CHAPTER 5

Experimental Research Designs

If there is an ideal against which all quantitative designs are compared, it is the true experiment. In health-related research, including studies of screening tests, diagnostics, prevention, and therapeutic interventions (DeMets & Fisher, 2008), this takes the form of the randomized clinical trial (RCT). There are many instances, however, in which employing the experimental design is difficult or impossible, premature, or unethical. For this reason, there are a variety of what are called *quasi-experimental designs*, as well as descriptive and observational designs. The experimental and quasi-experimental designs, along with their strengths and drawbacks, are discussed in this chapter.

EXPERIMENTAL DESIGN

Regular use of control groups in psychosocial and educational research dates back to about 1908. This is quite a bit later than its first use in the physical and biological sciences. Boring (1954) traced the recorded use of experimental controls back to experiments by Pascal in 1648 in France:

> Wanting to test the relationship of a column of mercury to atmospheric pressure, Pascal arranged for simultaneous measurements using exactly the same procedure to be done at the foot of a mountain, which was 1800 feet above sea level, and at the top of the mountain, which was 4800 feet above sea level. At the top of the

mountain, they took measurements inside and outside of a shelter as well on one side of the mountain and the other side, to check for possible influences from other factors.

On the way down the mountain, they took an additional measurement of the column of mercury finding the measurements at the three sites to be the following:

Top of the mountain	24.71 inches
Intermediate altitude	26.65 inches
Foot of the mountain	28.04 inches

There were no differences in the measurements taken inside or outside the shelter or on one side of the mountain compared to the other side. Their findings demonstrated the difference in atmospheric pressure at different altitudes.

Research Design

A research design includes the structure of a study and the strategies for conducting that study (Kerlinger, 1973). This plan, at minimum, spells out the variables that will be studied, how they will be studied, and their anticipated relationship to each other (Spector, 1981).

Experimental designs have been developed to reduce biases of all kinds as much as possible. We will review the major sources of bias in the section on threats to internal and external validity.

The primary difference between the true experiment and quasi-experimental designs is the degree of control that the researcher has over the subjects and variables of the study. Control is much easier to achieve in the laboratory than in the field. In nursing research, the "field" includes homes, hospitals, clinics, schools, the workplace, or wherever we find people with health concerns outside a facility that is specifically designed for the conduct of research such as a sleep lab.

Before considering the basic experimental designs, we will consider some additional ideas that underlie the experimental design.

Causation

In everyday conversation, the word *cause* is used frequently, sometimes casually. In research, however, we need to be careful how we use this term. Although we often hope to identify causes of health problems, many of our studies are not designed to do this.

The basic principle behind identifying a cause is based on the time sequence of variables. Davis (1985) calls this the "great principle of causal order: after cannot cause before" (p. 11). An example may make this clearer:

- Client A visited a sick friend the day after he began coughing and had an elevated temperature.

- Client B visited a sick friend the day before he began coughing and had an elevated temperature.
- Both blamed their illnesses on their sick friends. Are they both correct?

The principle of causation says no, Client A cannot have caught that cold from his friend. Client B, however, may have caught his cold from his sick friend. The time sequence, visiting the friend after becoming ill, rules out Client A's hypothesis on the basis of the principle of causation. The cause cannot come after the effect. The indefinite answer to Client B's hypothesis, as you probably have surmised, is because there are many other possible sources of infection (family and coworkers to name just a few), not just Client B's sick friend.

Multiple causes, indirect effects, and spurious effects occur frequently. These potential effects add considerable complexity to many of our research designs. Davis (1985) uses the example of height within a family to illustrate a spurious effect. If mother and father are both tall, this influences the height of their son and daughter genetically, a direct cause and effect relationship. However, the height of the son does not affect the daughter's height or vice versa, although it may appear to because of the high correlation. This apparent but false direct effect between the heights of the son and daughter is a spurious effect.

Indirect effects are likened by Davis (1985) to ripples on a pond. A chain of events or factors may lead to the ultimate effect (outcome):

> A couple is arguing with each other on their way home from a party. Brenda has accused Bart of drinking too much and acting very foolish in front of their friends. Bart denies both accusations. Their argument is escalating as Bart drives up the ramp into heavy traffic. As he turns to Brenda to tell her that she had also been acting foolishly, the car in front of him slows to avoid hitting a tire that fell off a truck. Bart's response is delayed just enough that he slams into the car in front of him. Bart's speed, following too closely, and the errant tire were multiple causes of the accident. The argument was an indirect effect because it contributed to his speeding and following too closely. Using a cell phone and/or texting while driving could have an effect similar to the argument.

