CHAPTER

6

Quantitative Research Designs: Experimental, Quasi-Experimental, and Descriptive

CHAPTER OUTLINE

LEARNING OUTCOMES

- Discuss five considerations when planning a research design.
- Explain the three essential components of experimental designs, and compare and contrast the following experimental designs: randomized controlled trials, crossover, factorial, and Solomon four group designs.
- Discuss the advantages and disadvantages of various experimental designs.
- Compare and contrast the nonequivalent control group and interrupted time series designs.
- Discuss the advantages and disadvantages of various quasi-experimental designs.
- Compare and contrast the descriptive cross-sectional, repeated cross-sectional, comparative, and descriptive correlational designs.
- Discuss the advantages and disadvantages of various descriptive designs.
- Read a research study and identify the design used and analyze study results.
- Distinguish between secondary data analysis and secondary research.

INTRODUCTION

Designing a research study requires making a number of decisions on the steps you will take to answer your research question(s). Like an architect, you need to prepare a blueprint for your project. If you have ever met with an architect before, you know that

the process usually starts with a lot of questions. Research design is no different. The following questions address a number of key design features that must be considered.

1. *What is the research question? Will there be an intervention?* Testing the effects of an intervention is the hallmark of experimental and quasi-experimental research. If there is an intervention with human participants, the researcher will assign participants to be exposed to the independent variable, such as a modified diet or nutrient supplement, or be part of the control group. Experimental and quasi-experimental designs are used to test a hypothesis.
2. *Instead of an intervention, will researchers observe study participants and take measurements?* For example, researchers might observe a group over a longer period of time to see if exposure to certain factors (such as a diet high in fruits and vegetables) affects their risk of disease. This type of design is called a cohort study design. It is commonly used in the field of **epidemiology**, a discipline within public health that looks at the rates of health-related states (such as disease) in different groups of people and why they occur, and then looks at how this information can be used to control health problems. Study designs used in epidemiology are discussed in Chapter 7.
3. *What are the variables?* What comparisons are going to be made between or within groups? Comparisons are needed to examine relationships between the independent and the dependent variable.
4. *When and how often will data be collected or measurements taken?* Many experimental studies measure the dependent variable at least before and after the intervention. Weight loss studies, for example, often take measurements for a year or more to see whether participants kept the weight off. Data may be collected at just one point in time, such as in a **cross-sectional study**, or more frequently. In a **longitudinal study**, participants are observed and measurements are taken over a long period of time. Longitudinal studies either go forward in time (prospective) or backward in time (retrospective).
5. *What will the setting be for the study?* The setting could be a hospital, community center, or other location. Some studies use multiple sites.
6. *In an intervention study with at least two groups, will the participants be randomly assigned to a group?* True experimental research involves random assignment to groups so participants each have an equal chance of receiving any of the treatments (including no treatment) under study. Quasi-experimental research does not have randomization of participants to groups.
7. *In a human intervention study, will participants, researchers, and staff be blinded from knowing to which group a participant was assigned?* Blinding helps to prevent or minimize sources of bias, such as expectation bias. **Expectation bias** is when researchers' expectations of what they believe the study results should be get in the way of accurately taking measurements and reporting results.
8. *What controls will be put in place to reduce the influence of extraneous variables?* **Extraneous variables** are factors outside of the variables being studied that might influence the outcome of a study and cause incorrect conclusions. A good quantitative design identifies and rules out as many of these competing explanations as possible.

A good research design helps you answer the research question while effectively reducing threats to design validity.

Quantitative research designs are often used to look at causal relationships, but they can also be used to look at associations or relationship between variables. Quantitative research studies can be placed into one of five categories, although some categories do vary

a bit from book to book. First are **experimental designs** with an intervention, control group, and randomization of participants into groups. Next are **quasi-experimental designs** with an intervention but no randomization. **Descriptive designs** do not have an intervention or treatment and are considered nonexperimental. They usually aim to provide information about relevant variables but do not test hypotheses. Good descriptive studies provoke the "why" questions of analytic (cause-and-effect) research. Two additional categories are epidemiologic and predictive correlational designs.

When you read about designs in this chapter, examples of studies are given to illustrate the design. The examples include some discussion of the results of statistical tests, as well as sample tables from the studies. In a quantitative study, statistics are often used to answer one of these questions:

1. Is there a *difference* among the groups?
   Example: "LA Sprouts: A Garden-Based Nutrition Intervention Pilot Program Influences Motivation and Preferences for Fruits and Vegetables in Latino Youth" (*Journal of the Academy of Nutrition and Dietetics*, 2012)
2. Is there an *association* or relationship among the variables?
   Example: "Preventable Incidence and Mortality of Carcinoma Associated With Lifestyle Factors Among White Adults in the United States" (*JAMA Oncology*, 2016)

You can often tell from the title of an article whether the study is looking at differences among groups or an association among variables.

Experimental and quasi-experimental designs have an intervention, so they involve questions about differences—often the difference between an outcome measured in the experimental and control groups. Correlational studies look at associations. **Table 6.1** shows examples of statistics that may be used to answer these two questions.

TIP

When you read a study, first read the abstract to determine whether there is an intervention. If so, the study is either experimental or quasi-experimental. If not, the study will fit into one of the other categories. If you see the word "association" in the title, the study is likely to be a descriptive, epidemiological, or predictive correlation design.

EXPERIMENTAL STUDY DESIGNS

To be considered an experimental design, the following must be present.

1. *An intervention or treatment*. The researcher manipulates the independent variable by, for example, requiring the intervention group to eat a diet that has been modified, take a supplement containing a nutrient or phytochemical, or take part in an educational program.
2. *Control for extraneous variables*. Various control techniques, such as randomization and having a control group, are used. Having a control group allows the researcher to compare and evaluate the performance of the experimental group on the outcome (dependent) variable.
3. *Randomization*. The researcher randomly assigns each participant to a group so that each person has an equal chance of being in either group. This removes the problem of selection bias so that comparable, balanced groups of similar size are

Table 6.1 Statistics That Look at Differences and Statistics That Look at Associations

Statistics That Look at Differences

Name	Test statistic	Purpose	Number of groups	Measurement level of dependent variable
Independent samples *t*-test	*t*	To test the difference between the means of two independent groups.	2	Interval/ratio
Paired samples *t*-test (or dependent *t*-test)	*t*	To test the difference between the means from two paired groups (such as before-and-after observations on the same subject).	2	Interval/ratio
One-way analysis of variance (ANOVA)	*F*	To test the difference among means of more than two independent groups for one independent variable (with more than one level).	More than two groups	Interval/ratio
Two-way analysis of variance (ANOVA)	*F*	To test the difference among means for two independent variables, of which each can have multiple levels.	More than two groups	Interval/ratio
Repeated measures ANOVA (one-way within-subjects)	*F*	To test the difference among three or more means in the same group over time. (Extended design of dependent samples *t*-test).	One group	Interval/ratio
Chi-square	χ^2	To analyze nominal and ordinal data to find differences between groups.	Two or more groups	Nominal/Ordinal

Statistics That Look at Associations

Name	Test statistic	Purpose	Measurement level of dependent variable
Pearson product-moment correlation	*r*	To measure the strength and direction of the relationship between two variables.	Interval/ratio
Spearman rank-order correlation	ρ	To measure the strength and direction of the relationship between two variables. (Nonparametric version of Pearson product-moment correlation)	Ordinal, interval, or ratio
Linear regression		To predict the value of a dependent variable and measure the size of the effect of the independent variable on a dependent variable while controlling for covariates.	Interval/ratio
Logistic regression		Same as linear regression; used when dependent value is binary.	Binary/dichotomous

formed. Randomization also forms the basis for statistical testing. To randomize participants, researchers first generate random numbers (using either computer software or a random number table) and use them to assign each participant to a group. This is referred to as **simple randomization**.

In a **randomized block design**, the participants are first split into homogeneous groups, or blocks, before being randomly assigned to either the treatment or control group within the block. Blocks are used to decrease the variability of the sample and to control the effects of a characteristic that could influence the outcome, such as sex, age, weight, or severity of disease. Each block generally needs at least 20 participants (Grove, Burns, & Gray, 2013). For example, a study of the effect of omega-3 fatty acids on adults with a history of heart disease may be randomized by age (see **Figure 6.1**).

Random assignment means that the groups will be comparable, and differences between the groups at the end of the experiment can be deduced as being a result of the intervention. Without randomization of participants into groups, a study is considered to be quasi-experimental.

TIP

Randomization into the treatment or control group is essential to an experimental study, but it is *not* essential that the participants be randomly selected from a target population *before* being randomly assigned to a group. Most randomized controlled trials do *not* use random sampling to pick who is in the study, but all do use random assignment to groups.

When reading research, you will come across two similar terms—control group and comparison group—and probably will wonder what the difference is. The terms are often used interchangeably, but there is a difference. Technically, a control group is chosen by random sampling of the target population, and a comparison group is chosen using

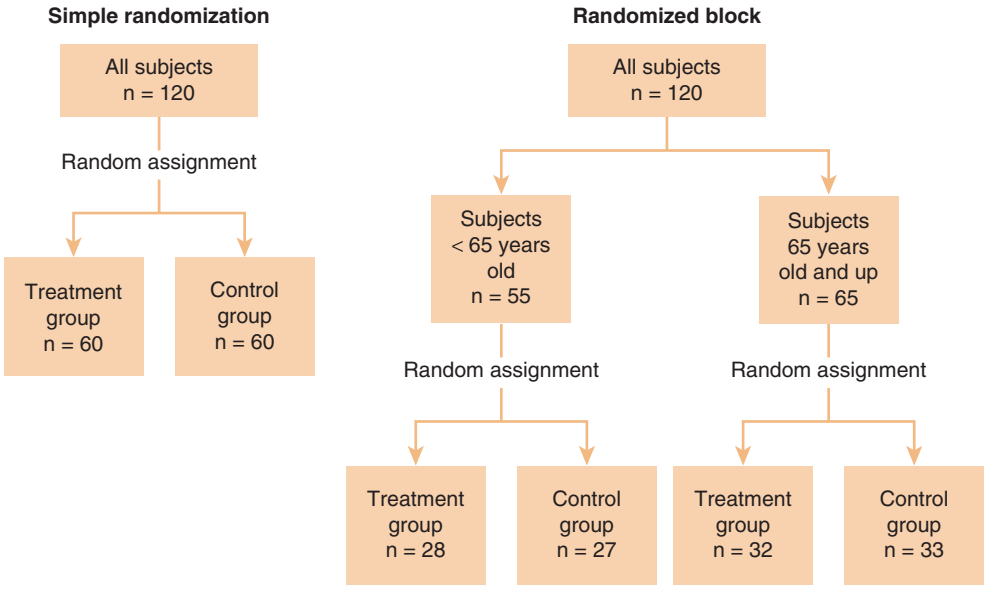

FIGURE 6.1 Simple Randomization and Randomized Block

a convenience sample. In much experimental research, the preference is for "control group." In some areas of epidemiology the preference is for "comparison group," in part because the studies do not have an intervention but do have groups they want to compare. We will use the terms *control group* and *comparison group* interchangeably unless noted.

