is equivocal: In the first case, the treatment effect is completely confounded with the intrinsic characteristics of the experimental unit; in the second case, the treatment effect is completely confounded with any systematic or random variation potentially associated with the different periods of observation. A better solution, and the only one really feasible, is to replicate the experiment with other experimental units. Provided that certain assumptions are verified, having observations from several units can allow us to separate the treatment effect from “subjects” and “temporal sequence” effects.

What we have been saying is the core content of Rubin’s causal model (Rubin, 1974, 2006, 2007; see also Rubin, 2004, for a pedagogical introduction), which defines a causal effect in terms of the mean difference in the potential outcomes between those who were submitted and those who were not submitted to a treatment, as long as the experimenter can guarantee what Rubin calls *stable-unit-treatment-value assumption* (SUTVA). Among other things, this assumption states that the potential outcome of one unit must not be affected by the actual assignment (to experimental conditions) of the remaining units. More precisely, SUTVA is a twofold assumption. First, it implies that there are *no hidden or different versions of the treatments*; that is, the selected treatment levels (or treatment level combinations) are administered without any modification from the beginning to the end of the experiment. Second, it implies *no interference or interaction between subjects* who are receiving different treatment levels (or treatment level combinations). If we are dealing with a randomized experiment, this means that the independence condition introduced by the initial randomization must be preserved during the experiment to avoid potential contamination effects resulting from interactions between subjects assigned to distinct experimental conditions. Some authors subsume randomization procedures under the

SUTVA rubric, despite the clear statement of Rubin (2007, 2010) that substantive assumptions must be distinguished from the underlying assignment mechanism.

Rubin's causal model can be seen as an elaboration of the three fundamental principles of experimentation introduced by Fisher (1935/1966): (1) *replication*, (2) *randomization*, and (3) *local control*. The first principle implies the recording of observations from several units to estimate causal effects (mean differences). Randomization guarantees that the estimate is a nonbiased one. Local control ensures more precision in the estimation (i.e., a more efficient nonbiased estimator; for a synthesis of the properties of statistical estimators, see Fox, 2009). Stated in another way, randomization rules out alternative explanations based on the intrinsic characteristics of experimental units, whereas local control reduces the magnitude of random noise (residual variability) in the experiment.

In addition to the elaboration of Fisher's principles, a distinctive feature of Rubin's causal model is the replacement of the *observed outcomes notation* with the *potential outcomes notation* underlying his contrafactual approach to experimentation (*randomized experiments*) and quasi-experimentation (*observational studies*, according to the terminology introduced by Cochran [1965, 1983] and popularized by Rosenbaum [2002, 2010] and Rubin himself [Cochran & Rubin, 1973; Rubin, 1973]). In the context of this introductory chapter, it is sufficient to remark that the potential outcomes notation, initially proposed by Neyman (1923/1990), constitutes a coherent framework for the analysis of randomized and nonrandomized experiments and is particularly relevant in cases where the conditions established by the initial randomization are broken throughout the experiment and the SUTVA is violated. We will revisit this issue in the final chapter, which is centered on practical matters and the guiding principles of data analysis.

For now, we will be returning to the basics of experimental design, reproducing an extended definition given by Kirk (1995), in which the nature and the structure of randomized experiments are clearly detailed:

The term experimental design refers to a plan for assigning subjects to experimental conditions and the statistical analysis associate with the plan. The design of an experiment to investigate a scientific or research hypothesis involves a number of interrelated activities:

- (1) Formulation of statistical hypotheses that are germane to the scientific hypothesis. A *statistical hypothesis* is a statement about (a) one or more parameters of a population or (b) the functional form of a population. Statistical hypotheses are rarely identical to scientific hypotheses; they are testable formulations of scientific hypotheses.

- (2) Determination of the experimental conditions (independent variable) to be used, the measurement (dependent variable) to be recorded, and the extraneous conditions (nuisance variables) that must be controlled.
- (3) Specification of the number of subjects (experimental units) required and the population from which they will be sampled.
- (4) Specification of the procedure for assigning the subjects to the experimental conditions.
- (5) Determination of the statistical analysis that will be performed. (pp. 1–2)

In the next section, these “interrelated activities” involved in experimental design are discussed in the broader context of the validity of causal inferences, and the main structural features of randomized experiments (*experimental factors*, *pseudofactors*, *classificatory factors*, and *outcome measures*) are conveniently described. Section 1.3 gives an overview of experimental designs in connection with methods of randomization and must be read as an advanced organizer for the core contents of Chapters 2 and 3. This introductory chapter ends with a brief section devoted to important terminological and notational issues.