Threats to Internal and External Validity

Internal validity is concerned with minimizing the effects of extraneous or confounding factors that may interfere with interpretation of the results of the experiment. Campbell and Stanley (1963) listed eight threats to internal validity:

1. *History:* What is happening at the same time the experiment is being conducted? Seasonal effects, patient transfers to different units, staff changes, reorganization, a natural disaster, even the beginning of a new school term can affect nursing research studies. For example, a study that is testing the effect of new infection

control policies on patient mortality can be confounded by a peak in the incidence of a particularly virulent influenza that raises the death rate of very young and very old patients.

2. *Maturation:* The effect of changes that occur naturally over time. These may include growth and development, growing older, or getting tired, hungry, or bored. For example, infants enrolled in a stimulation study will also be experiencing natural development of various cognitive abilities without the added stimulation. These developmental changes may confound the effects of the stimulation intervention.
3. *Testing:* The use of the same questions on pretest and posttest may affect how well subjects do at the second testing. For example, a questionnaire on attitudes toward people who are substance abusers may increase sensitivity to their problems. Likewise, nurses given a drug calculation test at the beginning of a study may practice on their own before the posttest is administered; keeping a food diary may change people's behavior by alerting them to poor eating habits, and so forth.
4. *Instrumentation:* Differences in the way different examiners complete observation or rating scales and in the instruments being used may directly affect the quality of the data obtained. This is primarily a question of reliability.
5. *Statistical regression:* The phrase *regression to the mean* describes the likelihood that subjects chosen because they score very high or very low on a particular test are likely to move closer to the mean (average) on subsequent tests without any intervention.
6. *Selection bias:* There may be differences, often subtle ones, in the way people are selected for the experimental treatment group and the comparison or no-treatment group. For example, people who are eager to exercise are easier to recruit for an exercise study, especially for the intervention group, than are people who do not want to exercise.
7. *Experimental mortality:* Differences may occur in the loss of subjects in the treatment group versus the control group. For example, the eager exercisers are more likely to complete a 6-week exercise program than an attention control educational program.
8. *Selection–maturation interaction:* Changes that are due to the interactive effect of selection bias and maturation may be mistakenly believed to be due to the effect of the experimental treatment. For example, a study of school-age children participating in a fitness program may be confounded by maturation in physical ability over time as well as the greater enthusiasm for exercise of those who complete the fitness program.

External validity is concerned with the degree to which the results of the study can be generalized to others. In fact, some argue that the rigorous controls of a true experiment

lead to results that may occur in an ideal situation but not fare so well in the real world of everyday practice (Toulany, McQuillan, Thull-Freedman, & Margolis, 2013). Campbell and Stanley (1963) listed four threats to external validity.

1. *Testing effect:* The effects of having been pretested may be sufficient to make the groups quite different from untested people to whom the results of the study will be generalized.
2. *Selection effect:* The rigorous criteria used to select subjects may limit generalizability. For example, in many pharmacological studies the subjects cannot have any illness other than the one for which the drug is intended. Although this eliminates the confounding effects of other illnesses, it also does not represent the reality of multiple comorbidities, especially in people with multiple chronic illnesses.
3. *Experiment effect:* Being involved in a carefully designed and implemented experimental study can be a very different experience from receiving the same treatment in ordinary care settings.
4. *Multiple treatment effect:* This threat occurs when the same subjects are exposed to more than one treatment (using the subjects as their own comparison group, for example). Campbell and Stanley (1963) comment that "the effects of prior treatments are not usually erasable" (p. 6).

There is no simple formula for addressing these threats to internal and external validity of an experimental design. Each research study poses different challenges that require thoughtful, often creative solutions. As Sandelowski and colleagues (2012) have observed, we actually have to reinvent these research methods every time we use them to accommodate the real world of research practice (p. 320).

The True Experiment

Although refinements are continuously being developed, the basic experimental design has remained consistent for quite some time and is often the "gold standard" against which other designs are measured (Thompson & Panacek, 2006). There are many elaborations on these basic designs; the most common of these will be explained.

Basic Experimental Designs

The simplest experimental design tests just one treatment that is compared with a no-treatment condition. Subjects (or participants, a term preferred by those who dislike the idea of people being "subjected" to an experiment) are randomly assigned to either the treatment or control (no-treatment) group and are measured or tested both before and after the experimental treatment is implemented.