Whether called a control or a comparison group, the researcher has a number of options to choose for this group:

1. *No intervention or treatment.*
2. *A* **placebo**. A placebo is an intervention with no effect, such as a dummy pill.
3. *Standard or usual health care*. In nursing studies, for example, patients in the control group typically receive "usual care" because no care would be unethical.
4. *A lower dose of treatment or an alternative treatment*. For example, in an experimental study examining the effectiveness of high-dose vitamin D in reducing falls and increasing lower extremity function, the control group received a low-dose of vitamin D_3, and the two experimental groups received higher doses (Bischoff-Ferrari et al., 2016).

When the control group receives no treatment, you will see a greater difference between the groups than with a placebo, usual care, or some treatment. This also makes it easier to show statistical significance.

Experimental designs are most useful with questions about therapy, such as "Which treatment options are most effective?" They can also help answer some questions about prevention, such as "Will a vitamin supplement prevent a condition?" But experimental designs are not useful for prognosis (likely course of a disease) questions. Cohort studies are much better at answering that type of question.

Keep in mind that in some situations experimental designs either cannot be used or would be unethical. Now we will look at specific experimental designs including randomized controlled trials and clinical trials, crossover designs, factorial designs, and Solomon four-group designs.

Randomized Controlled Trials and Clinical Trials

Randomized controlled trials (RCTs) are considered the "gold standard" for evaluating the effect of an intervention, treatment, or program. Participants are randomly assigned to the intervention or control group, and then followed forward in time (prospective) to compare the outcomes. Randomization, when done properly, creates equivalent groups so that differences between the groups can be attributed to the independent variable(s). RCTs are most often categorized as efficacy studies because they are designed to test hypotheses under ideal and controlled circumstances (as opposed to effectiveness studies, which are done under real-world conditions).

In an RCT, researchers need to carefully define how the participants will be randomized, the intervention, what the control group will do (if anything), what other research controls will be implemented, and other aspects of the experiment. Procedures for an RCT are normally documented in a study protocol. The **study protocol** explains the purpose of the study as well as all the details involved in carrying it out.

Random allocation of participants into groups involves implementing the random sequence in a way that conceals the sequence from anyone who enters participants in a study to prevent selection bias. If people who enter participants in a study know, or can detect, the upcoming allocations (such as the next participant will go into the experimental group), they may channel certain participants to a certain group. For example, a researcher may want a participant with a better prognosis to get into the experimental

group. This type of practice in a study may bias the estimate of the treatment effect by 30 to 40% (Moher et al., 1998; Schulz, Chalmers, Hayes, & Altman, 1995).

Pretest data, sometimes called baseline data, is useful along with demographic data to evaluate whether randomization really produced equivalent groups. Normally, this comparison of baseline data is discussed right at the beginning of the Results section and is displayed in one or more tables.

Randomization facilitates blinding, another feature of RCTs. *The Cochrane Handbook* defines **blinding** as follows:

> In general, blinding (sometimes called masking) refers to the process by which study participants, health providers and investigators, including people assessing outcomes, are kept unaware of intervention allocations after inclusion of participants into the study. Blinding may reduce the risk that knowledge of which intervention was received, rather than the intervention itself, affects outcomes and assessments of outcomes. (Higgins & Green, 2011, Box 8.11.a)

Appropriate blinding can help reduce sources of bias such as performance bias (systematic differences between groups in the care provided).

In a **single-blind study**, participants are not told whether they are in the experimental or the control group. In a **double-blind study**, two groups have been blinded—normally the participants and one or all of these groups: health care providers, data collectors, data analysts, and the researchers themselves (who may have a number of roles in the study such as data collector). Because the term *double-blind* lacks a standard definition, you will need to read the study to see who was really blinded.

When someone involved in a research study is responsible for assessing participant outcomes and knows which intervention a participant received, that person could bias how the outcome was measured (usually reporting greater effects in the treatment group). Lack of blinding in RCTs has been shown to inflate the effect of the intervention by 9% (Pildal et al., 2007).

Blinding is possible in some, but not all, studies. In a drug study, both groups can take pills as long as they look and taste the same. In a diet study where the intervention group follows a modified diet and the control group eats their usual diet, blinding is not possible. Also keep in mind that even when a researcher can use blinding, it is not a simple procedure and it does not always work perfectly.

Attrition is another concern when conducting RCTs. If dropouts and noncompliant participants are excluded from the data, it can cause a number of problems: it reduces sample size and may disrupt the balance of characteristics in each group, thereby biasing the results. For example, if more participants in the experimental group drop out than from the control group (perhaps the intervention caused some of this), this creates an imbalance. A technique called **intention-to-treat analysis** is used to prevent biases due to participant attrition. Intention-to-treat is the principle that all participants *are used in the statistical analysis,* regardless of whether they dropped out, did not receive all the treatments, or did not comply with the treatments. Intention-to-treat has both supporters and detractors and advantages and disadvantages, but it is generally preferred. Some researchers modify intention-to-treat by, let us say, excluding certain participants.

When reading results of an RCT, examine how the study handles the following to determine possible sources of bias:

1. Power calculation to determine sample size.
2. Randomization and allocation to groups.
3. Type(s) of blinding used (if any).
4. Follow-up of participants and intention-to-treat analysis (if used).

5. Data collection.
6. Precise, complete results.

In a **clinical trial**, researchers test new treatments, drugs, or medical devices with human participants to assess efficacy and safety. Clinical trials are intervention studies that often use an RCT design. Sometimes you also hear the term **controlled clinical trials** (CCT). Controlled clinical trials do use a control group but may not assign participants to the intervention or control group in a strictly random manner, making them quasi-experimental studies.

The FDA requires (and regulates) clinical trials before a new drug, medical products such as vaccines, or medical device is sold in the United States. Clinical trials are often done in stages or phases, each designed to answer a different research question (National Institutes of Health, 2008).

Phase I: Researchers test a new drug or treatment in a small group of people for the first time to evaluate its safety, determine a safe dosage range, and identify side effects.

Phase II: The drug or treatment is given to a larger group of people to see if it is effective and to further evaluate its safety.

Phase III: The drug or treatment is given to large groups of people to confirm its effectiveness, monitor side effects, compare it to commonly used treatments, and collect information that will allow the drug or treatment to be used safely.

Phase IV: Studies are done after the drug or treatment has been approved and marketed to gather information on the drug's effect in various populations and any side effects associated with long-term use.

Clinical trials may be carried out in multiple locations simultaneously. There are a number of advantages to such multicenter studies: increased sample size, a more representative sample, more cost-effective, and increased generalizability of results. Coordination and communication in multicenter studies is, of course, more challenging than a single-center trial.

TIP

The U.S. National Institutes of Health maintains ClinicalTrials.gov, a Web site that is a registry of more than 200,000 public and privately supported clinical studies of human participants in the United States and around the world. The Web site also contains results and is used by patients, researchers, students, and study record managers.

Most RCTs use a pretest-posttest design, as shown in **Table 6.2**. The pretest and posttest are designated respectively as "O_1" and "O_2." Think of "O" as an observation in which data are collected and measurements taken. Many RCTs are simply variations of these designs; some may use multiple experimental groups, perhaps receiving treatment that varies by intensity, frequency, or duration. Sometimes there may be more than one comparison group, such as one comparison group that receives no treatment and another comparison group that receives a placebo (as in a drug study).

Table 6.2 Randomized Controlled Trial Design with Pretest-Posttest

Random assignment	Experimental group	O_1	Treatment	O_2
	Control group	O_1		O_2

Let's take a look at a research study using the randomized controlled trial design. Ramly, Ming, Chinna, Suboh, and Pendek (2014) completed an RCT, which was also a Phase II clinical trial, about the effect of vitamin D supplements (independent variable) on cardiometabolic risks and health-related quality of life (dependent variables) with urban premenopausal women in Malaysia.

1. *Participants*. A power analysis using 80% power revealed that 88 participants would be required for each group. Participants were recruited, screened, and given baseline clinical and other measurements. One group was randomized and started the study in October 2012 (n = 93), and the second cohort was randomized and started in January 2013 (n = 99). A total of 171 participants completed the 12-month follow-up, but because the researchers followed an intention-to-treat protocol, the data from all 192 were used in the statistical analysis.
2. *Measurements*. The cardiometabolic risk factors that were measured included blood pressure, insulin resistance, triglycerides, and HDL. Additional data was collected such as BMI and blood levels of vitamin D as serum 25(OH)D. Baseline measurements were taken, as well as measurements at 6 months and 12 months (end of study). Participants completed a health-related quality-of-life questionnaire as part of the baseline data and again at 12 months.
3. *Intervention*. The experimental group received 0.5 grams of cholecalciferol powder taken orally by diluting the powder in warm water once a week for 8 weeks (equivalent to 7142 IU/day) and then once a month for 10 months (equivalent to 1667 IU/day). The control group received 0.5 grams of placebo taken orally exactly like the experimental group. The placebo group was provided with cholecalciferol for four months *after* the trial was done.
4. *Randomization and blinding*. The randomization sequence was created using software. The participants' names were matched with the random number sequence and printed on tubes filled with vitamin D powder or placebo. The tubes were identical and not labeled with their contents to maintain allotment concealment. Allocation of the participants into groups was only known by one staff member with no other involvement in the trial. Outcome measurements were concealed from the researchers, staff, and participants. Because many of the outcomes being measured involved blood work analyzed in a university hospital, there were not many ways for someone to change results.
5. *Statistical analysis*. Comparison of *baseline characteristics* of the participants used independent t-tests (with normally distributed data) or Mann–Whitney tests (with non-normally distributed data) for continuous variables, and chi-square tests for categorical variables.

