Characteristics of Experimental Research

There are several key ideas or characteristics involving experimental research that individuals must be aware of to truly grasp how experimental research differs from other research methods. These key ideas, as presented by Creswell (2008) include:

- Random assignment
- Control of extraneous variables
- Manipulation of treatment
- Measurement of outcomes
- Comparison of participant groups
- Possible threats to validity

**Random Assignment**

As discussed earlier, experiments involving random assignment are considered to “provide the most rigorous evidence” that a specific treatment produces a certain outcome. Random assignment is “the process of assigning individuals at random to groups or to different groups in an experiment” (Creswell, 2008, p. 300). Random assignment of participants to groups is the quality that distinguishes “true” experiments from the less rigorous quasi-experiments (Creswell, 2008). The purpose of random assignment is to insure that groups receiving different treatments are as reasonably equal or similar in any way that could possibly impact the outcome (or dependent variable) (Slavin, 2007). Thus, random assignment helps increase the likelihood that any personal characteristic that could bias a study’s outcome are evenly distributed among groups of participants. The process of evenly distributing potential bias is referred to as equating the groups. Such characteristics that could bias the outcome measured by a study are referred to as extraneous variables. Since personal characteristics that could bias a study’s outcome can never be completely eliminated, the process of randomly assigning participants to groups also serves to randomly distribute any potential bias. It is important to note that random assignment and random selection do not refer to the same process. Rather, random selection refers to “the process of selecting individuals so each individual has an equal chance of being selected” (Creswell, 2008, p. 645). Random selection, which serves a different purpose in research, often is not utilized during experimental research as it is not always a logical possibility (Creswell, 2008).

**Control of Extraneous Variables**

Extraneous variables must be controlled for in experimental studies. Extraneous variables are participant or environmental variables that “confuse relationships among the variables being studied” (Slavin, 2007, p. 384). Participant variables that act as extraneous variables include both intervening variables and organismic variables. Intervening variables include variables that are not directly observable (e.g., anxiety, boredom) and must be controlled for by researchers. While organismic variables cannot be altered by researchers, as they include variables such as participants’ gender, they too can be controlled. Experimenters can best control for extraneous variables by using randomization. Randomization requires that researchers randomly select study participants as well as randomly assign participants to treatment groups. Researchers can employ a variety of practices to ensure that extraneous variables are controlled for through random selection and random assignment such as by using a table of random numbers. By using either a
table of random numbers or simply flipping a coin, researchers can be certain that participants were selected and assigned to groups by pure chance and, thus differences amongst participants should be equally distributed between treatment groups. Randomization is an integral element of experimental research as it allows researchers to create treatment groups that are relatively equal, thus, any differences noted after the researcher has implemented the experimental treatment can be attributed to the treatment (Gay et al., 2006).

Researchers should also be cognizant that the effectiveness of randomization can be greatly impacted by the amount of study participants. For instance, a researcher who randomly selects and randomly assigns six participants to two groups (a control and treatment group) can be less certain that extraneous variables have been effectively equalized than a researcher who employs random selection and random assignment with a study of 150 participants. In situations where researchers are prevented from randomly selecting participants, researchers should at least randomly assign participants to groups. When researchers are not able to use either random selection or random assignment, researchers should at least randomly select which group will receive the treatment and which group will act as the control group.

Additional methods exist that researchers can utilize to control for extraneous variables. One such method is to hold certain variables constant throughout an experiment for both the control and treatment group. This practice is often utilized with certain environmental variables such as the learning materials used, participants’ prior exposure to certain variables, and the place and time the experiment occurs (Gay et al., 2006). For instance, if a high school principal is interested in studying the effectiveness of a specific science program, the principal might control for extraneous variables by holding certain variables constant for both the control and treatment group by selecting two classes to act as a control and treatment group that meet at the same time each morning. This will allow the principal to be certain that any differences noted in learning between the treatment group and control group are not due to the time of the morning during which the instruction was taught and cannot be due to one class of students being less alert than the other class. The principal might also decide to control for extraneous variables by holding another variable constant: the learning materials used by the teachers. Thus, the principal provides the same learning materials for each class of students to ensure that the quality of learning materials is equal for both experimental groups.