This simple but elegant design is the basis for most of the variations on the experimental design in research involving human subjects. It can be represented by a set of symbols (Campbell & Stanley, 1963, p. 13):

	Pretest	Treatment	Posttest
R	O_1	X	O_2
R	O_3		O_4

R = Randomized, that is, subjects are randomly selected and randomly assigned to the treatment group

O = Observation or testing

X = The treatment

Note that subjects are randomly selected and randomly assigned to the treatment or control group. In addition, if examiners (data collectors) are used, they should be blinded to treatment group assignment; in other words, the examiners should not know if a person is part of the experimental or control group.

Technically, if subjects are randomly selected and assigned to treatment or control groups, it should not be necessary to pretest them because the groups are, by definition, equivalent due to randomization. The symbolic representation of this posttest-only design is (Campbell & Stanley, 1963, p. 25):

	Treatment	Posttest
R	X	O_2
R		O_4

In reality, however, when relatively small numbers of subjects are involved, randomization may not produce exactly equivalent groups. One group may be a little older, for example, or have a higher average pretest weight. If known, these differences can be controlled statistically. If not known, a difference at posttest may be attributed to the experimental treatment although it really reflects the failure of randomization to produce equivalent groups.

Given the two concerns about the effect of pretesting and about the possibility that randomization may fail to produce equivalent groups, there is a more complex (and more resource-consuming) design that addresses them. This is the Solomon four-group design that uses two treatment groups, one pretested and the other not pretested, and two no-treatment control groups, one pretested and the other not pretested. Using the same symbols, the Solomon four-group design looks like this (Campbell & Stanley, 1963, p. 24):

	Pretest	Treatment	Posttest
R	O_1	X	O_2
R	O_3		O_4
R		X	O_6
R			O_8

If you want to test both immediate posttreatment outcomes and long-term outcomes (3 or 6 months posttreatment, for example), there is a nice design that controls for the effect of immediate posttesting (Campbell & Stanley, 1963, p. 32):

	Pretest	Treatment	Immediate Posttreatment	3 Months Posttreatment
R	O_1	X	O_2	
R	O_3		O_4	
R	O_5	X		O_6
R	O_7			O_8

There are many instances in which this long-term effect is equally or even more important than the immediate effect. Maintaining the weight loss of people who had been obese would be an example.

Additional treatment groups may be added as well. For example, one study (Paul et al., 2007) compared the results of using honey, dextromethorphan, and no treatment for coughs due to upper respiratory infections in children. This study used two experimental treatments, honey and the ingredient common to many over-the-counter cough medications, and a control group that received nothing for their cough. A study using two experimental treatment groups and one control group with randomization and posttesting only would be diagrammed as follows:

	Treatment	Posttest
R	X_1	O_2
R	X_2	O_4
R		O_6

Randomized Clinical Trials

A clinical trial is "an experiment in which a group of individuals is given an intervention and subsequent outcome measures are taken" (Cook & DeMets, 2008, p. 10). Similarly, the National Institutes of Health (NIH) (2014) define clinical trials as:

> A research study in which one or more human subjects are prospectively assigned to one or more interventions (which may include placebo or other control) to evaluate the effects of those interventions on health-related biomedical or behavioral outcomes.

Likewise, a Phase III clinical trial is defined by NIH (2013) as follows:

> Phase III clinical trials usually involve several hundred or more human subjects for evaluating an experimental intervention in comparison with a standard or controlled intervention or comparing two or more existing treatments. Often the aim is to provide evidence leading to a change in health policy or standard of care. The

definition includes pharmacologic, non-pharmacologic, and behavioral interventions for disease prevention, prophylaxis, diagnosis, or therapy. Community trials and other population-based intervention trials are included.

In a randomized clinical trial, individuals are randomly assigned to intervention or no-intervention groups or "arms" of the study. In a randomized, blinded clinical trial, the raters or examiners who conduct the pretesting and posttesting are completely ignorant of treatment group assignments. Even more control is achieved if the study uses a double blind design where neither the researchers, the examiners, nor the subjects know who is receiving active treatment and who is receiving an inactive placebo that appears the same as the actual treatment. Achieving double blind control is obviously much easier when testing a medication or device that delivers treatment without researcher intervention. Behavioral interventions are difficult to cloak in an effective manner. If, for example, you are comparing two types of therapy, the subject may be blinded as to whether it is the experimental therapy being received, and independent raters may also be blinded. Those who are providing the treatment, however, cannot be blinded although they can be kept ignorant of the study hypothesis.