   The *outcome measurements* were examined statistically using a linear mixed effects model. This model is an extension of a linear regression model, and it works well here because there are multiple measurement points and some data was missing. The linear mixed effects model provides the **mean difference**, which shows the absolute difference between the groups, and estimates how much the intervention changed the outcome on average compared with the control group. Probability is used to determine whether the means are significantly different.
6. *Results*. **Table 6.3** shows some of the results. The first column shows the mean value of the outcome (such as systolic blood pressure) for the intervention group, and then for the placebo group. The third column shows the mean difference, which is simply the mean for the intervention group minus the mean for the placebo group. For example, the mean systolic blood pressure for the intervention group at 12 months was 125.8, and 123.9 for the placebo group. The mean

Table 6.3 Summary of Selected Outcome Measurements Over Time Using Linear Effects Model

	Intervention (*n* = 93) Mean (95% CI)	Placebo (*n* = 99) Mean (95% CI)	Mean difference (95% CI) between treatment groups
Systolic BP			
Baseline	121.6 (118.4 to 124.8)	118.9 (115.8 to 121.9)	2.71 (–1.71 to 7.13)
6 months	126.3 (122.9 to 129.6)	123.9 (120.7 to 127.2)	2.38 (–2.27 to 7.04)
12 months	125.8 (122.6 to 128.9)	123.9 (120.8 to 126.9)	1.89 (–2.56 to 6.35)
Diastolic BP			
Baseline	77.77 (75.56 to 79.99)	76.79 (74.65 to 78.94)	0.976 (–2.107 to 4.059)
6 months	79.74 (77.38 to 82.10)	79.23 (76.97 to 81.55)	0.508 (–2.805 to 3.820)
12 months	77.52 (75.22 to 79.81)	76.76 (74.53 to 78.99)	0.757 (–2.441 to 3.954)
Se Glucose (mmol/l)			
Baseline	5.07 (4.88 to 5.25)	4.93 (4.75 to 5.12)	0.13 (–0.13 to 0.39)
6 months	5.14 (4.97 to 5.32)	5.07 (4.89 to 5.24)	0.07 (–0.17 to 0.32)
12 months	5.04 (4.83 to 5.26)	5.11 (4.90 to 5.32)	–0.07 (–0.37 to 0.23)
Se insulin (mU/L			
Baseline	13.81 (10.38 to 17.24)	11.07 (7.74 to 14.39)	2.74 (–2.04 to 7.51)
6 months	13.11 (11.18 to 15.04)	12.17 (10.27 to 14.06)	0.943 (–1.75 to 3.65)
12 months	13.93 (11.47 to 16.38)	12.74 (10.36 to 15.12)	1.19 (–2.23 to 4.61)
HOMA-Insulin Resistance			
Baseline	3.72 (2.25 to 5.19)	2.47 (1.04 to 3.91)	1.25 (–0.81 to 3.31)
6 months	3.12 (2.57 to 3.67)	2.84 (2.29 to 3.38)	0.28 (–0.49 to 1.05)
12 months	3.19 (2.61 to 3.78)	2.99 (2.42 to 3.56)	0.21 (–0.61 to 1.03)
TG (mmol/l)			
Baseline	1.15 (1.03 to 1.26)	1.17 (1.06 to 1.29)	–0.03 (–0.19 to 0.14)
6 months	1.38 (1.25 to 1.51)	1.19 (1.07 to 1.32)	0.19 (0.01 to 0.37)*
12 months	1.36 (1.23 to 1.49)	1.22 (1.09 to 1.35)	0.14 (–0.33 to 0.33)
HDL-C (mmol/l)			
Baseline	1.45 (1.36 to 1.54)	1.44 (1.35 to 1.52)	0.01 (–0.11 to 0.13)
6 months	1.43 (1.36 to 1.49)	1.50 (1.44 to 1.57)	–0.08 (–0.18 to 0.02)
12 months	1.52 (1.44 to 1.59)	1.53 (1.43 to 1.58)	0.02 (–0.09 to 0.13)

* Significant at $P < 0.05$

Modified from "Effect of Vitamin D Supplementation on Cardiometabolic Risks and Health-Related Quality of Life among Urban Premenopausal Women in a Tropical Country—A Randomized Controlled Trial," by M. Ramly, M. F. Ming, K. Chinna, S. Suboh, & R. Pendek, 2014, *PLOS ONE*, *9*, e110476. Reprinted with permission.

difference is 1.89. In this table, if these means were statistically significant, the authors would place an asterisk next to the mean difference.

There was no significant effect of vitamin D on blood pressure, insulin resistance, triglycerides, or HDL between the groups (all $P > 0.05$) except for the effect on triglycerides at 6 months. The results from the

health-related quality-of-life questionnaire showed small but significant improvement in vitality (mean difference: 5.041; 95% CI: 0.709 to 9.374) and mental component score (mean difference: 2.951; 95% CI: 0.573 to 5.329) in the intervention group compared to the placebo group.

The CONSORT 2010 checklist (see Appendix C) describes information to include when reporting a randomized trial. (CONSORT stands for Consolidated Standards of Reporting Trials.) The CONSORT group, a panel of experts, developed this checklist to increase the transparency of RCTs and to reveal when there are deficiencies (Schulz, Altman, & Moher, 2010).

Advantages of RCTs include good internal validity and the use of powerful statistical tests, such as analysis of variance (ANOVA), to analyze data. RCTs often can be used in meta-analysis. RCTs are the most appropriate research design to answer research questions on treatment or therapy. As for disadvantages, randomized experiments do tend to be costly and time-consuming. RCTs often suffer from noncompliance and dropouts (sometimes due to side effects), and participants may respond differently because they know they are being observed and assessed (known as the **Hawthorne effect**). As in many research studies, their external validity (ability to generalize results) may have limitations. Threats to external validity are minimized when a broadly representative sample is used and the setting is not too controlled.

Crossover Designs

In a crossover design, each participant acts as a member of both the experimental and the control group. Studies designed to compare two different groups of participants are referred to as **between-groups design**. Crossover designs are **within-groups design** because the researcher is making comparisons within the same participants.

The most common crossover design is the two-period, two-treatment design. Participants are randomly assigned to receive either the treatment in period 1 and the control in period 2, or the reverse. For example, in a drug study, one participant initially received the active drug and then later received the placebo. To avoid **carryover effects** (when exposure to a treatment affects outcomes in a later period), researchers build in a period of time—called a **washout period**—between treatments for the effect of the treatment to disappear.

Crossover studies include a design feature known as **repeated measures**. When you see "repeated measures" in a study, it means multiple, repeated measurements are being taken, not just a pretest and posttest.

Troup et al. (2015) used a crossover design to study the effects of black tea intake (specifically the flavonoids in tea) on blood cholesterol levels on participants with mild hypercholesterolemia. Participants were block-randomized by sex to drink either 5 cups/day of black tea or 5 cups/day of the placebo for the first 4 weeks. The placebo was a caffeinated beverage that looked and tasted like tea but contained no flavonoids. After the 3-week washout period, participants switched assignments. **Figure 6.2** shows the design for this study.

During the treatment periods, participants were provided and consumed a low-flavonoid diet. During the run-in periods (13 days each), participants drank the tea-like placebo. Participants were allowed to add sugar, but not milk, to either beverage (milk reduces the antioxidant capacity of tea).

The study did not show that black tea significantly changed the lipid profile of the participants. As in the study just mentioned, the researchers here also looked at the mean difference of an outcome, such as LDL-C. None had a *p*-value below 0.05, the level of significance.

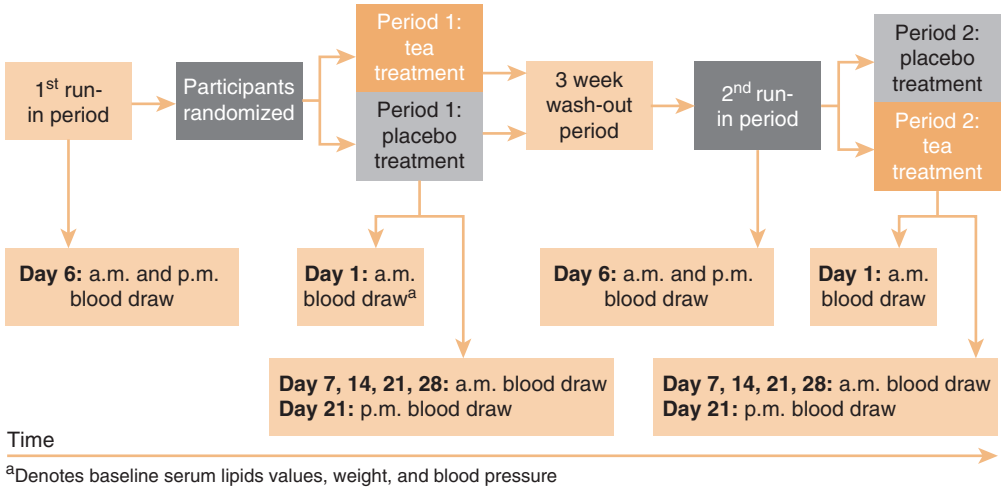

FIGURE 6.2 Crossover Study Design Overview, including timing of Biological Sample Collection Points, in a Study of the Effect of 5 Cups Per Day of Black Tea on Serum Cholesterol Concentrations (n = 57).

Reproduced from "Effect of Black Tea Intake on Blood Cholesterol Concentrations in Individuals with Mild Hypercholesterolemia: A Diet-Controlled Randomized Trial." by R. Troup et al., 2015, *Journal of the Academy of Nutrition and Dietetics*, *115*, 265. Copyright 2015 by the Academy of Nutrition and Dietetics. Reprinted with permission.

APPLICATION 6.1

What might have happened if the placebo was not caffeinated (the black tea, of course, had some caffeine)? What might have happened if the participants could tell the difference between the regular tea and the placebo? Could this have affected the results?

One advantage of the crossover design is that you do not have to worry about the comparability of two groups as you would in a parallel-group experiment (experimental and control groups), so this improves internal validity. Because each participant acts as his or her own control, you minimize the effect of confounding variables. You also can use smaller groups than in parallel-group studies.