Other methods that researchers can employ to control for extraneous variables include: matching groups, comparing subgroups or homogenous groups, using a single group as both the control and treatment group, and using certain statistical procedures to equate groups on particular variables (e.g., analysis of covariance is one such statistical procedure). Matching typically requires that the researcher identify two participants that are similar on certain characteristics and then assign one participant to the treatment group and the other participant to the control group. For instance, when using matching to ensure that groups are equal in terms of gender, a researcher might select a pair of two males and then assign one male to the treatment group and the other male to the control group. In this way, the researcher could be fairly certain that differences between participants are equally distributed between both the control and treatment group. Researchers then should exclude any participants from the study who cannot be paired as being similar to another participant according to a specific characteristic. This practice, of course, could require that the researcher eliminate many participants from a study especially if
the researcher attempts to match participants on more than one variable. Another matching technique typically employed by researchers requires that participants be ranked according to a particular characteristic. Pairs are then decided according to the two participants who ranked the highest with the next two participants being paired together and so forth. The type of matching does not require researchers to exclude participants who might not be easily matched to another participant on a certain characteristic, unfortunately, this advantage also proves to be the major disadvantage of this type of matching as some participants are not as precisely matched as they may have been with the previously mentioned matching technique. The use of matching has decreased recently due to statistical procedures such as analysis of covariance.

Another way to control for extraneous variables is to compare subgroups or homogenous groups. This can be accomplished by selecting participants based on their similarity with regard to a certain variable. For instance, researchers might decide to implement reading comprehension strategies with students who are already fluent readers. Thus, the researcher might select fourth grade participants who read between 94-145 words per minute. Several disadvantages are associated with this procedure. First, it decreases the amount of participants available to researchers. Second, it decreases the generalizability of the study’s findings to fourth graders who read between 94-145 words per minute. Another approach that researchers might consider more satisfactory is stratified sampling. As described by Gay et al. (2006), stratified sampling requires researchers to “form different subgroups representing all levels of the control variable” (p. 249). For instance, rather than simply including students who read 94-145 words per a minute in the study, researchers would create subgroups such as: excelling readers (120 + wpm), high average readers (94-119 wpm), low average readers (68-93 wpm), and struggling readers (67 wpm and lower). Half of each subgroup would then be equally assigned to the control group and treatment group to ensure that extraneous variables are equally distributed among groups (Gay et al., 2006).

Another way for researchers to control for extraneous variables is by exposing a single group of participants to both treatments so that the group acts as its own control group. For instance, if a teacher were interested in whether one type of spelling instruction were more effective than another, the teacher might teach his class with the one type of spelling instruction first, collect the necessary data, and then instruct the class with the second type of spelling instruction. There are several obvious disadvantages to controlling for extraneous variables in this way. First, the technique is often not feasible due to real world constraints. Secondly, researchers cannot be certain of the impact of the first treatment on participants’ reaction to the second treatment. For instance, if a teacher were to study the effects of two classroom management programs, to implement a second class-wide approach after the first program has been implemented not only would not be very feasible, but also might impact the students’ reactions to the second program in ways that cannot be accounted for (Gay et al., 2006).

**Manipulation of Treatment**

True experimental research differs from all other methods of research in that experimental research requires the manipulation of at least one independent variable. This basically means that researchers decide which variable will serve as the independent variable to be manipulated and which group of participants will receive this treatment (Gay et al., 2006). For example, a teacher might decide to manipulate the reading strategies he uses during guided reading instruction.
After randomly assigning the students in his class to two groups, the teacher decides to teach the first group of students Think-Aloud strategies and the second group how to effectively use graphic organizers. The teacher then compares the reading progress both groups have achieved at the end of the experiment.