Nurses, physicians, and other healthcare providers employ a number of practices that lack research-based evidence to support their use. The following example cited by DeMets and Fisher (2008) demonstrates the potentially powerful impact of a well-designed clinical trial on patient care practices:

> The story of the use of high-dose oxygen for premature infants has been told often but bears repeating. Given the respiratory challenges in many preemies, it seemed reasonable to treat them with high doses of oxygen. This became standard practice, but it was eventually noted that retrolental fibroplasia was also increasing, leading to blindness in many of these infants. A review of patient records indicated that it particularly affected the infants who had received the high doses of oxygen. But this evidence did not convince practitioners. In fact, in one trial where infants were randomized to high- or low-dose oxygen, the nurses increased the oxygen setting at night for the infants randomized to the low-dose group because they believed this was the better treatment. You can see why a randomized clinical trial comparing administration of high- and low-dose oxygen to premature infants was a challenge both in terms of research ethics and in its implementation. A later clinical trial in which 800 premature infants were randomized to low-dose or high-dose oxygen produced convincing evidence: 23% of the infants in the high-dose group were blinded versus 7% in the low-dose group. DeMets and Fisher (2008) note that as many as 10,000 infants were probably blinded during the time that high-dose oxygen was the accepted treatment. "A widely used but untested intervention," they concluded, "was ultimately shown to be harmful" (p. 16).

There are a number of stories like this. Here are just two more:

1. Intermittent positive pressure breathing (IPPB) used a relatively complex and expensive device to deliver bronchodilators to the lungs. When compared to the

use of a much cheaper, simpler, handheld nebulizer, the clinical effect was found to be the same.

2. When a trial of coronary artery bypass graft (CABG) was undertaken, many surgeons were hesitant to have their patients randomized to the medical (nonsurgical) treatment group. The randomized clinical trial demonstrated that CABG was not superior to medical therapy in people with less advanced disease and fewer occluded coronary blood vessels (DeMets & Fisher, 2008).

It is interesting to note the reluctance to give up an accepted practice that is evident in two of these three examples. Once a practice is accepted, it takes a large, well-designed (i.e., difficult to argue against) study to reverse an entrenched practice. The first example also illustrates the thinking and nonexperimental research that often precede use of an experimental design. This is the discovery phase in which observation (an increase in blindness in premature infants, for example), development (an easy-to-use nebulizer), or astute questioning of expensive or untested practices preceded the experimental study. It is not at all unusual for a considerable amount of preliminary work to have been done before conducting a clinical trial.

Variations of the Randomized Clinical Trial

There are several common variations of the basic randomized clinical trial. These are the parallel groups, randomized block design, stratified random sample, paired or matched comparisons, and crossover design. Each of these is explained in the following sections (Fleiss, 1999; Piantadosi, 2013).

Parallel Groups

The simplest and most commonly used design, this is the basic true experiment described earlier.

Randomized Block Design

This is a form of directed or constrained randomization of assignment to treatment group that ensures that equal numbers of subjects are assigned to each of two or more treatment groups. For example, if you use 6 subjects per block to create a final sample of 60, you need 10 blocks altogether. Within each block of 6, you then create a random sequence of equal numbers of treatment and no-treatment assignments for the subjects. Each block would have a different random sequence: treatment, no treatment, no treatment, treatment, treatment, no treatment, for example.

Stratified Random Sample

Certain characteristics of subjects can be expected to affect their response to the experimental treatment. Males may lose weight more easily than females, younger people may complete rehabilitation more quickly than older people, and individuals with severe

allergies may have a higher level of motivation to implement allergen control measures than those with mild seasonal allergies. There may be good reasons for ensuring that equal numbers of males and females, smokers and nonsmokers, or other groupings of people are assigned to each treatment group. This can be done by using a stratified random sample in which subjects are randomly selected in equal numbers from each stratum, or group, such as males and females. Randomized blocks and stratification are often combined in a single design.

Paired or Matched Comparisons

In some studies, matched pairs of individuals are randomized to treatment and no-treatment control groups. People can be matched on any of a number of characteristics: gender, age, ethnic group, income, severity of illness, comorbidities, functional impairment, and so forth.

Crossover Designs

Each subject serves as his or her own control in this design. Subjects receive each of the treatments to be tested, one at a time, often with a washout period in between to reduce the possibility of carryover of the effect of the first treatment on the second treatment, and so forth. An example of a crossover study would be to test infants' tolerance of three or four different formulas, one formula at a time. Do not confuse crossover designs with cross-sectional designs which are descriptive, not experimental in nature.