## 1.2 Experimental Design and Validity of Scientific Inferences

Scientific hypotheses are conjectural statements about the relationships between theoretical constructs, empirically represented by particular events or realizations (called operationalizations or measurements). In the nomothetic tradition, relational or causal connections specified in the hypothesis are (ideally) theory driven. Inferences from data (particular observables) to the hypothesis (corroboration, falsification) are the realm of the scientific enterprise and must be distinguished from statistical inferences, which are about estimating population parameters from sampling particulars (Meehl, 1990). Study designs are the embodiment of the theoretical hypothesis, and their structural features define the conditions and constraints of scientific inference.

We can easily define the structural features of experimental designs by relying on the well-known Lewinian truism, which states that “in general terms, behavior ( $B$ ) is a function ( $F$ ) of the person ( $P$ ) and of his environment ( $E$ ),  $B = F(P, E)$ ” (Lewin, 1946/1997, p. 337). First, *behavioral measures* ( $B$ ) are *dependent variables* (outcome measures). Second, *environmental* or *situational variables* ( $E$ ) susceptible of being manipulated are *experimental factors* (independent variables, treatments, or interventions). Third, *personal* or *dispositional variables* ( $P$ ) are *classificatory factors*, which are

ideally controlled by random assignment of units to experimental conditions. Finally, *pseudofactors*—that is, environmental or situational variables other than the focal or primary experimental factors—can be incorporated in the design (as secondary experimental factors), locally controlled (by holding them constant), or statistically handled (by measurement and subsequent modeling as covariables). When *classificatory factors*—conceptualized either as randomization restrictions (blocks or strata) or as measured covariables (substantive dispositional moderators)—are included in the design, they are occasionally labeled *passive independent variables* and contrasted with *true* (i.e., manipulated) or *active independent variables*.

A substantive classification of classificatory factors, experimental factors, outcome measures, and pseudofactors is given in Table 1.1, where some disciplinary traditions in experimental research are also identified. Using the terminology proposed by Cronbach (1982) to describe the basic

**Table 1.1** Classification of Experimental Factors, Classificatory Factors, Pseudofactors, and Outcome Measures in Randomized Experiments

<b>I. Experimental Factors (Independent Variables)</b>	<b><i>Treatments</i></b>
A. Physical Manipulations	
<ul style="list-style-type: none"> <li>• Variations in physical settings (<i>ecological tradition</i>)</li> <li>• Variations of specific stimuli (<i>experimental psychology tradition</i>)</li> </ul>	
B. Biological Manipulations	
<ul style="list-style-type: none"> <li>• Biophysiological treatments and interventions (e.g., drugs, chemical therapies, surgical interventions; <i>pharmacological and evidence-based medicine traditions</i>)</li> <li>• Physical exercise and diet regimen (<i>sports and nutritional sciences traditions</i>)</li> </ul>	
C. Psychosocial Manipulations	
<ul style="list-style-type: none"> <li>• Variations in social stimuli and situations (<i>Festinger's tradition in social psychology</i>)</li> <li>• Variations in the <i>cognitive definition</i> (instructional manipulations) of social situations, physical settings and stimuli, or internal states (<i>experimental psychology tradition; social cognition experiments tradition; Schachter's tradition in social psychology</i>)</li> <li>• Variations of response contingencies (e.g., schedules of reinforcement; <i>Skinner's behaviorist tradition</i>)</li> <li>• Systematic psychological (e.g., psychotherapies) or social interventions (e.g., educational programs; <i>social experimentation tradition; evaluation studies tradition</i>)</li> </ul>	
D. Combinations of Physical, Biological, and Psychosocial Manipulations	

---

## II. Classificatory Factors (Personal or Dispositional Variables) *Units*

- A. Biosocial Markers (gender, age, nationality, ethnicity, socioeconomic status, educational level, political and religious affiliations, sexual orientation, family and relationships status, structural characteristics of social and professional networks, etc.)
- B. Physical Attributes and Organismic Variables
- C. Personality Traits and (Enduring) Motivational-Emotional Dispositions
- D. Cognitive Abilities and Styles
- E. Frames of Reference (e.g., ideologies, shared social representations, etc.), Values, and Social Attitudes