Types of Experimental Research

Pre-Experimental Designs:
One-Shot Case Study
The one-shot case study, as with other pre-experiments, is considered a pre-experiment (or a weak experiment) because the design lacks essential components of a true experiment: random assignment of participants to groups and manipulation (Suter, 2006). One-shot case studies occur when a researcher selects a single group of participants to act as the treatment group, implements a treatment, and then administers a posttest. This research design is problematic as no threats to validity, such as maturation, history, and mortality, are controlled for. As a pretest measure is absent in this type of research design, it is impossible to tell where participants were performing prior to the implementation of the treatment. Gay et al. (2006) recommends that, “if you have a choice between using this design and not doing a study – select another study” (p.251). This design is diagrammed as:

Group 1 → Treatment Implemented→ Posttest Administered

One-Group Pretest-Posttest Design
The one-group pretest-posttest design is considered a pre-experimental design because it includes only one group of participants. Consider the following example of the one-group pretest-posttest design: Mr. South has asked his principal for permission to utilize a new mathematics textbook with which to teach his class fractions. At the beginning of the school year, the principal randomly assigns students to Mr. South’s class from all of the fifth graders enrolled at the school. Mr. South administers an assessment prior to beginning instruction into fractions to determine his students’ knowledge. Mr. South then teaches his students fraction concepts with the new mathematics textbook. After instruction, Mr. South again administers an assessment to determine students’ knowledge of fractions post instruction with the new textbook. This design can be diagrammed as:

Students Assigned to Group → Pretest Administered → Treatment Implemented → Posttest Administered

There are several disadvantages to this type of experimental design. First, the absence of a control group does not allow researchers to make comparisons between groups. Another problematic component of the one-group pretest-posttest design is the possible impact of the pretest on students’ posttest scores. Finally, researchers cannot be certain that observed changes in posttest scores cannot be contributed to extraneous variables. Perhaps students’ increased posttest scores were impacted more by Mr. South’s excitement towards the new math text and not necessarily the intended treatment. Such factors contribute to the overall consensus of the one-group pretest-posttest design as not being a particular strong experimental design (Mertler & Charles, 2008).
Static-Group Comparison Design

Another weak pre-experiment design is the static-group comparison design. While you may not be familiar with its name, you likely are familiar with research that has employed the static-group comparison design. There are two elements of true experimental designs that are absent in studies that utilize the static-group comparison design. First, this design does not include the administration of pretest to ensure that participants in the control and treatment groups are similar on specific variables. Secondly, and perhaps most problematic, this design does not require that participants are randomly assigned to either the control or treatment group. Rather, intact groups of participants are selected to act as the control and treatment group without any certainties that the groups are similar prior to the implementation of treatment on the outcome variable. This is what the name of the design refers to with ‘static’; specifically, that groups are already intact or inactive prior to the experiment (Suter, 2006).

Consider the following example of a static-group comparison design experiment: Ms. March, an elementary principal, decides to implement a new writing program in all four sections of fourth grade for the upcoming school year. At the end of the year, Ms. March and the fourth grade teachers gather to analyze the students’ writing gains made during the course of the year when the new writing program was implemented. In order to compare their students writing outcomes with other fourth graders they look at the writing scores of fourth graders at a nearby elementary school that did not implement the new writing program. In order to control for extraneous variables that could impact the findings of Ms. March’s experiment, she should consider a technique called matching. Ms. March could use matching by selecting fourth graders from a school as alike to the school that she serves on as many important variables as possible such as: ability at the beginning of the year, socioeconomic status, and motivation (Suter, 2006). This design can be diagrammed as:

```
Intact Groups → Group 1 → No Treatment Implemented → Posttest Administered
              → Group 2 → Treatment Implemented → Posttest Administered
```

True Experimental Designs:

Two Group Posttest-Only Design

As indicated by the name of this design, the two group posttest-only design includes two groups: a control and an experimental group. Participants are randomly assigned to either the control or experimental group while the treatment (this is what has been identified as the independent variable) is only given to the treatment group. Both the control and experimental groups are then tested (the groups are tested on the dependent variable, or outcome variable). Consider the following example: Mr. South decides to continue his study for a second year of the educational outcomes of teaching with fractions with the new mathematics textbook. Again, the principal agrees at the beginning of the school year to randomly assign students to Mr. South’s fifth grade from all of the fifth graders enrolled. However, this year Mr. South teaches two sections of mathematics to fifth graders. In the first section (who act as the control group), Mr. South does not use the new mathematics textbook to teach fractions and instead uses the old math text. In the other section (who act as the treatment group), Mr. South teaches fractions to the students using the new math textbook. Mr. South then administers an assessment to both sections of fifth grade math that he taught. This design can be diagrammed as:
One major disadvantage of this design is the absence of a pretest measure. Without a pretest measure researchers cannot be certain that participants in the control and experimental groups were equivalent prior to the implementation of the treatment for the treatment group. Consider how posttest scores would be impacted in Mr. South’s study if many of the students who were in the control group had received tutoring during summer break in fractions? Or if many of the students who were in the treatment group had received tutoring during summer break in fractions while the vast majority of students in the control group had not? Either way Mr. South’s findings would not likely accurately depict the effect of teaching fractions with a new mathematics textbook (Mertler & Charles, 2008).

