Finding the Answer with Numbers:
Overview of Quantitative Research Designs

The common feature in all quantitative research designs is an emphasis on using numerical descriptions and/or numerical comparisons. This overview is not about manipulating numbers. Instead, it is about what you will do with the participants in your study in order to produce the numerical values that represent what happens in your study.

There are many good reasons why quantitative research methodology may be your correct choice, but let's also be sure that you're choice is for the right reason. Do any of these apply to you?

1. *I want to do a quantitative study because qualitative studies aren't really scientific.* No offense intended, but this would be a really foolish justification for choosing a quantitative research approach. The statement is simply incorrect. Anyone who believes this is demonstrating a lack of understanding about the scientific method.

2. *I want to do a quantitative study because too much time is required to do qualitative research.* This rationale is based, at best, on a half-truth. It is true that researchers using qualitative methodology often spend more time with their research participants. But, a variety of other factors will contribute to the amount of time you will invest in a research study. Ultimately the amount of time that is required depends more on what you want to know and how you go about obtaining the answer, than on whether a quantitative or qualitative approach is used.

3. *I want to do a quantitative study because my program advisor told me I had to.* Your
program advisor is more likely to be saying that you must have research questions that require a quantitative approach.

Quantitative Decision Tree

Just Describing
1. Descriptive-Survey
2. Correlation-Relationship

Predicting an Effect
1. Control Group Comparison
   a. experimental
   b. quasi-experimental
   c. ex post facto
2. Self-Anchored Comparison
   a. simple time series
   b. multiple baseline
   c. alternating treatment
3. Correlation-Prediction

Just Describing

It is not unusual for a research question to describe some current situation. An instructor may want to know how former students now feel about the quality of the instruction in the courses. A school psychologist might be interested in the typical age of children referred for evaluation. A school administrator could be interested in the percentage of teachers who hold graduate degrees.

Studies to describe current status are often used in educational research to gather data that can be used to make policy decisions. For example, the answer to the question about the number
of graduate degrees would be needed to determine the financial impact of a change in the district’s salary schedule. These types of studies come in two basic “flavors”: descriptive-survey and correlation-relationship.

**Descriptive-Survey**

This is a widely used design often involving the administration of some form of questionnaire survey. If your research question is focused only on describing the current situation, and if quantitative information is sufficient to answer the question, this is a quite satisfactory and relatively simple design.

Traditionally these surveys were in the form of paper and pencil questionnaires. More recently, you are much more likely to receive the invitation to participate through an email message that directs you to a web page for online responses to the questions.

In case you might be thinking that this design seems in some ways very similar to a qualitative research design, you would be correct. What makes this a quantitative design is that the emphasis will be on data provided in numerical form. Major limitations in use of this design are mostly associated with the sample of participants and/or the quality of the questionnaire used to gather the data.

**Correlation-Relationship**

Like the descriptive-survey strategy, this design is also characterized by procedures involving gathering current status information. The difference is that the correlation-relationship design provides a little more descriptive information, specifically information about the relationships between and among the factors being described.

To illustrate, consider a simple survey in which the target population is graduate students. The instrument is a questionnaire on which participants provide their age, number of credits
being taken in the current semester, number of study hours per week, number of hours per week in a full or part-time job, and current GPA.

A research objective of this design would be to provide information about currently enrolled graduate students. The outcome of a descriptive-survey design would be the relevant data for each of the four questions (i.e. the average response for age, number of credits, and so forth). A correlation-relationship design would simply add information about relationships among these variables, typically expressed as **correlation coefficients**. For example, the researchers might expect a negative relationship between hours of employment and hours of study; as one goes up, the other goes down. In essence, this design could be thought of as just a descriptive survey with more extensive analysis of the responses.

Primary limitations of the design are also quite similar to those in the descriptive-survey approach, mostly involving the selection of participants and quality of the instrument used to obtain the data. With the correlation-relationship design the number of participants also becomes an important concern. Correlation coefficients, the numerical tool for describing the relationships, are not particularly stable when the number of participants is small. A minimum of 50 participants is a typical recommendation when there is intent to generalize from a sample to a larger population.