FACTORIAL DESIGNS

In real life, variables rarely exist in isolation, so some designs include more than one independent variable. One design that manipulates two or more independent variables (or treatments) is a **factorial design**. In this design, the independent variables are referred to as **factors**. The simplest factorial design includes two factors, and each factor has two levels, resulting in a 2 × 2 factorial design. The first number "2" refers to the number of levels of the first independent variable, and the second number "2" refers to the number of levels for the second independent variable.

For example, a factorial, double-blind design was used to test the effect of zinc and multivitamins supplements on growth of infants in Tanzania (Locks et al., 2016). The two independent variables, or factors, were zinc supplements and multivitamin supplements. The two levels of the zinc supplements were administration of just the zinc supplement *or* administration of the zinc supplement with the multivitamin supplement. The two levels of the multivitamin supplements were administering just the multivitamin *or* administering it with the zinc supplement. Because one level of each variable is identical (infants

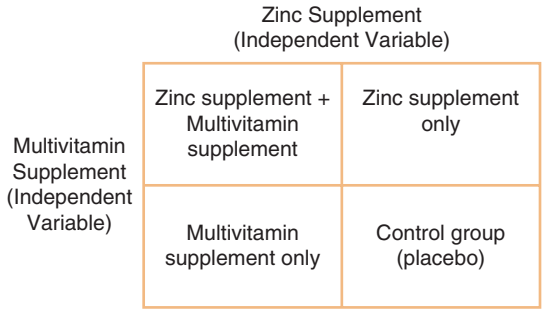

FIGURE 6.3 Example of a 2 × 2 Factorial Design

receiving both the zinc and multivitamin supplements), you can use three groups and a control group as shown in **Figure 6.3**. Locks et al. found that daily zinc supplements starting in infancy had small but significant improvement in weight for age.

This was an example of a simple factorial design. You could, for example, design a diet/exercise study with four diets and three exercise programs. That would result in 12 combinations of diet/exercise, leading to a 4 × 3 factorial design.

Solomon Four-Group Design

The **Solomon four-group design** (**Table 6.4**) is a combination of the pretest-posttest design and the posttest only design. In this design, participants are randomly assigned to one of two intervention groups or one of two control groups. Both intervention groups receive the same intervention; the only difference is that one of these groups receives the pretest, the other does not. Likewise, only one of the control groups receives the pretest. Posttest measures are collected on all four groups to assess the effect of the independent variable. Some researchers modify this design and use just one control group, which receives both the pretest and the posttest.

Atlantis, Salmon, and Bauman (2008) used a Solomon four-group design to explore the effects of television advertisements (independent variable) promoting physical activity on children's preferences for physical or sedentary activities (dependent variables). The children were randomized to one of two treatment groups or one of two control groups. The treatment groups watched a television show with standard advertisements and also advertisements promoting more physical activity instead of sedentary activity. The control groups watched the same show but without the advertisements promoting physical activity. One experimental group and one control group were assessed before *and* after watching the television show for their choices, preferences, and ratings of physical and sedentary activities. The other groups were only assessed after watching the television show. The study did not show any significant differences between groups.

This type of design is useful when a researcher thinks the outcomes could be biased by exposure to the pretest. In general, the Solomon four-group design is considered a very rigorous design that strengthens both internal and external validity. As you can

Table 6.4 Solomon Four-Group Design

Random assignment	Experimental group 1	O_1	Treatment	O_2
	Control group 1	O_1		O_2
	Experimental group 2		Treatment	O_2
	Control group 2			O_2

imagine, this design is more time consuming for researchers and also requires a large sample due to the four groups.

QUASI-EXPERIMENTAL DESIGNS

Quasi-experimental designs have an intervention and manipulation of the independent variable, but they lack a key feature of experimental studies—randomization. Because we are unsure if the groups are truly equivalent, quasi-experimental designs are ranked lower than experimental studies as sources of evidence.

Two of the most popular quasi-experimental designs are nonequivalent control group and time series designs. Be cautious when you see a quasi-experimental study that does not have a control group. Without a control group, a study has little, if any, external or internal validity.

NONEQUIVALENT CONTROL GROUP DESIGNS

In most cases, **nonequivalent control group design** is similar to the classic experimental design except that participants are not randomly assigned to groups. Often researchers use natural groups or assign participants to groups using a nonrandom method. Sampling is still going on in terms of choosing the study's participants. It is just that participants do not have the same chance of being in either the experimental or control group, and as a result, the groups are not necessarily equivalent.

Some researchers match participants at the group level based on demographic or other possible confounding variables. The more similar the groups are, the closer the design approximates an experimental study. Researchers confirm whether two groups are comparable (especially on the dependent variable) at baseline by collecting and analyzing pertinent data, but that may not include all baseline differences in active variables.

Table 6.5 shows a nonequivalent control group design with a pretest and posttest. There are a number of variations on this design, such as posttest only with a control group (sometimes a pretest is not possible or would flaw the results) or pretest and posttest with two comparison treatments and a routine care comparison group.

McAleese and Rankin (2007) carried out a nonequivalent control group study to "determine whether adolescents who participated in a garden-based nutrition intervention would increase their fruit and vegetable consumption more than those participating in a nutrition education intervention without any garden activities (McAleese & Rankin, 2007, p. 662). This study appears in Appendix A.

1. *Participants*: A convenience sample of 99 sixth-grade students in three different schools were the participants. Two schools had the experimental groups and one school had the control group.
2. *Measurements*: All students took pretests (three 24-hour food recalls) and posttests (three 24-hour food recalls).
3. *Intervention*: Both experimental groups participated in a 12-week nutrition education curriculum, "Nutrition in the Garden." Experimental school 2 also participated in gardening activities, maintaining and harvesting a garden with vegetables, herbs, and strawberries. The control group received no intervention.

Table 6.5 Nonequivalent Control Group Pretest/Posttest Design

Experimental group	O_1	Treatment	O_2
Control group	O_1		O_2

4. *Results*: Using repeated measures ANOVA, results showed that students in experimental school 2 ate significantly more fruits, vegetables, vitamin A, vitamin C, and fiber after the intervention compared to before the intervention. For example, fruit consumption increased by 1.13 servings ($P < 0.001$) for students in experimental school 2. They also ate more fruits and vegetables than the other experimental group or the control group.

This design is not as strong in controlling for threats to internal and external validity as is a true experimental design. However, the hallmark of a good quasi-experimental study is that the researchers have instituted controls and sometimes a more natural setting is advantageous. This design can show the effect of an intervention, as well as associations and trends. When randomization is impractical or unethical, a nonequivalent control group design can be useful.

APPLICATION 6.2

In the results table for the McAleese and Rankin study (Appendix A), did the control school or experimental school 1 increase consumption of fruits or vegetables? What does the column marked "F" mean? Which additional statistical test was done (in addition to repeated measured ANOVA) to pinpoint which groups were significantly different from each other?

Interrupted Time Series Designs

An interrupted time series design includes several waves of observation in which the dependent variable is measured before *and* after an intervention (which is the "interruption"). The **simple interrupted time series design** is shown in **Table 6.6**, and you can see that this design does not include a control/comparison group. The use of multiple observations/measurements does strengthen this design. It can also help assess trends in scores and decrease the chance of regression to the mean. **Regression to the mean** occurs when very high or low pretest scores of participants move closer to the mean on the posttest (due to natural variability), which may lead to an inaccurate conclusion that the intervention resulted in a treatment effect. Having several posttest scores is helpful because it is unlikely that small differences will be maintained if the treatment really has no effect.

Repeated measurements can create concerns about testing effects, instrumentation, and consistency in measurements. You also need to consider whether some unanticipated events occurred during this time and whether attrition will be more of an issue due to the multiple points of measurement over a longer time frame than many other designs.

The addition of a control group to this design strengthens the validity of the findings. Researchers can now look for differences in trends between the groups and control

Table 6.6 Times Series Designs

Simple Interrupted Time Series Design	
Experimental group	$O_1\ O_2\ O_3\ X\ O_4\ O_5\ O_6$
Interrupted Time Series Design with Control Group	
Experimental group	$O_1\ O_2\ O_3\ X\ O_4\ O_5\ O_6$
Control group	$O_1\ O_2\ O_3\ O_4\ O_5\ O_6$

Table 6.7 Abstract of a Study Using a Simple Interrupted Time Series Design

Objectives	We examined changes in meal selection by patrons of university food-service operations when nutrition labels were provided at the point of selection.
Methods	We used a quasi-experimental, single-group, interrupted time-series design to examine daily sales before, during, and after provision of point-of-selection nutrition labels. Piecewise linear regression was employed to examine changes in the average energy content of entrees and a paired *t* test was used to detect differences in sales across the periods.
Results	The average energy content of entrees purchased by patrons dropped immediately when nutrition labels were made available at point of selection and increased gradually when nutrition information was removed. There was no significant change in number of entrees sold or in revenues between the two periods.
Conclusions	Use of nutrition labels reduced the average energy content of entrees purchased without reducing overall sales. These results provide support for strengthening the nutrition labeling policy in food-service operations.

Reproduced from "Improving Patrons' Meal Selections Through the Use of Point-of-Selection Nutrition Labels," by Y. H. Chu, E. A. Frongillo, S. J. Jones, & G. L. Kaye, 2009, *American Journal of Public Health*, *99*, p. 2001. Reprinted with permission.

more for the history effect. Table 6.6 shows the **interrupted time series design with control group**. Some researchers also call this a *multiple time series design*. Basically a *multiple* time series design has more than one group.

Interrupted time series designs are flexible and can be used in a number of situations. This type of design is especially useful in the evaluation of community interventions when RCTs are impractical and too expensive, and you want to focus on measuring changes in behaviors and outcomes over time.

Table 6.7 contains an abstract of a study using a simple interrupted time series design (Chu, Frongillo, Jones, & Kaye, 2009). **Figure 6.4** displays the results of the study.

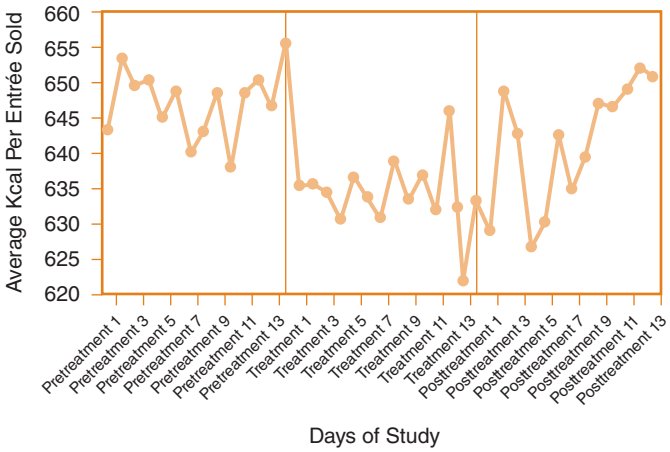

FIGURE 6.4 Average Energy Content of Entrées Sold Per Day in a Food-Service Operation before, during, and After Provision of Nutrition Information at Point of Selection: Columbus, Ohio, October 25–December 8, 2004.