**Two Group Pretest-Treatment-Posttest Design**
The two group pretest-treatment-posttest design is considered a strong experimental design for several reasons. First, the design includes the usage of both a control and treatment group. Second, the design also requires that a pretest and posttest be administered to both the control and treatment group. This obviously serves to ensure that participants in both groups are as equivalent as possible on certain variables prior to the implementation of the treatment for the treatment group (Mertler and Charles, 2008). Consider the following example of a two group pretest-treatment-posttest design: Mr. South decides to continue his research into the effects of teaching fractions with the new mathematics textbook for a third year. Again, the principal agrees to randomly assign students to the two sections of mathematics that Mr. South teaches from among all of the fifth graders enrolled in the school. Mr. South then administers a pretest to both sections of fifth grade math (both the control and treatment group) to ensure that the students begin the experiment with roughly the equivalent skills in fraction concepts. Mr. South then teaches one section of fifth graders math with the old textbook (this section acts as the control group) while Mr. South teaches the other section of fifth graders math with the new textbooks (this section acts as the treatment group). Mr. South then administers a posttest to both the control and treatment groups to determine the learning outcomes of the experiment. This design can be diagrammed as:

Group 1 → Pretest Administered → No Treatment Implemented → Posttest Administered
Group 2 → Pretest Administered → Treatment Implemented → Posttest Administered

**Matched Control Group Design**
The matched control group design requires researchers to match participants prior to assigning them to either the control or treatment group. In the matched control group design, after participants have been randomly selected from a population they are rank ordered according to a specific variable closely related to the posttest measure. Participants are then paired according to the rankings. For instance, participants who ranked 1 and 2 would be paired as well as participants who ranked 3 and 4 on a specific variable and so on. Participants from each pair are then randomly assigned to either the control or treatment group. For instance, the participant who ranked as 1 might be randomly assigned to the control group while the participant who ranked as...
2 would be assigned to the treatment group and so forth through the rankings. This procedure ensures that, through matching participants with another participant who scored fairly similarly according to a specific variable related to the posttest measure, the participants in the control and treatment group are as similar as possible prior to the implementation of the treatment variable. The experimental group would then receive the treatment. The control group would not receive the treatment. Both groups would be administered post tests.

Consider the following example of a matched control group design: Ms. Simpson wants to determine the effectiveness of a new reading fluency program that uses a variety of strategies such as poetry to increase students’ fluency. Ms. Simpson begins by administering oral reading probes to all of her third grade students. Ms. Simpson then ranks students according to the results of the oral reading probes. She then matches the student who ranked 1 with the student who ranked 2 and so on. Ms. Simpson then randomly selects an individual from each pair to be placed in the control group and places the other member of the pair in the treatment group. Ms. Simpson then begins instructing the treatment group with the new reading fluency program each day while the control group is receiving keyboarding instruction in the school’s technology laboratory. Ms. Simpson uses the old reading program to teach the control group fluency while the treatment group is out of the classroom in the technology laboratory. At the conclusion of eight weeks Ms. Simpson administers a posttest to both the control and treatment group to determine whether the new reading program substantially impacted the treatment groups’ fluency scores. This design could be diagrammed as:

Matching of Students → Group 1 → No Treatment Implemented → Posttest Administered