## III. Pseudofactors (Environmental or Situational Variables) *Settings*

- A. All the variables classified under *I—Experimental Factors*—but not being the focal target (i.e., active independent variables) in the current experiment
- B. Socio-Institutional and Ecological Contexts and Temporal Structure of Experiments

## IV. Outcome Measures (Dependent Variables)<sup>a</sup> *Observations*

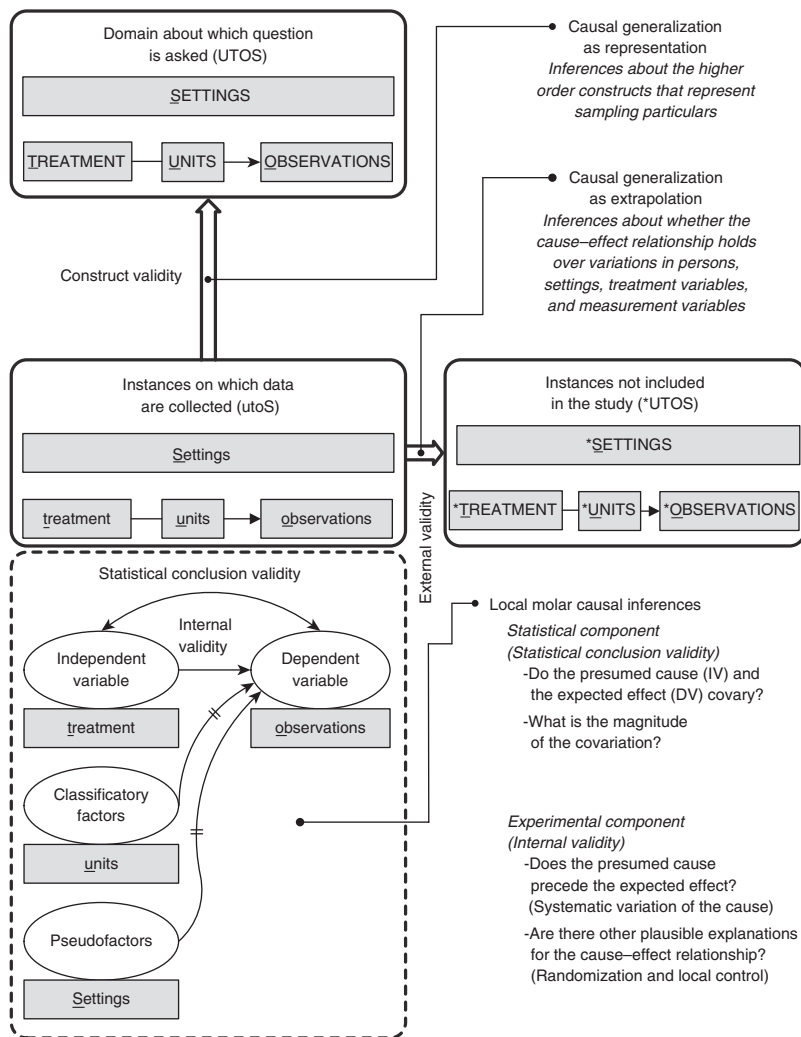
- A. Self-Report Measures (rating scales; questionnaires and interviews; etc.)
  - B. Observational Measures
    - Overt behaviors (including verbal behavior and expression of behavioral intentions, as well as performances in standardized tasks or psychological and educational tests)
    - Biophysiological measures
  - C. Accretion and Erosion Measures (“behavioral fossils”)
- 

*Note.* This classification is restricted to randomized experiments with human beings or animals as experimental units, omitting typical manipulations and measures in agricultural, physical, or technological research.

<sup>a</sup>Classification of self-report and observational measures is based on Aronson, Ellsworth, Carlsmith, and Gonzalez (1990). Accretion and erosion measures are extensively presented and discussed in Webb, Campbell, Schwartz, Sechrest, and Grove (1981).

elements of experimental designs (UTOS: *units, treatments, observations, and settings*), the relationships between these elements are depicted in the lower left panel of Figure 1.2.