Reproduced from "Improving Patrons' Meal Selections Through the Use of Point-of-Selection Nutrition Labels," by Y. H. Chu, E. A. Frongillo, S. J. Jones, & G. L. Kaye, 2009, *American Journal of Public Health*, *99*, p. 2003. Reprinted with permission

TIP

When reading a study with an interrupted time series design, you want to take a good look at the chart (such as Figure 6.4) that shows the trend of the measurements before, during, and after the intervention. The graph is split into the three segments. The average kcalories per entrée sold was high before the intervention, and then went down during the intervention. It is also interesting that once the intervention stopped, the average kcalories slowly moved back to the preintervention numbers. Looking at a graph such as this can give you a quick mental picture of what happened during a study.

DESCRIPTIVE QUANTITATIVE DESIGNS

Descriptive designs collect information about variables without changing the environment or manipulating any variables, so they do not look at possible cause and effect. They are different from observational designs in that they do not include comparison groups. According to Grove, Burns, and Gray (2013), descriptive designs "may be used to develop theory, identify problems with current practice, justify current practice, make judgments, or determine what others in similar situations are doing" (p. 215).

Descriptive designs range from cross-sectional surveys (at one or multiple points in time) to comparative designs (comparing two groups) to correlations (relationships between two variables). You can think of many descriptive designs as creating a snapshot. We now take a look at three common descriptive designs.

Descriptive Cross-Sectional and Repeated Cross-Sectional Design

In a cross-sectional study, data is collected at one point in time. A purely **descriptive cross-sectional study** provides basic information about prevalence (number of existing cases of a disease or health condition in a population) and distribution, as you can see in these examples.

- Using data from the National Health and Nutrition Examination Survey (2009–2012), researchers reported that men consumed an average of 14.6 cups of water per day, and women consumed 11.6 cups of water per day (Rosinger & Herrick, 2016). This indicates that Americans seem to be taking in adequate fluids.
- You hear frequent media reports about how American adults and children are managing their weight. Using data from the National Health and Nutrition Examination Survey (2013–2014), a group of researchers announced that 16% of children and adolescents (ages 2–19 years) are overweight and 17% are obese (Skinner, Perrin, & Skelton, 2016).
- If you ever wondered how many people really use menu labeling at fast-food or chain restaurants, researchers found that of adults who noticed nutrition labeling at fast-food or chain restaurants, 25% reported frequent use of the information, 32% reported moderate use, and 43% reported they never used it (Lee-Kwan, Pan, Maynard, McGuire, & Park, 2016). The data came, again, from a national survey: the Behavioral Risk Factor Surveillance System. In this study, data from 17 states was used.

A repeated cross-sectional study generally collects the same data at multiple points in time, and usually includes both descriptive and inferential statistics (to look at the

differences over time). Repeated cross-sectional studies can tell us about trends, patterns, or stages of development. For example, Larson, Story, Eisenberg, and Neumark-Sztainer (2016) used a repeated cross-sectional design to examine meal and snack patterns in adolescents in the Minneapolis/St. Paul secondary schools from 1999 to 2010. Food frequency questionnaires and surveys were the instruments used to gather data.

Selected results showed modest but significant changes: the frequency of eating breakfast and lunch increased over time, and the adolescents consumed fewer snacks with high kcalories and few nutrients (i.e., empty kcalorie foods/drinks). The total sample results row in **Table 6.8** shows how the frequency of eating breakfast and lunch changed from 1999 to 2010. For example, breakfast frequency went from 3.7 mean days/week to 4.2 mean days/week ($p < 0.001$). The *P*-values were calculated using two-sample *t* tests. (The 1999 sample was weighted so that you can see the trends over time independent of demographic shifts in the population.)

Advantages of cross-sectional studies include that they are relatively inexpensive, can estimate prevalence of an outcome of interest as well as assess risk factors, and do not have loss to follow-up. Cross-sectional studies are useful for generating hypotheses and for public health planning. Because they are only a snapshot, you cannot use this type of design to make causal inferences.

APPLICATION 6.3

Using Table 6.8, for which groups (such as male or black) did the frequency of eating breakfast change significantly (assume $P < 0.05$) between 1999 and 2010?

COMPARATIVE DESIGN

In a comparative design, the researchers measure the dependent variable in two or more groups, but they do not manipulate the independent variable. Descriptive and inferential statistical tests can be used to look at the differences between the groups.

For example, Mathias, Jacquier, and Eldridge (2016) looked at two groups of children aged 4 to 18 years old: those who ate lunch and those who didn't eat lunch on a given day. Using data from a 24-hour recall administered as part of the National Health and Nutrition Examination Surveys, the researchers compared the dietary intakes of the children who ate lunch with those who did not eat lunch to see if not eating lunch affected nutrient and kcalorie intake for the day. The independent variable, whether the child did or did not eat lunch, was not manipulated because this is not an experimental or quasi-experimental study. The dependent variable, the nutrient and kcalorie content of the children's diets, was measured for everyone in both groups.

TIP

As you can see from the studies just discussed, descriptive studies do not just provide descriptive statistics. Descriptive studies can compare groups and use inferential statistics such as *t*-tests to look at the relationship between variables. The next topic, correlation, also looks at the relationship between variables.

Table 6.8 Secular Trends from 1999 to 2010 in Adolescent Meal Patterns by Sociodemographic Characteristics: Minneapolis-St. Paul, MN, Project EAT (Eating and Activity in Teens)

			Breakfast frequency (mean days/wk)			Lunch frequency (mean days/wk)		
Characteristic	**1999[aa] (n)**	**2010 (n)**	**1999[aa]**	**2010**	***P* value[b]**	**1999[a]**	**2010**	***P* value[b]**
Total Sample	2,598	2,540	3.7	4.2	<0.001	5.6	5.8	<0.001
Sex								
Male	1,181	1,175	4.1	4.4	0.001	5.9	5.9	0.49
Female	1,348	1,365	3.4	4.0	<0.001	5.4	5.7	<0.001
School Level[c]								
Middle school	1,148	1,136	4.1	4.3	0.20	5.9	6.0	0.25
High school	1,335	1,404	3.4	4.1	<0.001	5.4	5.7	<0.001
Ethnicity/Race[d]								
White	540	499	4.3	4.7	0.005	5.6	6.0	<0.001
Black	638	706	3.5	4.3	<0.001	5.4	5.7	0.002
Hispanic	414	435	3.2	4.0	<0.001	5.5	5.8	0.07
Asian	546	520	3.8	3.8	0.94	6.0	5.9	0.32
Native American	98	92	3.5	4.1	0.10	5.2	5.7	0.05
Mixed/Other	293	288	4.0	4.0	0.96	5.7	5.6	0.50
Socioeconomic Status[e]								
Low	936	973	3.4	3.9	0.002	5.7	5.8	0.36
Low middle	560	556	3.4	4.0	<0.001	5.4	5.9	<0.001
Middle	436	430	3.7	4.4	<0.001	5.5	5.8	0.03
High middle	335	320	4.3	4.6	0.16	5.7	5.8	0.65
High	199	193	4.9	5.0	0.63	5.8	6.1	0.16

[a] The 1999 sample was weighted to allow for an examination of secular trends in meal patterns independent of demographic shifts in the population. For example, estimates of weekly breakfast frequency within the low socioeconomic status group in 1999 and 2010 are mutually controlled so that sex, school level, and ethnicity/race makeup are the same in the low socioeconomic status group in the 1999 sample as in the 2010 sample.

[b] *P* values represent testing to examine weighted mean differences in meal frequency between 1999 and 2010.

[c] Middle school represents students enrolled in 6th to 8th grades and high school represents students enrolled in 9th to 12th grades.

[d] Adolescents could choose >1 ethnic/racial category; those responses indicating multiple categories were coded as "Mixed/Other." Because there were few participants who identified themselves as Hawaiians or Pacific Islanders, these participants were also included in the "Mixed/Other" category.

[e] The prime determinant of socioeconomic status was the higher education level of either parent with adjustments made for student eligibility for free/reduced-price school meals, family public assistance receipt, and parent employment status.

Reproduced from: "Secular Trends in Meal and Snack Patterns among Adolescents from 1999 to 2010," by N. Larson, M. Story, M. E. Eisenberg, & D. Neumark-Sztainer, 2016, *Journal of the Academy of Nutrition and Dietetics*, *116*, 243. Reprinted with permission.

Descriptive statistics were used to show the percentage of children and adolescents who did not eat lunch: 7% ± 1% (standard error) for 4- to 8-year-olds, 16% ± 2% for 9- to 13-year-olds, and 17 ± 1% for 14- to 18-year-olds. Linear regression was used to show that missing lunch was associated with significantly lower intake of many micronutrients for the day in all age groups.

Descriptive Correlational Design

Correlation is a statistical procedure used to measure and describe the relationship or association between two variables. The researcher may not know whether the variables are related, or may suspect that one influences the other. In either case, no attempt is made to manipulate an independent variable in correlational designs, so you cannot conclude that the relationship is causal simply based on correlation.

Before a correlation coefficient can be calculated, you need to draw a **scatterplot** with the quantitative data, as seen in **Figure 6.5**. Each dot in the scatterplot represents one variable (*x* or *y*) from one person or observation. The values for the *x* variable are placed on the *x*-axis (horizontal axis), and the values for the *y* variable are on the *y*-axis (vertical axis).

The variable for the *y*-axis should be the **outcome variable**, which may also be called the response or dependent variable. The variable for the *x*-axis should be the **predictor variable**, which also may be called the explanatory or independent variable. For example, in a study on carbohydrate intake and dental caries, the researcher wants to see if increasing carbohydrate intake (predictor or independent variable) increases the number of cavities (outcome or dependent variable). So the number of cavities should be on the *y*-axis and carbohydrate intake should be on the *x*-axis. This way, when you look at the scatterplot, as the carbohydrate intake (predictor variable) increases along the horizontal axis, you can see how it affects the number of cavities (outcome variable) on the vertical axis.