Types of Control Groups

In the design of clinical trials, there are several different ways to construct the control or comparison group. The true experiment model uses a no-treatment control group. This is a group of individuals who are enrolled in the study but who will be involved in only the pretesting, if done, and the posttesting.

When the individuals are patients or clients of a healthcare facility, medical ethics rarely allows care to be withheld. Instead, they may receive the usual or customary care, becoming part of a comparison group rather than a no-treatment control group.

Both the control and comparison groups are usually concurrent groups; that is, groups are tested or observed within the same time frame as the treatment group. This is important when a temporal factor is likely to affect outcomes.

An alternative to the concurrent control or comparison group is the historical control group (Piantadosi, 2013). The data on a historical control group are obtained from previous studies of individuals who did not receive the experimental treatment (DeMets & Fisher, 2008). An example would be the use of known mortality rates for people with certain advanced cancers. The results of a treatment that dramatically increases survival rates can be compared to these known mortality rates.

Hypothesis vs. Research Question

Randomized clinical trials are designed to be hypothesis-testing research studies. A hypothesis is a statement that specifies the population of interest, the intervention to be tested, and the outcome that is expected. The following are examples based on nursing studies:

- Middle school students who receive an individually tailored physical activity program and nurse counseling will increase their reported physical activity more than those who receive printed information on physical activity (based on Robbins, Gretebeck, Kazanis, & Pender, 2006).
- Use of a culturally appropriate Spanish language home-based educational intervention will improve knowledge, attitudes, and preventive practices of urban Latino households (based on Larson, Ferng, McLoughlin, Wang, & Morse, 2009).

The null hypothesis states that no difference will be found between the treatment and control groups, that the treatment has had no effect. Technically, this is the hypothesis that is tested statistically.

Research questions generally do not specify the expected outcomes of the study, although they should otherwise be as specific as possible. They are more appropriate for nonexperimental studies. A few examples include "What are the factors affecting the rate of influenza vaccination in low-income urban Latino families?" and "What is the average number of hours of physical exercise of middle school students in low-income urban neighborhoods?"

QUASI-EXPERIMENTAL DESIGNS

There are many instances where you cannot use a true experimental design or meet the criteria for a randomized clinical trial. The following are some examples:

- *When your treatment (intervention) cannot be simultaneous with the no-treatment data collection period but has to follow it:* For example, a sweeping change in the responsibilities of nursing assistants cannot be contained within the experimental units of an organization if nursing assistants float to other units and/or if nursing assistants ever talk with nursing assistants on other units, so preintervention data must be collected before the intervention is implemented.
- *When pretesting may have a strong effect on individuals in the no-treatment group:* For example, doing poorly on tests of endurance and balance may in itself convince participants that they need to become more active, so no-treatment controls may also improve their fitness. In some cases, omitting the pretest may be the only solution to this design problem.

- *Randomization may fail to produce groups that are precisely equivalent in terms of characteristics that may be important to the outcome of the study:* For example, Dunbar and colleagues (2014) randomized individuals who had both heart failure and diabetes into an integrated self-care program group and a usual care group, but one group ended up with a higher proportion of African Americans, and those in the other group were slightly older. Randomization also may fail when participants refuse assignment to a less desirable intervention. For example, individuals with mild depressive symptomology were invited to participate in a study of the effect of art therapy compared with passive listening to music. The individuals who agreed to participate were eager to be involved in "something that might do me some good" and were so disappointed if they were assigned to the music group that several dropped out of the study, leaving the researcher with nonequivalent groups.

Basic Quasi-experimental Designs

There is an almost endless variety of quasi-experimental designs. Some of these variations are very weak; others are relatively strong if used and interpreted appropriately. We will consider four progressively weaker designs that are commonly used. All are based on descriptions by Cook and Campbell (1979) and are represented symbolically with the same notation used for the experimental designs. If you keep in mind the threats to internal and external validity described earlier in this chapter, you will be able to recognize the weaknesses of these designs.

The first quasi-experimental design employs all the elements of the experimental design except randomization. The control group design with a pretest and posttest includes an intervention (X), pretesting and posttesting (O), and a no-treatment control group (Cook & Campbell, 1979, p. 104).