Figure 1.2 is a graphical representation of the widely known Campbellian approach to the validity of scientific inferences (Campbell, 1957;



**Figure 1.2** Components of the Campbellian validity typology

Campbell & Stanley, 1966; Cook & Campbell, 1979; Shadish & Cook, 2009; Shadish et al., 2002). To state it briefly, scientific claims are evaluated by the degree to which the underlying design allows the researcher to *make* and to *generalize* local causal inferences. Making valid local causal inferences is synonymous with giving unequivocal evidence of covariance



between the presumed causes and the expected effects and, simultaneously, ruling out alternative explanations for the observed relationship. That is, the validity of local causal inferences depends on the statistical and experimental components of the research design (see *statistical conclusion validity* and *internal validity* in the lower right panel of Figure 1.2). Generalization of causal relationships requires a twofold approach: (1) generalizing from sampling particulars to higher order constructs (causal generalization as representation, or *construct validity*) and (2) generalizing the local causal relationship to other persons and instances not included in the study (causal generalization as extrapolation, or *external validity*) (see the two upper panels of Figure 1.2).

The four dimensions of the validity of scientific inferences (external, construct, internal, and statistical conclusion validity) can be thought of as the organizing principles for the main areas of the methodological field (*sampling, measurement, design, and analysis*), matching, term by term, the critical challenges that all researchers must deal with: (1) to extrapolate from observed samples to target populations, (2) to guarantee that their empirical realizations adequately represent the theoretical constructs involved in the scientific hypothesis, (3) to establish causality, and (4) to model the data ensuring that the underlying statistical relationships are not spurious. The first section of Chapter 4—focusing on practical matters related to planning and monitoring randomized experiments—is organized around the most common strategies to overcome potential drawbacks in local (Subsection 4.1.2) and generalized (Subsection 4.1.1) causal inferences.

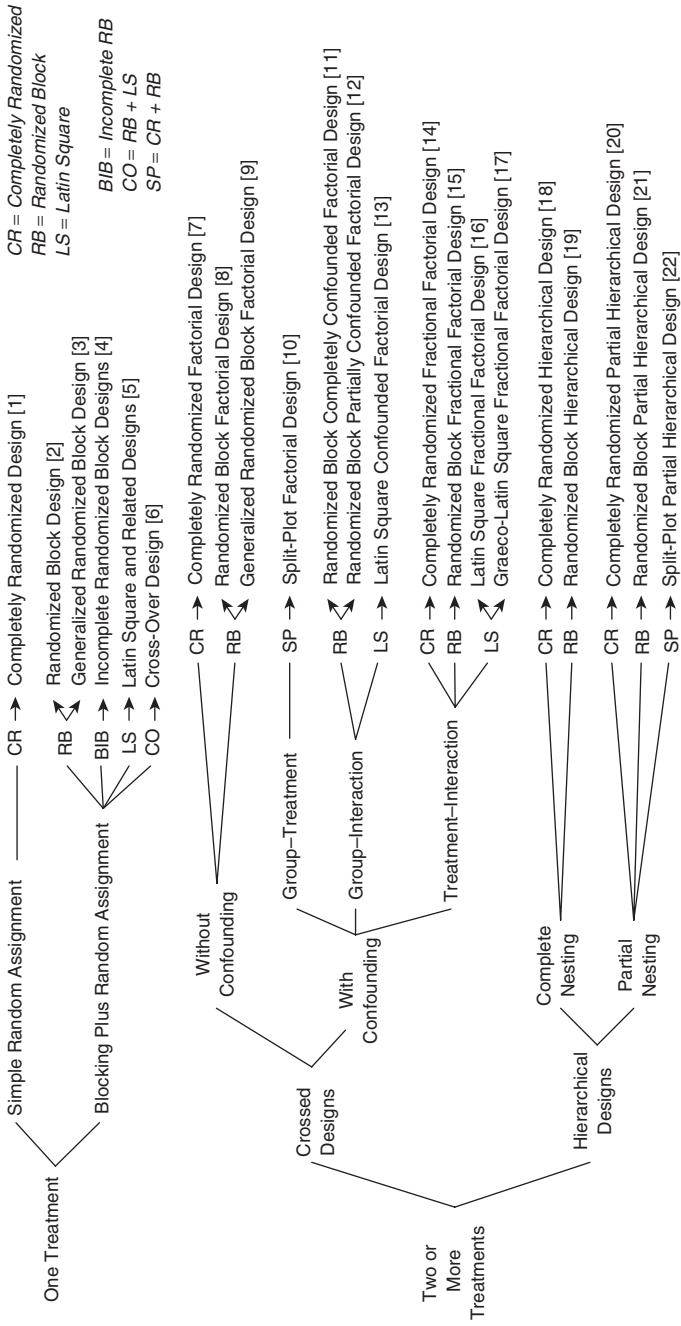
### 1.3 Randomized Experiments and Methods of Randomization