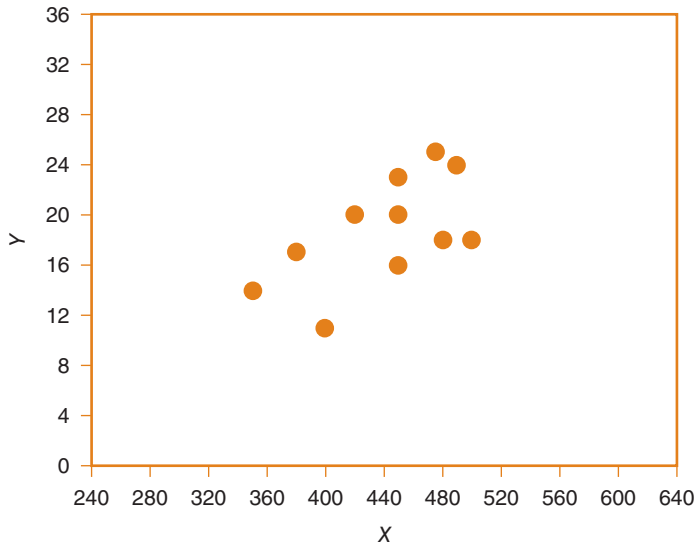

FIGURE 6.5 Example of a Scatterplot

Reproduced from *Basic Biostatistics: Statistics for Public Health Practice* (2nd ed.) by B.B. Gerstman, 2015, Burlington, MA: Jones & Bartlett Learning, p. 336. Reprinted with permission..

There are several things to look for on a scatterplot: form, direction, strength, and outliers.

1. *When the dots are closely grouped along what appears to be a relatively straight line, the form is linear.* If all the points lie on a line, then we have perfect linear correlation (this is rare!). Correlation can only be used if the shape in the scatterplot is linear. The correlation coefficient tells you how well the variables fit on a straight line. Scatterplots may also be curved, curvilinear, or random.
2. **Figure 6.6** shows scatterplots with *positive and negative directions* (discussed in a moment). Sometimes you may see both a positive and a negative trend in a scatterplot, in which case you may need to split the data into subgroups.
3. *The strength of the relationship can be seen by how tightly clustered the points are along the form,* whether the form is a straight line, a curvy line, or other shaped line. The relationship is stronger when the data points are close to an imaginary line that you draw through the points. The relationship is weak when the points are all over the place, with no pattern.
4. You should also *check for outliers* because the correlation coefficient can be greatly affected by just one outlier.

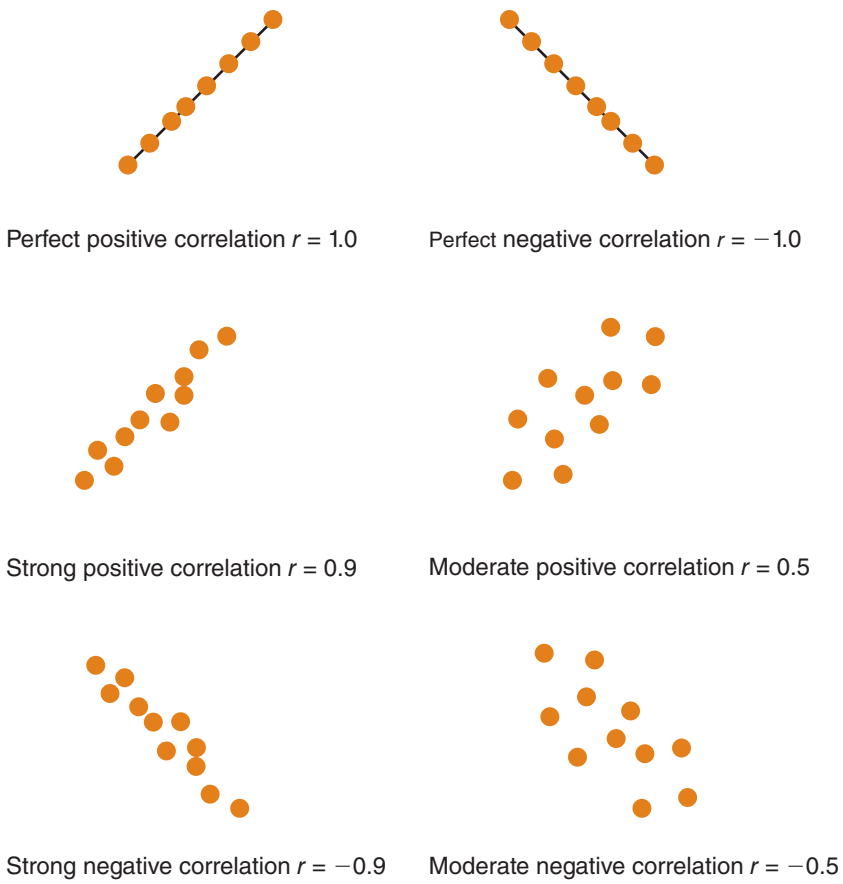

FIGURE 6.6 Scatterplots Showing Different Correlations (Positive and Negative)

Reproduced from *Basic Biostatistics: Statistics for Public Health Practice* (2nd ed.) by B.B. Gerstman, 2015, Burlington, MA: Jones & Bartlett Learning, p. 339. Reprinted with permission.

Before you can use correlation, you have to be sure that there is a linear relationship and that the variables are quantitative. You also may have to make decisions about how to handle outliers.

A descriptive correlational design has the following characteristics:

1. Two variables are clearly identified and defined.
2. Data for each of the two variables are collected.
3. There is one group of participants.
4. There is no intervention or treatment going on before, during, or after data collection.
5. Data is collected at one general point in time.
6. A correlation coefficient is calculated. The correlation coefficient is a summary statistic, similar to the mean, that summarizes the strength and direction of a linear relationship.

For continuous variables that are normally distributed, a Pearson correlation coefficient (r) is calculated. If either or both variables are ordinal, the Spearman rank-order correlation (also called Spearman's rho, r_s) is used. This is a nonparametric test that has the same values as the Pearson's correlation coefficient. Correlational techniques such as Cramer's V can be used for nominal data. When researchers are trying to understand the relationship between two variables *while controlling for the effects of other variables*, they use linear regression.

The correlation coefficient r measures both the strength and the direction of the relationship between the two variables, from no relationship (0) to a perfect positive or negative linear relationship (+1 or −1). The closer a correlation coefficient is to 1.0 or −1.0, the stronger the relationship between the variables.

There are no generally accepted standards for interpreting whether a correlation is considered strong or weak, but some would say that a strong correlation should be at least ± 0.7-0.8. To interpret a correlation coefficient, you can use its statistical significance to see whether the correlation differs significantly from 0 (no correlation), although some researchers suggest paying more attention to the size of the correlation (Guyatt, Walter, Shannon, Cook, Jaeschke, & Heddle, 1995). Also, keep in mind that large sample sizes can make small correlations look significant.

The direction of the relationship can be positive or negative. When two variables are positively correlated, it means that they vary together. For example, as BMI increases, so does systolic blood pressure: this is a positive relationship in which both variables increase. If both variables decrease at the same time, that is also a positive relationship. In a negative relationship, as one variable increases, the other one decreases. So in a negative correlation, the variables move in opposite directions and are said to have an inverse relationship. For example, researchers may find that as participants in a weight loss program spend more time exercising and planning meals, their BMI decreases.

One thing to note at this point is a key difference between scatterplots and correlation. Scatterplots use the units the variables are measured in, such as BMI, whereas the formula to calculate the correlation coefficient standardizes the variables (using z scores), so changes in scale or units of measurement do not affect its value. Therefore, the correlation coefficient has no units.

Descriptive correlational studies sometimes lay the groundwork for testing a hypothesis in a later study. A common mistake with correlation is that people think a high correlation coefficient demonstrates causation. Correlation does not demonstrate cause and effect; the independent variable is not manipulated and extraneous variables (sometimes called lurking variables) are not controlled. In some cases, one variable may actually cause the other, but other research designs will be needed to prove it.

ADDITIONAL TYPES OF DESIGNS

As you read more research articles, you will find some additional types of designs.

1. *Secondary data analysis.* Researchers use data already collected in another study. Then they use a traditional research design (obviously not experimental or quasi-experimental) to answer a research question. For example, Eicher-Miller, Khanna, Boushey, Gelfand, and Delp (2016) used data from the National Health and Nutrition Examination Survey (1999–2004) to test a hypothesis. They hypothesized that the diet quality of American adults varies depending on how and when they distribute their energy/nutrient intake over the day. **Table 6.9** lists sources of data sets for NHANES and other studies. When using a national data set, be aware that they often oversample underrepresented groups, so you need to read in detail about the sampling methods.

Table 6.9 Examples of Sources of Data Sets for Secondary Data Analysis

Centers for Disease Control and Prevention		
	Behavioral Risk Factor Surveillance System	http://www.cdc.gov/brfss/
	Youth Risk Behavior Surveillance System	http://www.cdc.gov/healthyyouth/data/yrbs/index.htm
	National Health and Nutrition Examination Survey	http://www.cdc.gov/nchs/nhanes/nhanes_questionnaires.htm
	National Health Interview Survey	http://www.cdc.gov/nchs/nhis/index.htm
	Nutrition, Physical Activity, and Obesity: Data, Trends, and Maps	https://nccd.cdc.gov/NPAO_DTM/Summary
National Institutes of Health		
	Clinical Trials (some include study results)	http://www.clinicaltrials.gov
	National Heart, Lung, and Blood Institute, Framingham Heart Study	http://www.framinghamheartstudy.org
	Women's Health Initiative	http://www.whi.org
U.S. Department of Agriculture		
	Infant Feeding Practices, Study II and Its Year Six Follow Up	http://www.cdc.gov/breastfeeding/data/ifps/index.htm
	What We Eat in America	http://www.ars.usda.gov/Services/docs.htm?docid=18354
National Collaborative on Childhood Obesity Research (A Collaboration Among CDC, NIH, USDA, and Robert Wood Johnson Foundation)		
	Catalogue of Surveillance Systems (for childhood obesity)	http://nccor.org/nccor-tools/catalogue/index
Nurses' Health Study http://www.nurseshealthstudy.org		
U.S. Renal Data System http://usrds.org		

2. *Methodological designs.* Methodological designs are used to test the reliability and validity of instruments used to measure variables in research. For example, Boucher et al. (2006) asked participants to take an adaptation of Block's food frequency questionnaire (FFQ). The same respondents also completed two 24-hour diet recalls via telephone. The researchers evaluated the agreement of 32 nutrient intakes between the adapted FFQ and the diet recalls for each respondent. Correlation coefficients showed moderate to high validity.
3. *Secondary research.* Original research is considered primary research, and review articles are known as secondary research because they analyze data already collected in primary research. Systematic reviews critically appraise and pool data from multiple single studies, often using meta-analysis, to answer a research question.