Pretest	Treatment	Posttest
O_1	X	O_2
O_3		O_4

The primary weakness in this very common design is that the two (or more if multiple interventions are tested) groups may not be equivalent. Even more difficult to address is the fact that they may be different in ways that are not known to the researcher and, therefore, will not be detected or measured in any way. Known differences in the two groups such as differences in mean age or income may be controlled statistically; however, many unknown sources of bias may still be present. If, for example, the first 50 volunteers are placed in the experimental group and the second 50 volunteers are placed in the control group, then you may have more eager participants in the experimental group, or more people who get up early, or people who are more assertive, have more initiative, and so forth. These characteristics may influence willingness to try a new intervention.

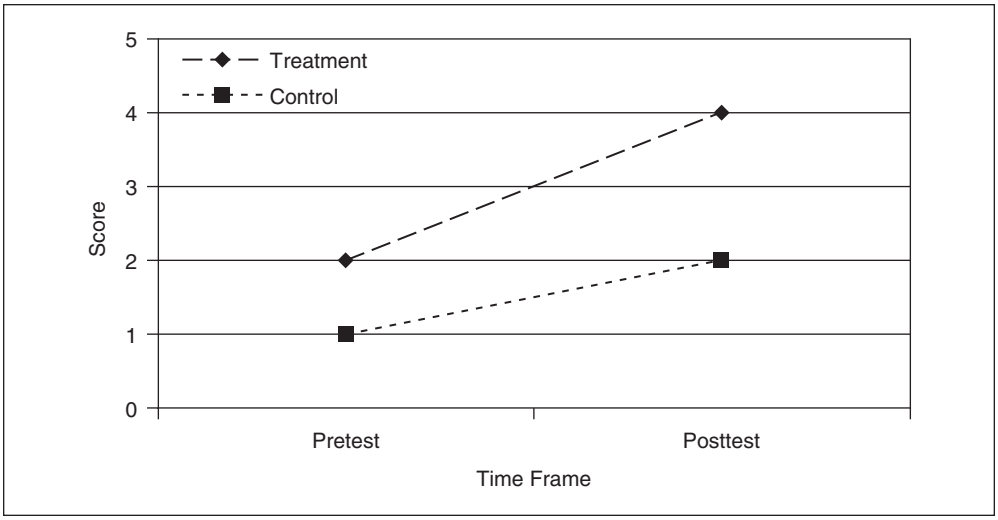

Figure 5-1 First pattern: nonrandomized control group with pretest and posttest design.

Source: Data from Ruth Tappen.

Cook and Campbell (1979) devote many pages to the variety of outcomes that may be encountered when using this quasi-experimental design. We will consider just two to touch on the types of concerns that may arise given the direction of the outcome.

The first pattern is one that looks good initially, but there are reasons to be skeptical about the contribution of the experimental treatment to the better outcome for the treatment group. As you can see in **Figure 5-1**, the control group began at a lower level and progressed only 1 unit by posttest whereas the treatment group began at a higher level and progressed 2 units by posttest. (The units are scores on the primary outcome measure, which could be an improvement in function measured by degree of assistance needed or an increase in lung capacity or fewer depressive symptoms, and so forth.) It appears that the experimental treatment group progressed at twice the rate of the control group. Does this mean that the experimental treatment was effective? Not necessarily. Cook and Campbell (1979) point out that the individuals given the experimental treatment (not randomly assigned in this quasi-experimental design) may well have been improving at a faster rate before the study began. This is an example of the selection—maturation interaction threat to internal validity.

Another pattern is more promising. In this pattern (**Figure 5-2**) there is a crossover of the experimental treatment group and control group means from pretest to posttest: the experimental treatment group began at a lower level but ended at a higher level than the control group, which did not change at all. In other words, the experimental treatment group has "overtaken" the control group (Cook & Campbell, 1979, p. 111).