In Section 1.1, experimental designs are contrasted with nonexperimental and quasi-experimental designs and classified into two broad categories according to the researcher's manipulation strategy: (1) *between-subjects* versus (2) *within-subjects designs*. Additionally, between-subjects designs are split into *completely randomized designs* and *restrictedly randomized designs*, on the basis of the restrictions (blocking or stratifying) imposed on random assignment procedures (see Figure 1.1). This classification scheme can be elaborated to account for other features of experimental design, such as the number of treatments, the pattern of treatment level combinations, and the introduction of control devices other than randomization and blocking (e.g., single and multiple pretest–posttest observations, covariates, or concomitant variables).

Relying on the *elementary building blocks* of randomized experiments (*CR*—*completely randomized*, *RB*—*randomized block*, and *LS*—*Latin square* basic designs) and their possible combinations, Kirk (1995, 2003a) proposes a very useful classification of factorial extensions of one-treatment designs. Excluding the examples of four “systematic designs” and two miscellaneous designs, Kirk (2003a, p. 11) lists 34 randomized designs, classified on the basis of six criteria: (1) the *number of treatments* (one vs. two or more treatments), (2) the *absence/presence of covariates* (analysis of variance [ANOVA] vs. analysis of covariance [ANCOVA] designs), (3) the *randomization procedures* used (simple random assignment vs. blocking prior to random assignment), (4) the *structure of factorial designs* (crossed vs. hierarchical designs), (5) the *absence/presence of confounding in crossed designs*, and (6) the *type of hierarchical designs* (complete vs. partial nesting). Ignoring ANCOVA designs, whose structure is similar to the equivalent ANOVA designs, and grouping together the variants and extensions of *incomplete block* (Designs 4) and *Latin square* (Designs 5) designs, we get the 22 designs listed in Figure 1.3.

Adopting an analysis-centered perspective, Kirk (1995, 2003a) subsumes within-subjects or repeated measures designs under the rubric *cross-over design*, a special type of the randomized block design where each block is formed by a single experimental unit that is observed in all experimental conditions (see Chapter 3, Table 3.3, Design 6N). This classification of within-subjects designs is consistent with the underlying statistical model of Design 6N, which for computational purposes is precisely the same. However, as Kirk (1995) observes, compared with designs containing homogeneous but different units per block, the cross-over design has distinct interpretations and generalizes to different target populations. More important in the context of this book, the counterbalancing procedures applying to Design 6N and to other types of cross-over designs are quite different from the random assignment procedures used in the between-subjects designs, and therefore, we have chosen to handle within-subjects designs randomization in a separate chapter (Chapter 3). For the rest, Figure 1.3 can be taken as an outline for the description of random assignment and blocking procedures in between-subjects designs (Chapter 2). The distinct types of the one-factor cross-over design (Designs 6 in Figure 1.3) and their factorial extensions are listed and labeled in Table 3.3 (Chapter 3). The main randomization methods described and illustrated in Chapters 2 and 3 are named and sequentially numerated in Tables 1.2 and 1.3.

This arrangement of the core contents of this monograph fits a widely used organization scheme in statistical analysis and research



**Figure 1.3** Classification of randomized experiments (based on Kirk, 1995, 2003a)

**Table 1.2** Methods of Randomization of Between-Subjects Experimental Designs

Method of Randomization		Comments	Design <sup>a</sup>
<i>Nonrestricted Randomization</i>			
SRA-ep	Simple Random Assignment with equal probabilities (Method 1)	—	1, 7, 14, 18, 20
SRA-up	Simple Random Assignment with unequal probabilities (Method 2)	—	1, 7, 14, 18, 20
SRA-es	Simple Random Assignment with forced equal sizes (Method 3)	—	1, 7, 14, 18, 20
SRA-us	Simple Random Assignment with forced unequal sizes (Method 4)	—	1, 7, 14, 18, 20
SRA-es-s	Simple Random Assignment with forced equal sizes—Sequential Assignment (Method 5)	Variation of Method 3 (Time Blocking)	1, 7, 14, 18, 20
SRA-us-s	Simple Random Assignment with forced unequal sizes—Sequential Assignment (Method 6)	Variation of Method 4 (Time Blocking)	1, 7, 14, 18, 20
<i>Restricted Randomization: Blocking</i>			
BRA-rb	Blocked Random Assignment with one blocking variable (Method 7)	Extension of Method 3	2, 3, 8, 9, 15, 19, 21
BRA-2s	Two-Step Blocked Random Assignment (Method 8)	Method 3 combined with Method 7	4, 10, 11, 12, 22
BRA-Ls	Two-Way Blocked Random Assignment: Latin Squares (Method 9)	Extension and restriction of Method 7	5A, 13, 16