RESEARCHER INTERVIEW: Intervention Research

Leslie Cunningham-Sabo, PhD, RDN
Associate Professor, Department of Food Science and Human Nutrition, Colorado State University, Fort Collins, CO

1. **Briefly describe the areas in which you do research.**
My research focus is childhood obesity prevention. Most of my projects take place in public schools, but with some emphasis on families at home too.
2. **With your experience in intervention research, what should students/practitioners know about this area of research?**
Intervention research focuses on the development, implementation, and evaluation of programs or projects with the study population. You might read several different types of articles related to intervention research. For example, I recently published an article with my research team that described in detail the protocol (scientific procedures) our multicomponent study followed. It did not include any study results but focused on the design of our Fuel for Fun: Cooking with Kids Plus Parents and Play project.

   Much more frequently you will find articles describing some or all of the components of an intervention study and the related results. Depending on how complex the study is, one article may describe all or just part of the intervention and results.
3. **What do you enjoy most about the research process?**
I really love the creativity of the research process. You get to answer important questions, which lead to even more questions! For example, right now we are trying to understand how best to engage parents in a school-based intervention with fourth graders. Will busy parents find it easier to connect via Facebook or through a blog? Or will they prefer materials sent home with their child that they complete together? That is just one example of the questions we try to address through our research.

   Other things I enjoy about the research process are that I get to work with faculty and students from different academic disciplines (e.g., education, exercise science, public health) in addition to nutrition and dietetics. I also really enjoy mentoring students and younger professionals. I feel a sense of pride and contribution when they learn, grow professionally, and achieve their academic goals. Most important, though, is the opportunity to make a difference and improve the health and quality of life of the children and families that are part of our intervention study.
4. **What tip(s) do you have for practitioners who want to do practice-based research?**
Research can contribute to improvement in all practice settings. Think about areas where your team is struggling; it could be with delivery of client services or client outcomes, or even how your team works together. Discuss what information you need to gather to learn more about the situation. Who else do you need to involve? How will you gather the resources needed to do this work? Where are the sources of internal or external funding? Gain the necessary approvals

and move forward. Collect the information systematically, and prepare reports or presentations formatted for the audience (e.g., funder, superiors, target audience). The bottom line: identify a significant problem you want to address, find others who share your interest and commitment, gain the resources you need, and keep stakeholders apprised of your progress and results. Start small, gain experience, and achieve your initial goals. This can lead to bigger research projects in the future.

SUMMARY

1. Some research design features to consider include whether there will be an intervention, where it will take place, what comparisons will be made between or within groups, what is going to be measured, when measurements will be taken, how you will get the participants, how you will split participants up into groups if needed, and how you will control extraneous variables.
2. Quantitative research designs are often used to look at causal relationships, but they can also be used to look at associations or relationship between variables. First are experimental study designs with an intervention, control group, and randomization of participants into groups. Next are quasi-experimental designs with an intervention but no randomization. Descriptive designs do not have an intervention or treatment and are considered nonexperimental.
3. Most research questions look at either differences among groups or an association or relationship among variables.
4. Although randomization into the treatment or control group is essential to an experimental study, it is *not* essential that the participants are randomly selected from a target population *before* being randomly assigned to a group. Most randomized controlled trials do *not* use random sampling to pick who is in the study. Random assignment means the groups are comparable, so that differences between them at the end are deduced as being caused by the intervention.
5. A control or comparison group may receive no treatment, a placebo, standard or usual health care, or a lower dose of treatment or an alternative treatment.
6. Experimental designs are most useful for questions about therapy or treatment options.
7. Randomized controlled trials (RCTs) are considered the "gold standard" for assessing causality and determining efficacy in intervention research. Participants are randomly assigned to the intervention or control group, and then followed forward in time to compare the outcomes. In addition, blinding may be used. Most RCTs use a pretest-posttest design.
8. Intention-to-treat analysis is used to prevent biases due to participant attrition.
9. When reading a RCT, pay attention to how sample size was determined, how participants were randomized and allocated to groups, if blinding was used, if intention-to-treat analysis was used, how data was collected, and how complete and precise the results are.
10. Clinical trials are intervention studies that often use an RCT design. The FDA requires clinical trials before new drugs or medical products or devices are sold in the United States. Clinical trials are often done in stages or phases.
11. The most common crossover design is the two-period, two-treatment design. Participants are randomly assigned to receive either the treatment in period 1 and the control in period 2, or the reverse. To avoid carryover effects (when exposure to a treatment affects outcomes in a later period), researchers build in a period of time—a washout period—between treatments for the effect of the treatment to disappear.
12. A factorial design manipulates two or more independent variables (or treatments). In this design, the independent variables are referred to as factors. The simplest factorial design includes two factors, and each factor has two levels, resulting in a 2 × 2 factorial design. The first number "2" refers to the number of levels of the first independent variable, and

the second number "2" refers to the number of levels for the second independent variable.

13. The Solomon four-group design combines the pretest-posttest design and the posttest only design, which results in a very rigorous design.
14. Two of the most popular quasi-experimental designs are nonequivalent control group and time series designs. The nonequivalent control group design is similar to the classic experimental design except that participants are not randomly assigned to groups. An interrupted time series design includes several waves of observation where the dependent variable is measured before *and* after an intervention. This design may or may not have a control group.
15. Descriptive designs collect information about variables without changing the environment or manipulating any variables, so they do not assess cause and effect. Descriptive designs include descriptive cross-sectional (collects information at one point in time, such as the prevalence of childhood obesity), repeated cross-sectional, comparative, and descriptive correlational designs.
16. In a comparative design, the researchers measure the dependent variable in two or more groups but do not manipulate the independent variable. Descriptive and inferential statistical tests can be used to look at the differences between the groups.
17. The correlation coefficient *r* measures both the strength and the direction of the relationship between the two variables, from no relationship (0) to a perfect positive or negative linear relationship (+1 or −1). The closer a correlation coefficient is to 1.0 or −1.0, the stronger the relationship between the variables.
18. Two additional kinds of studies you will find in journals are those using secondary data analysis and methodological designs.

REVIEW QUESTIONS

1. A prospective longitudinal study:
   A. goes backward in time
   B. goes forward in time
   C. takes measurements only at one point in time
   D. any of the above
2. _____ variables are factors outside of the variables being studies that could influence the outcome of a study.
   A. independent
   B. dependent
   C. extraneous
   D. extra
3. If a study has an intervention but no randomization, in which group of study designs does it belong?
   A. experimental
   B. quasi-experimental
   C. descriptive
   D. observational
4. A control or comparison group may receive:
   A. no intervention or treatment
   B. a placebo
   C. standard or usual health care
   D. a and b only
   E. a, b, and c
5. An excellent design to test treatment options is:
   A. correlational design
   B. comparative design
   C. randomized controlled trial
   D. repeated cross-sectional design
6. Most RCTs use a posttest only design.
   A. true
   B. false
7. Crossover studies use repeated measures.
   A. true
   B. false
8. Nonequivalent control group designs generally use:
   A. randomization
   B. sampling
   C. blinding
   D. repeated measures

9. A study design that is used to gather information about how much calcium Americans are taking in each day would most likely be a:
   A. clinical trial
   B. comparative design
   C. descriptive cross-sectional design
   D. analytic cross-sectional design
10. The correlation coefficient tells you about:
   A. the direction of the relationship
   B. the strength of the relationship
   C. whether the independent variable caused the dependent variable to change
   D. a and b only
   E. a, b, and c
11. What are the three main features of experimental research?
12. Explain what intention-to-treat analysis is and why it is used.
13. Why could the Solomon four-group design be more rigorous than an RCT?
14. Describe four characteristics of a descriptive correlational design.
15. Explain what secondary data analysis is and give an example.

CRITICAL THINKING QUESTIONS

1. Read this article (see Appendix B) describing a randomized controlled trial and answer the following questions.

   Rajkumar, N., Karthikeyan, V. S., Manwar, S., Sistla, S. C., & Kate, V. (2013). Clear liquid diet vs soft diet as the initial meal in patients with mild acute pancreatitis: A randomized interventional trial. *Nutrition in Clinical Practice, 28,* 365–370.

   A. What is the objective for this study?
   B. Did this study take a prospective, retrospective, or cross-sectional approach?
   C. What are the independent and dependent variables?
   D. Describe the setting and the participants.
   E. What was the primary study outcome? Were there secondary outcomes that were measured? If so, describe.
   F. What statistical test was used to compare continuous variables? What *P*-value was considered significant?
   G. What were the results?
   H. What did the researchers conclude?
   I. Discuss the generalizability of this study to the United States.
2. Read this quasi-experimental study and answer the following questions.
   Humphrey. L., Clifford, D., & Neyman Morris, M. (2015). Health at Every Size college course reduces dieting behaviors and improves intuitive eating, body esteem, and anti-fat attitudes. *Journal of Nutrition Education and Behavior, 47,* 354–360. doi: 10.1016/j.jneb.2015.01.008
   A. What is the objective for this study? If the researchers had stated a hypothesis for this study, what do you think it would be?
   B. Why is this a quasi-experimental study?
   C. What are the independent and dependent variables?
   D. Who are the participants? Describe the intervention for the experimental group and comparison groups. What did the control group do?
   E. How many students completed both pretests and posttests by group? Briefly describe each instrument.
   F. ANOVA was used to compare mean scores of posttests among the three groups. ANOVA can tell you that there are significant differences between the groups, but it cannot pinpoint which groups. So which post hoc test was used to identify groups that had significantly different scores?
   G. Table 6.3 compares the mean scores for each instrument in two ways. First, it compares the pretest and posttest score within each group (using paired sample *t*-tests). Second, it compares the posttest scores between the three groups (using ANOVA with post hoc analysis). From pretest to posttest, where did the HAES

class experience statistically significant differences? For which instruments was the HAES posttest score significantly different from *both* the comparison and control group?