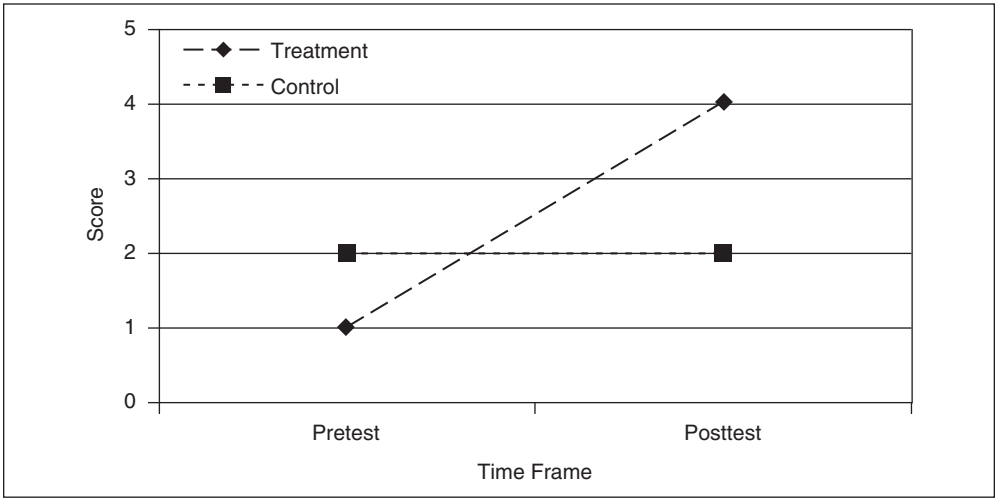

Figure 5-2 Second pattern: nonrandomized control group with pretest and posttest.

Source: Data from Ruth Tappen.

The tentative conclusion that this is due to the intervention itself is more plausible than in most other patterns for several reasons:

- There is no apparent ceiling (the highest either group could achieve) or floor (the lowest score possible) effect evident here.
- Selection–maturation does not explain the difference because the control group means did not change; few maturation patterns (child growth and development patterns, for example) match the pattern in Figure 5-2.
- It also does not appear that a regression to the mean has occurred.

Although this second pattern is more supportive of the hypothesis that the experimental treatment had a substantive effect on the outcomes, this conclusion still remains far weaker without randomization.

The second quasi-experimental design also has an experimental treatment and a control group (that is, two nonequivalent groups) but lacks randomization or pretest information on either group (Cook & Campbell, 1979, p. 98):

Treatment	Posttest
X	O_2
	O_4

The information about the two groups' initial status is missing and must be assumed. Any assumption on the part of the researcher that the groups are equivalent is very

tentative and may be incorrect. Patients from the same nursing care unit, clients of the same clinic, or infants from the same nursery may all differ on critical characteristics if not randomly selected and randomly assigned to treatment or control group. Without a pretest, there is virtually no information about their equivalence on variables of interest, especially the outcome variables.

Why, then, you might ask, is this design ever used? Often, a researcher encounters a situation calling for study where it is too late or not possible to collect pretest data. An example would be a comparison of the degree of posttraumatic stress disorder in hurricane victims who received crisis intervention compared with those who did not. It is not possible in this case to test the prehurricane status of the people affected, although it can be estimated through postdisaster interviewing.

The third quasi-experimental design is the one-group pretest–posttest design in which there is only an experimental treatment group and no comparison or control at all (Cook & Campbell, 1979, p. 99).

Pretest	Treatment	Posttest
O_1	X	O_2

You can see that there is information about how much change occurred in the experimental group, but arguments for ascribing it to the experimental treatment have a very weak basis. In fact, the change, if there is one, could be due to an entirely different factor. For example, a school nurse may have been concerned about low immunization levels in incoming kindergartners. To remedy this, the school nurse sent a letter to all their parents about 6 weeks before the school term began in the fall. The rate of completed immunizations rose dramatically from 88% the previous year to 98% the new fall term. The school nurse was very pleased, of course, but was it possible to credit the letters sent home? In fact, the health department had received a small grant to place advertisements on local television and radio at the same time that the letters were sent out. Neither the school nurse nor the health department could differentiate between these two possible contributors to the improved immunization rates. In addition, neither was aware that the only pediatrician in town had, with the pediatric nurse practitioner in the practice, begun more actively urging parents to complete their children's immunizations as well, another potential contributing factor.

Parenthetically, if there were several elementary schools in the same locality and only one school nurse sent letters, the response (immunization rate) could be compared across two schools. Of course, the schools themselves may not be comparable, and the parents and children in each school may differ on such important factors as income and confidence in the safety and efficacy of immunizations. For example, there are parents who fear that the mercury in some immunizations may be a cause of autism spectrum disorders and resist having their children immunized for this reason.

This one-group design is often used when resources are limited or it is difficult or impossible to recruit a control group. At best, a positive outcome only suggests an effect for the experimental treatment.