Method of Randomization		Comments	Design <sup>a</sup>
BRA-GLs	Blocked Random Assignment Via Graeco-Latin Squares (Method 10)	Extension of Method 9	5B, 17
<i>Restricted Randomization: Stratifying</i>			
StrRA-c	Stratified Random Assignment: Nonsequential procedure (Method 11)	Extension of Method 3 plus Last Replication Correction	Section 2.4
StrRA-s	Stratified Random Assignment: Sequential procedure (Method 12)	Extension of Methods 3 (Time Blocking)	Subsection 2.5.3
<i>Restricted Randomization: Minimizing Treatment Imbalance</i>			
MIN	Minimization (Method 13)	Combination of Methods 1 and 2	Subsection 2.5.4

<sup>a</sup>For design designation, see Figure 1.3.

**Table 1.3** Methods of Randomization of Within-Subjects (Cross-Over) Experimental Designs

Method of Randomization		Comments
<i>Nonsequential Counterbalancing</i>		
RC-ro	Random Counterbalancing (Method 14)	Variation of Method 7
PC-Ls	Positional Counterbalancing (Method 15)	Variation of Method 9
<i>Sequential Counterbalancing</i>		
SC-nr	Nonrestricted Sequential Counterbalancing (Method 16)	Application of Method 3 to Specific Sequences of Treatments
SC-rs	Restricted Sequential Counterbalancing: The Same Sequences per Group (Method 17)	Extension of Method 16
SC-rd	Restricted Sequential Counterbalancing: Different Sequences per Group (Method 18)	Extension of Method 17

*Note.* These methods apply to variations and factorial extensions (Designs 6A–6V—see Table 3.3) of Design 6 (Cross-Over Design—see Figure 1.3).

design reference books (e.g., Anderson, 2001; Keppel & Wickens, 2007; Maxwell & Delaney, 2004; Tabachnick & Fidell, 2007; Winer, Brown, & Michels, 1991). Shadish et al. (2002, p. 258), adopting a more methodological approach and omitting blocking and stratifying procedures, present diagrams for nine randomized experiments classified according to three criteria: (1) the *inclusion of pretests* (posttest-only designs vs. pretest–posttest designs), (2) the *number of experimental factors* (one-factor vs. two-factor designs), and (3) the *manipulation strategies* used (between-subjects vs. within-subjects designs). From the randomization perspective adopted here, eight designs are completely randomized designs (three variations of the *pretest–posttest design*, three variations of the *posttest-only design*, the *longitudinal design*, and the *factorial design*), and the remaining one (*cross-over design*) is a within-subjects or repeated measures design.

#### 1.4 Terminological and Notational Issues

The designation under which *randomized experiments* are known varies according to the dominant traditions in each scientific discipline (e.g., *randomized controlled trials* or *randomized clinical trials* in medicine and the health sciences) and, even within the same discipline, researchers use different labels, as is the case in some areas of psychology, where *true experiments*, defined as the “gold standard” of experimentation, are contrasted with quasi-experiments, which are done without prior random assignment of subjects to experimental conditions (see Figure 1.1). The same goes for the labels currently applied to the major categories of experimental designs: (a) *between-subjects designs* are also called *independent group designs* and (b) *within-subjects designs* are known under the interchangeable names of *repeated measures designs* and *cross-over designs* (see Figure 1.1). The terminology is even more diverse when we consider the designations given to the manipulated, controlled, and measured variables in a randomized experiment, as the reader can notice on a careful inspection of Table 1.4.