→ Group 2 → Treatment Implemented → Posttest Administered

Factorial Design
Factorial designs can be best described with an example. Thus, consider the following scenario: Mrs. Scott is an elementary principal who is interested in implementing a new reading program next year. Mrs. Scott is curious under what conditions the program is most successful. Since Mrs. Scott wants to compare two different conditions or treatments, Mrs. Scott has decided to use a factorial design while implementing the reading program with four classrooms this year. Mrs. Scott is interested in studying the impact on comprehension of the amount of instructional time dedicated to reading (either 1 hour or 1 ½ hours) and the instructional setting (either small group or whole class). The factorial design that Mrs. Scott plans to implement for her research would be symbolized as:
The variables that Mrs. Scott has chosen to study, instructional time and instructional setting, are considered *continuous variables* as they could have many values (e.g., age, attitude). Dichotomous variables are categorical variables that only have two levels (e.g. high and low, gender) (Slavin, 2007). As illustrated with the example of the research Mrs. Scott intends to implement, factorial designs are characterized as allowing researchers to investigate two (or more) variables both individually as well as how they interact with one another. Mrs. Scott’s research proposal is an example of a 2x2 factorial design (this is the simplest factorial design) as it includes two factors (instructional time and instructional setting) and variables may be manipulated either individually or together. Since factorial designs can be a variety of sizes (e.g., 2x3, 3x3), in order to determine the amount of cells in a given factorial design, an individual needs to only multiple the number of each factors together. For instance, if not provided a picture of Mrs. Scott’s factorial design an individual would be able to determine that a 2x2 factorial design has a total of 4 cells (Gay et al., 2006).

Factorial designs are often employed in order to research the interaction effects between two variables. Interaction effects can be most simply defined as occurring “when the total effect is different than the sum of its parts” (Suter, 2006, p. 273). Often the easiest way to understand interaction effects is with an example. For instance, consider the relationship of diet and exercise. While dieting might allow an individual to lose four pounds or exercise without dieting might decrease a person’s body weight by three pounds, the combination (or interaction effect) of diet and exercise will result in the greatest weight loss, seven pounds or more. In this case, the sum is greater than the parts; however, an interaction effect can also occur when the sum is less than its parts (Suter, 2006).

While each 2x2 factorial design can only have one interaction effect (factorial designs including a greater number of variables can have more interaction effects), a 2x2 factorial design can have up to two main effects. Whereas, interaction effects consider the relationship between two variables, main effects consider “differences between groups on one grouping” (Aron, Aron, Coups, 2008, p. 458). For instance, as with all other 2x2 factorial designs, in the earlier example of Mrs. Scott’s study, there are two possible main effects. Look at the following table. If these are the results of Mrs. Scott’s study, was there any main effects? Was there an interaction effect?

<table>
<thead>
<tr>
<th>Whole Class</th>
<th>Group 2 (reading comprehension)</th>
<th>Group 4 (reading comprehension)</th>
</tr>
</thead>
</table>

The variables that Mrs. Scott has chosen to study, instructional time and instructional setting, are considered *continuous variables* as they could have many values (e.g., age, attitude). Dichotomous variables are categorical variables that only have two levels (e.g. high and low, gender) (Slavin, 2007). As illustrated with the example of the research Mrs. Scott intends to implement, factorial designs are characterized as allowing researchers to investigate two (or more) variables both individually as well as how they interact with one another. Mrs. Scott’s research proposal is an example of a 2x2 factorial design (this is the simplest factorial design) as it includes two factors (instructional time and instructional setting) and variables may be manipulated either individually or together. Since factorial designs can be a variety of sizes (e.g., 2x3, 3x3), in order to determine the amount of cells in a given factorial design, an individual needs to only multiple the number of each factors together. For instance, if not provided a picture of Mrs. Scott’s factorial design an individual would be able to determine that a 2x2 factorial design has a total of 4 cells (Gay et al., 2006).

Factorial designs are often employed in order to research the interaction effects between two variables. Interaction effects can be most simply defined as occurring “when the total effect is different than the sum of its parts” (Suter, 2006, p. 273). Often the easiest way to understand interaction effects is with an example. For instance, consider the relationship of diet and exercise. While dieting might allow an individual to lose four pounds or exercise without dieting might decrease a person’s body weight by three pounds, the combination (or interaction effect) of diet and exercise will result in the greatest weight loss, seven pounds or more. In this case, the sum is greater than the parts; however, an interaction effect can also occur when the sum is less than its parts (Suter, 2006).