H. What did this study demonstrate?

3. Read this descriptive comparative study and answer the following questions.

Zizza, C. A., Sebastian, R. S., Wilkinson, C., Isik, Z., Goldman, J. D., & Moshfegh, A. J. (2015). The contribution of beverages to intakes of energy and MyPlate components by current, former, and never smokers in the United States. *Journal of the Academy of Nutrition and Dietetics, 115,* 1939–1949. doi: 10.1016/j.jand.2015.07.015

A. What was the purpose of this study? If the researchers had stated a hypothesis, what do you think it would be?

B. Is this study an example of secondary data analysis? If so, where did the data come from?

C. What are the independent and dependent variables?

D. Why is this a descriptive study?

E. Who are the three groups being compared?

F. Table 6.2 shows mean beverage intake in grams by beverage group and smoking status for men/women. Do male current smokers drink significantly more total beverages than nonsmokers? Do female current smokers drink significantly more total beverages than nonsmokers? Compare male and female current smokers to nonsmokers in terms of coffee and alcoholic beverage consumption. Note: Within the category of men or the category of women, means with *different* superscript letters (x, y, z) differ significantly as determined by the *t*-test.

G. Interpret Figure 6.2. If you have trouble doing that, read in the article where Figure 6.2 is discussed.

H. What did the researchers conclude?

SUGGESTED READINGS AND ACTIVITIES

1. Boushey, C., Harris, J., Bruemmer, B., Archer, S. L., & Van Horn, L. (2006). Publishing nutrition research: A review of study design, statistical analyses, and other key elements of manuscript preparation, Part 1. *Journal of the American Dietetic Association, 106,* 89–96. doi: 10.1016.j.jada.2005.11.007
2. Bruemmer, B., Harris., J., Gleason, Ph., Boushey, C. J., Sheean, P. M., Archer, S., & Van Horn, L. (2009). Publishing nutrition research: A review of epidemiologic methods. *Journal of the American Dietetic Association, 109,* 1728–1737. doi: 10.1016/j.jada.2009.07.011
3. Sheean, P. M., Bruemmer, B., Gleason, P., Harris, J., Boushey, C., & Van Horn, L. (2011). Publishing nutrition research: A review of multivariate techniques–Part 1. *Journal of the American Dietetic Association, 111*, 103–110. doi: 10.1016/j.jada.2010.10.010

REFERENCES

Atlantis, E., Salmon, J., & Bauman, A. (2008). Acute effects of advertisements on children's choices, preferences, and ratings of liking for physical activities and sedentary behaviours: A randomised controlled pilot study. *Journal of Science and Medicine in Sport, 11*, 553–557.

Bischoff-Ferrari, H., Dawson-Hughes, B., Orav, E. J., Staehelin, H. B., Meyer, O., Theiler, R., . . . Egli, A. (2016). Monthly high-dose vitamin D treatment for the prevention of functional decline: A randomized clinical trial. *JAMA Internal Medicine, 176,* 175–183. doi: 10.1001/jamainternmed.2015.7148

Boucher, B., Cotterchio, M., Kreiger, N., Nadalin, V., Block, T., & Block, G. (2006). Validity and reliability of the Block98 food-frequency questionnaire in a sample of Canadian women. *Public Health Nutrition, 9,* 84–93. doi: http://dx.doi.org/10.1079/PHN2005763

Boushey, C., Harris, J., Bruemmer, B., Archer, S. L., & Van Horn, L. (2006). Publishing nutrition research: A review of study design, statistical analyses, and other key elements of manuscript preparation, Part 1. *Journal of the American Dietetic Association, 106,* 89–96. doi: 10.1016.j.jada.2005.11.007

Bruemmer, B., Harris., J., Gleason, Ph., Boushey, C. J., Sheean, P. M., Archer, S., & Van Horn, L. (2009). Publishing nutrition research: A review of epidemiologic methods. *Journal of the American Dietetic Association, 109*, 1728–1737. doi: 10.1016/j.jada.2009.07.011

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research.* Belmont, CA: Wadsworth Cengage.

Chu, Y. H., Frongillo, E. A., Jones, S. J., & Kaye, G. L. (2009). Improving patrons' meal selections through the use of point-of-selection nutrition labels. *American Journal of Public Health, 99,* 2001–2005. doi: 10.2105/AJPH.2009.153205

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design & analysis issues for field settings*. Boston, MA: Houghton Mifflin Company.

Eicher-Miller, H. A., Khanna, N., Boushey, C. J., Gelfand, S. B., & Delp, E. J. (2016). Temporal dietary patterns derived among the adult participants of the National Health and Nutrition Examination Survey 1999–2004 are associated with diet quality. *Journal of the Academy of Nutrition and Dietetics, 116,* 283–291. doi: 10.1016/j.jand.2015.05.014

Forbes, D. (2013). Blinding: An essential component in decreasing risk of bias in experimental designs. *Evidence-Based Nursing, 16,* 70–71. doi: 10.1136/eb-2013-101382

Grove, S. K., Burns, N., & Gray, J. R. (2013). *The practice of nursing research: Appraisal, synthesis, and generation of evidence*. St. Louis, MO: Elsevier Saunders.

Gupta, S. K. (2011). Intention-to-treat concept: A review. *Perspectives in Clinical Research, 2,* 109–112. doi: 10.4103/2229-3485.83221

Guyatt, G., Walter, S., Shannon, H., Cook, D., Jaeschke, R., & Heddle, N. (1995). Basic statistics for clinicians: 4. Correlation and regression. *Canadian Medical Association Journal, 152*, 497–504.

Higgins, J., & Green, S. (Eds.). (2011). *Cochrane handbook for systematic reviews of intervention* (Version 5.1.0). Retrieved from http://handbook.cochrane.org

Karanicolas, P. J., Farrokhyar, F., & Bhandari, M. (2010). Blinding: Who, what, when, why, how? *Canadian Journal of Surgery, 53*, 345–348.

Larson, N., Story, M., Eisenberg. M. E., & Newmark-Sztainer, D. (2016). Secular trends in meal and snack patterns among adolescents from 1999 to 2010. *Journal of the Academy of Nutrition and Dietetics, 116,* 240–250. doi: 10.1016/j.jand.2015.09.013

Lee-Kwan, S. H., Pan. P., Maynard, L. M., McGuire, L. C., & Park, S. (2016). Factors associated with self-reported menu-labeling usage among US adults. *Journal of the Academy of Nutrition and Dietetics, 116,* 1127–1135. doi: 10.1016/j.jand.2015.12.015

Locks, L. M., Manji, K. P., McDonald, C. M., Kupka, R., Kisenge, R., Aboud, S., . . . Duggan. C. P. (2016). Effect of zinc and multivitamin supplementation on the growth of Tanzanian children aged 6–84 weeks: A randomized, placebo-controlled, double-blind trial. *American Journal of Clinical Nutrition, 103,* 910–918. doi: 10.3945/ajcn.115.120055

Mathias K. C., Jacquier, E., & Eldridge, A. L. (2016). Missing lunch is associated with lower intakes of micronutrients from foods and beverages among children and adolescents in the United States. *Journal of the Academy of Nutrition and Dietetics, 116,* 667–676. doi: 10.1016/j.jand.2015.12.021

McAleese, J. D., & Rankin, L. L. (2007). Garden-based nutrition education affects fruit and vegetable consumption in sixth-grade adolescents. *Journal of the American Dietetic Association, 107,* 662–665. doi: 10.1016/j.jada.2007.01.015

Moher, D, Pham, B., Jones, A., Cook, D. J., Jadad, A. R., Moher, M., . . . Klassen, T. P. (1998). Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? *Lancet, 352,* 609–613. doi: 10.1016/S0140-6736(98)01085-X

National Institutes of Health. (2008). *FAQ ClinicalTrials.gov—Clinical trial phases*. Retrieved from https://www.nlm.nih.gov/services/ctphases.html

Pildal, J., Hróbjartsson, A., Jørgensen, K. J., Hilden, J., Altman, D. G., & Gøtzsche, P. C. (2007). Impact of allocation concealment on conclusions drawn from meta-analyses of randomized trials. *International Journal of Epidemiology, 36,* 847–857.

Ramly, M., Ming, M. F., Chinna, K., Suboh, S., & Pendek, R. (2104). Effect of vitamin D supplementation on cardiometabolic risks and health-related quality of life among urban premenopausal women in a tropical country—A randomized controlled trial. *PLOS ONE, 9,* e110476. doi: 10.1371/journal.pone.0110476

Rosinger, A., & Herrick, K. (2016). *Daily Water Intake Among U.S. Men and Women, 2009–2012*. (NCHS Data Brief No. 242). Retrieved from http://www.cdc.gov/nchs/products/databriefs/db242.htm

Schulz, K. F., Altman, D. G., & Moher, D. (for the CONSORT Group). (2010). CONSORT 2010 statement: Updated guidelines for reporting parallel group randomised trials. *PLoS Medicine, 7,* e1000251. doi:10.1371/journal.pmed.1000251

Schulz, K. F., Chalmers, I., Hayes, R. J., & Altman, D. G. (1995). Empirical evidence of bias: Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA, 273*, 408–412. doi:10.1001/jama.1995.03520290060030.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* Belmont, CA: Wadsworth Cengage.

Sheean, P. M., Bruemmer, B., Gleason, P., Harris, J., Boushey, C., & Van Horn, L. (2011). Publishing nutrition research: A review of multivariate techniques–Part 1. *Journal of the American Dietetic Association, 111,* 103–110. doi: 10.1016/j.jada.2010.10.010

Skinner, A. C., Perrin, E. M., & Skelton, J. A. (2016). Prevalence of obesity and severe obesity in US children, 1999–2014. *Obesity, 24,* 1116–1123. doi: 10.1002/oby.21497

Troup, R., Hayes, J. H., Raatz, S. K., Thyagarajan, B., Khaliq, W., Jacobs, D. R., . . . Gross, M. (2015). Effect of black tea intake on blood cholesterol concentrations in individuals with mild hypercholesterolemia: A diet-controlled randomized trial. *Journal of the Academy of Nutrition and Dietetics, 115*, 264–271. doi: 10.1016/j.jand.2014.07.02