Finally, a word of caution concerning the precise meaning of the key terms *treatment*, *treatment level* (also called *treatment arm* in some areas of medicine and the health sciences), *treatment level combination*, and *experimental condition*. *Experimental conditions* correspond to the differential features of the independent variable(s) manipulation or treatment(s) implementation. In one-factor experiments, the number of experimental conditions is identical to the *levels of the treatment*,

**Table 1.4** Common Designations Given to Manipulated, Controlled, and Measured Variables in Randomized Experiments

Causes ( <i>Explanatory Factors</i> )				Effects
Experimental Factors and Pseudofactors <sup>a</sup> <i>Environmental/situational variables</i>		Classificatory Factors <sup>a</sup> <i>Personal/dispositional variables</i>		
Manipulated	Controlled	Included in the design	Controlled by randomization	Observations
<ul style="list-style-type: none"> <li>• Experimental factor</li> <li>• Treatment</li> <li>• Independent variable (active)</li> <li>• Experimental variable</li> <li>• Stimulus (variable)</li> <li>• Intervention</li> <li>• Program</li> <li>• Primary factor</li> </ul>	<ul style="list-style-type: none"> <li>• Pseudofactor<sup>b</sup></li> <li>• Settings<sup>b</sup></li> <li>• Contextual variable<sup>b</sup></li> </ul>	<ul style="list-style-type: none"> <li>• Blocking variable</li> <li>• Matching variable</li> <li>• Stratifying variable</li> <li>• Independent variable (passive)</li> <li>• Prognostic factor</li> </ul>	<ul style="list-style-type: none"> <li>• Subject variable</li> <li>• Intrinsic variable</li> <li>• Individual characteristics or attributes</li> <li>• Individual difference variable</li> <li>• Personality variable</li> <li>• Organismic variable</li> </ul>	<ul style="list-style-type: none"> <li>• Dependent variable</li> <li>• Measure</li> <li>• Outcome (measure)</li> <li>• Response (variable)</li> <li>• Behavioral variable</li> </ul>

<sup>a</sup>Classificatory factors (i.e., personal or dispositional variables) and pseudofactors (i.e., all environmental or situational variables with the exception of the focal or primary experimental factors) are generically named *nuisance*, *extraneous*, or *confounding* variables. When explicitly measured and incorporated in data analysis, they are also referred to as *covariates* or *concomitant variables*.

<sup>b</sup>Pseudofactors, settings, or contextual variables (sometimes called *nonspecific factors*) are ideally controlled by holding them constant throughout the experiment. Alternatively, they can also be incorporated in the design as covariates or secondary experimental factors. In some circumstances, pseudofactors can be handled as blocking variables (e.g., *time blocking*).

while in (multi)factorial experiments, this number equals the number of *treatment level combinations* included in the design. To make a fair trade-off between clarity of exposition and economy of words, avoiding misunderstandings and giving the reader a consistent frame of reference, we have adopted in this monograph the notational system depicted in Table 1.5.

**Table 1.5** Notational System for Treatments and Treatment Levels in Between-Subjects and Within-Subjects Experimental Designs

Design	Treatments	Treatment Levels
Between-Subjects Designs (Chapter 2)		
One treatment <sup>a</sup>	T	T <sub>1</sub> , T <sub>2</sub> , T <sub>3</sub> , . . .
Two or more treatments <sup>b</sup>	A	A <sub>1</sub> , A <sub>2</sub> , A <sub>3</sub> , . . .
	B	B <sub>1</sub> , B <sub>2</sub> , B <sub>3</sub> , . . .
	C	C <sub>1</sub> , C <sub>2</sub> , C <sub>3</sub> , . . .
	[. . .]	[. . .]
Within-Subjects Designs (Chapter 3)		
One treatment <sup>c</sup>	<i>No labeling</i>	A, B, C, . . .
Two or more treatments		
Only within-subjects treatments <sup>d</sup>	W	W <sub>1</sub> , W <sub>2</sub> , W <sub>3</sub> , . . .
	X	X <sub>1</sub> , X <sub>2</sub> , X <sub>3</sub> , . . .
	[. . .]	[. . .]
	At least one between-subjects treatment <sup>e</sup>	
Within-subjects treatments	<i>The same as previous</i>	
Between-subjects treatments	T <sub>A</sub>	T <sub>A1</sub> , T <sub>A2</sub> , T <sub>A3</sub> , . . .
	T <sub>B</sub>	T <sub>B1</sub> , T <sub>B2</sub> , T <sub>B3</sub> , . . .
	[. . .]	[. . .]

<sup>a</sup> Subsections 2.2.1 and 2.3.1 to 2.3.4, and Sections 2.4 and 2.5<sup>b</sup> Subsections 2.2.2 and 2.3.5<sup>c</sup> Sections 3.1 to 3.6<sup>d</sup> Subsection 3.7.1<sup>e</sup> Subsection 3.7.2