While each 2x2 factorial design can only have one interaction effect (factorial designs including a greater number of variables can have more interaction effects), a 2x2 factorial design can have up to two main effects. Whereas, interaction effects consider the relationship between two variables, main effects consider “differences between groups on one grouping” (Aron, Aron, Coups, 2008, p. 458). For instance, as with all other 2x2 factorial designs, in the earlier example of Mrs. Scott’s study, there are two possible main effects. Look at the following table. If these are the results of Mrs. Scott’s study, was there any main effects? Was there an interaction effect?
There is an interaction effect as how many comprehension questions students answered correctly after a certain amount of instructional time depended on the type of setting in which students received instruction. There also were two main effects as students answered averaged a greater number of comprehension questions answered correctly after an hour and half of instruction compared to an hour of instruction. Likewise, students also averaged a greater number of comprehension questions answered correctly when they’d received small group instruction in comparison to whole class reading instruction. It is often easier to determine whether interaction effects or main effects occurred in a study by looking at a graph of the data. Look at the following graph depicting the results of Mrs. Scott’s study to determine whether you can locate the interaction effects and main effects with greater ease.

As you can see, when factorial designs are graphed one variable typically is on the horizontal axis while the other is represented with different colors or designs of the bars. The vertical axis is typically reserved for the results of the dependent (outcome) variable.
Multiple steps are involved in completing experimental research. In experimental research procedures are rarely deviated from by researchers once they have been selected. Lodico et al. (2006) lists the following steps as integral to experimental research:

**Step One: Select a Research Topic**
As with other types of research, researchers prior to conducting experimental research must select a topic to research further. Topics are typically selected based on personal interest, a review of recent literature, and the experiences of the researcher.

**Step Two: Review the Literature and Specify a Research Question**
Literature reviews are typically exhaustive in scope and require researchers to determine the methodology previously used to study the topic of interest, variables that have been studied in association to the topic, and the findings of previous studies. Research questions are typically generated based on researchers’ exploration of previously published literature of the topic. For instance, while some researchers select a research question based on a paucity of research into one area of a topic, other researchers select a research question that is similar to those previous tested in order to replicate study findings.

**Step Three: Develop a Research Hypothesis**
Lodico et al. (2006) defines a research hypothesis as “an educated guess that states the expected outcome of the study” (p. 181). Researchers are ‘educated’ about a specific topic through the literature review that they conduct. With the information gathered from the literature review, researchers must state a hypothesis that indicates what they believe the causal relationship between variables will look like. The following is an example of a research hypothesis:

Students who participate in a peer-tutoring spelling intervention will spell more words correctly on weekly spelling tests in comparison to students who practice weekly spelling words by writing each spelling word correctly ten times.

Variables in a research hypothesis must be operationally defined. This means that researchers must clearly define how each variable will be manipulated and/or measured. For instance, they simply state in a research hypothesis that students’ who participate in a peer-tutoring intervention will have higher spelling achievement is not to operationally define what ‘spelling achievement’ looks like. Instead, researchers should define what spelling test will be used to compare students’ spelling achievement and, in some cases, what score students are required to earn on a specific test.

Three types of research hypotheses exist in experimental research. The above example of a research hypothesis states that students who participate in a peer-tutoring spelling intervention will spell more words correctly than students who practice spelling words via another method is an example of a directional hypothesis. Directional hypotheses differ from other types of research hypotheses in that researchers specify the expected outcome or direction of the relationship between variables in directional hypotheses. Directional hypotheses are typically used when sufficient evidence exists from previously completed research that allows researchers to predict the direction of group differences. For instance, another example of a directional hypothesis would be the following:
Students who participate in an after-school tutoring program will perform significantly better in reading as measured by the Iowa Assessments than students who do not participate in the after-school mentoring program.

Another type of hypothesis question is the nondirectional hypothesis. With nondirectional hypotheses, researchers typically state that some difference will occur between two groups but does not specify the direction of the difference between groups. The following hypothesis is an example of a nondirectional hypothesis:

**There will be a significant difference** between students who participated in the after-school tutoring program and students who did not participate in the after-school tutoring program in reading achievement as measured by the Iowa Assessments.