The fourth and final quasi-experimental design to be discussed is the one-group posttest-only design. This design lacks randomization, comparison with an equivalent group, even information on the experimental treatment group before treatment is given. The symbolic representation of this design is as follows (Cook & Campbell, 1979, p. 96):

Treatment	Posttest
X	O

Despite its limitations, there are times when this extremely simple design is appropriate. For example, it may be used to identify the source of a contaminant in food that has sickened a large number of people. In this case, however, prior information exists that can be used at "posttest": knowing the manufacturer of the contaminated food, the manufacturing process used, and the ingredients included in the food allows one to trace the path of the contaminant to this recipient.

In most instances, however, this last and simplest design is inadequate if you want to study cause and effect related to a nursing intervention and patient outcomes.

CONCLUSION

Careful planning is essential to all research. The best time to do the work necessary to ensure that you will have data of the type and quality needed to achieve your study objectives is before you begin the study (Piantadosi, 2013). Once the people are selected, the treatment provided, and the data collected, it is very difficult to fill in the gaps, reduce preventable biases, or correct other mistakes that were made in planning the study. Experienced researchers appreciate the importance of careful planning before beginning a study. If you are planning a quantitative study, this includes consultation with a clinically oriented biostatistician to be certain that the data obtained will be sufficient to address the questions posed and conduct the desired analyses.

REFERENCES

Boring, E. G. (1954). The nature and history of experimental control. *American Journal of Psychology*, *67*(4), 573–589.

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Boston, MA: Houghton Mifflin.

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

Cook, T. D., & DeMets, D. (2008). *Introduction to statistical methods for clinical trials*. Boca Raton, FL: Chapman & Hall/CRC.

Davis, J. A. (1985). *The logic of causal order*. Newbury Park, CA: Sage.

DeMets, D., & Fisher, M. R. (2008). Introduction to clinical trials. In T. D. Cook & D. L. DeMets (Eds.), *Introduction to statistical methods for clinical trials* (pp. 1–28). Boca Raton, FL: Chapman & Hall/CRC.

Dunbar, S. B., Butts, B., Reilly, C. M., Gary, R. A., Higgins, M. K., Ferranti, E. P., … & Butler, J. (2014). A pilot test of an integrated self-care intervention for persons with heart failure and concomitant diabetes. *Nursing Outlook*, *62*(2), 97–111.

Fleiss, J. L. (1999). *The design and analysis of clinical experiments*. New York, NY: John Wiley & Sons.

Kerlinger, F. N. (1973). *Foundations of behavioral research.* New York, NY: Holt, Rinehart & Winston.

Larson, E. L., Ferng, Y. H., McLoughlin, J. W., Wang, S., & Morse, S. S. (2009). Effect of intensive education on knowledge, attitudes and practices regarding upper respiratory infections among urban Latinos. *Nursing Research*, *58*(3), 150–165.

National Institutes of Health (NIH). (2014). Notice of Revised NIH Definition of "Clinical Trial" (NIH Notice No. NOT-OD-15-015). Retrieved from http://grants.nih.gov/grants/guide/notice-files/NOT-OD-15-015.html

National Institutes of Health (NIH). (2013). Grants and Funding Glossary and Acronym List: Phase III Clinical Trials. Retrieved from http://grants.nih.gov/grants/glossary.htm#PhaseIIIClinicalTrial

Paul, I. M., Beiler, J., McMonagle, A., Shaffer, M. L., Duda, L., & Berlin, C. M. (2007). Effect of honey, dextromethorphan and no treatment on nocturnal cough and sleep quality for coughing children and their parents. *Archives of Pediatric and Adolescent Medicine*, *161*(12), 1140–1146.

Piantadosi, S. (2013). *Clinical trials: A methodologic perspective.* New York, NY: John Wiley & Sons.
✧ Probably the most readable of the many books on RCTs.

Robbins, L. B., Gretebeck, K. A., Kazanis, A. S., & Pender, N. J. (2006). Girls on the Move program to increase physical activity participation. *Nursing Research*, *55*(3), 206–216.

Sandelowski, M., Voils, C. I., Leeman, J., & Crandell, J. L. (2012). Mapping the mixed methods–mixed research synthesis terrain. *Journal of mixed methods research*, *6*(4), 317–331.

Spector, P. E. (1981). *Research designs*. Beverly Hills, CA: Sage.

Thompson, C. B., & Panacek, E. A. (2006). Research study designs: Experimental and quasi-experimental. *Air Medical Journal*, *25*(6), 242–246.

Toulany, A., McQuillan, R., Thull-Freedman, J. D., & Margolis, P. A. (2013). Quasi-experimental designs for quality improvement research. *Implementation Science*, *8*(Suppl 1), S3.