The final type of research hypothesis, the null hypothesis, is an implicit element of all experimental research. As defined by Lodico et al. (2006), “the null hypothesis states that no significant difference between the variables is expected after the treatment is applied” (p. 182). The purpose of most experimental research is to disprove the null hypothesis while instead providing evidence to support the research hypothesis. Null hypothesis are the most common hypothesis used in educational research. The following is an example of a null hypothesis:

**There will be no difference** in reading achievement as measured by the Iowa Assessments between students who participated in the after-school tutoring program and students who did not participate in the after-school tutoring program.

**Step Four: Select and Assign Participants to Groups**

Experimental research typically includes researchers randomly selecting and then randomly assigning participants to two groups: a control group and a treatment group. As earlier discussed, random selection and random assignment is especially important in experimental research as it allows researchers to generalize findings of the study to the population of interest. What exactly does it mean to generalize experimental findings back to a population? Consider this scenario. An administrator is interested in knowing whether an after-school tutoring program significantly impacts students’ reading achievement in the school district. The population that the administrator is interested in generalizing research findings to in this case is the school district. The administrator randomly selects students from the district to participate in the after-school tutoring program (treatment group) and compares their reading achievement to the other students in the district who did not participate in the after-school program (control group). Whether the administrator finds that the after-school tutoring program improved or had no affect on students' reading achievement, the administrator will be able to generalize the findings to the population of interest (students in the entire district).

**Step Five: Select an Instrument**

Researchers select or design instruments to use in experimental research based on which instrument will best measure the dependent variable. For instance, in the earlier example of an
administrator implementing an after-school tutoring program in order to determine whether students who participate in the after-school program have substantially higher reading scores than students who did not participate in the program, obviously must base students’ reading achievement on results of some kind of instrument that is reliable and more importantly that the inferences we make from it are valid, such as the Iowa Assessments. Researchers often select instruments from those used in previous study of a specific topic. However, researchers also must be certain that any instruments they select are reliable and valid.

*Step Six: Administer Experimental Treatments*

Usually, once researchers design a research plan and answer certain questions about how experimental treatments will be administered to the treatment group, there is little deviation from this plan through the course of the experiment. Since the goal of experiments is often to provide evidence that a certain treatment caused differences in outcomes for the control and treatment groups, researchers must be certain to differentiate treatment between the control and treatment groups. When designing a research plan regarding how experimental treatments will be administered researchers should consider the following questions suggested by Lodico et al. (2006): “Specifically, what happens to the participants in each group? How does the treatment for the experimental group differ from that of the control group? (p. 187)”

*Step Seven: Collect and Analyze Data*

Keep in mind that how data is analyzed should be carefully worked out before the research begins. For more information regarding the methods used to analyze data from experimental studies see the *Statistics* section.

*Step Eight: Make a Decision about the Hypothesis*

Researchers must ultimately decide whether there is substantial evidence to support rejecting the null hypothesis. Since the null hypothesis states that there is no difference between the treatment and control groups, researchers generally embark upon research in order to provide evidence to the contrary; that the control and treatment groups differed as a result of the treatment implemented. Researchers determine whether substantial evidence exists for them to reject the null hypothesis based on probability. To learn more about what must occur in order for researchers to reject the null hypothesis see the *Statistics* section.

*Step Nine: Formulate Conclusions*

Once researchers have conducted statistical analyses of the data they have collected it is important that they formulate conclusions regarding the implications of their study. At times, the outcome of the study may be that the researchers have collected data that rejects the null hypothesis and supports the research question, might be a researcher’s ultimate goal, obtaining evidence to the contrary might prove to be just as important. For instance, if a district implements an after-school tutoring program with hopes that the program will increase students’ reading achievement only to collect data that proves that the program does not substantially impact students’ achievement, a district likely will save themselves the money, time, and man-
power that would have been necessary to implement such a program across the district when it would likely have not produced the desired outcome. Thus, studies in which researchers do not collect evidence to support their research hypothesis can be just as important.
References