**Predicting an Effect**

The designs on the right-hand side of the decision tree are needed when your research question involves more than a description of current status. Your inquiry, for example, might be whether different basal reading programs result in different outcomes on standardized achievement tests. How well could you predict scores on the reading tests if you knew which basal reading program was used? Or, you might want to know if the number of counseling
sessions used in a typical case was contingent on whether the service was delivered in a private practice setting or in an agency.

These designs are more complex than those in which you only want to describe current status. You have to decide how you want to analyze the data and decide how you will create group(s) for comparison.

When you want to know if you can predict an outcome, you have several alternative approaches. There are three broad categories and two of the categories, control group comparison and self-anchored comparison, have three subcategories each.

**Overview of Prediction Designs**

In the decision tree, the first broad category, **control group comparison**, involves just what the name implies. Something is done, or has been done, to one group and not to another. What you want to know is: Does it make any difference?

If you actually create the differences between the groups for the independent variable (the thing “done” by you to them), and assignment to those groups was truly random, you are using what is called an **experimental** design. If you were not able to randomly assign participants to the groups, your design is **quasi-experimental**, meaning “almost an experiment.” In both instances, you play an active role in making the groups different from one another. This is in marked difference from the **ex post facto** design in which you do not actively "make" the groups different and then observe the outcome. Instead, for the independent variable (the intervention), you use preexisting differences between groups (e.g. gender, age). Likewise, for the dependent variable you may use data that already exist (e.g. school records, case files). This strategy is called ex post facto because it occurs “after the fact.”
All of the control group comparison strategies involve comparing one group of participants to another group of participants. The **self-anchored comparison** strategies in the decision tree hold the participants constant and observe what happens over time. Measurement occurs before and after the experimental conditions have been implemented. The key feature in all self-anchored designs is the self-comparison. Unlike the other strategies, it can be used with just one participant, but the self-comparison designs can work equally well when the focus is on a larger group, for example a classroom.

Finally, the **correlation-prediction** research strategy differs from the other two broad categories because of a primary emphasis on using correlation coefficients to make specific outcome predictions. Its primary distinguishing feature is not in how the comparisons are conceptualized but in how the data are analyzed.

**Control Group Comparison**

*Experimental.* The logic of the true experimental design is straightforward. In essence it simply involves creating comparison groups by randomly assigning participants to one group or the other. When you use this design you have more control over extraneous features than is available in most of the other choices, but as we’ve mentioned that control comes with a cost. You cannot use this technique with existing data; you must create the conditions that make the persons in the groups different from one another. And, even when you are going to be creating those conditions, your design meets the criteria for identification as an experiment only if participants are randomly assigned.

Random assignment to the treatment conditions is a crucial element if you want to do use the experimental design, but can you be absolutely certain that this will insure that the control
and experimental groups are alike on the extraneous variables? The answer is no. Randomization may balance those other variables between the groups, but it may not. Researchers often include a pretest in the experimental design to provide evidence of equity between the experimental and control groups. If it is needed at a later time, the researcher can then adjust the mean scores during the analysis phase in such a way to reflect the pre-existing conditions (analysis of covariance).

**Quasi-Experimental.** In a quasi-experimental design, you will “make” the groups different, but you are not able to meet the criteria of random assignment to the groups. This happens frequently in educational research because of the logistical requirement to use intact classroom groups. For example, consider a university professor teaching two sections of a course on research methods. In one section, the researcher will display an outline of lecture notes on the screen during the class session. In the other section, the class procedure will be essentially the same except that the lecture notes will not be displayed. The independent variable is the display of the lecture notes. The dependent variable will be scores on a final examination.

Use of a true experimental design is not possible in this example because the researcher cannot control who actually enrolls in which section of the course. Especially if there are more than two conditions of the independent variable, researchers may randomly assign which group gets which condition, but this alone is not sufficient to meet the criteria to be classified as an experimental design.

**Ex Post Facto.** Schools and agencies have stored in their files a vast amount of information that could be analyzed by researchers to investigate cause-prediction relationships. Using such data has many advantages, not the least of which is that the study does not require any additional interference in the lives of the participants. Let's say, for example, that a research
question investigates the time that students spend "on-task". The dependent variable, school achievement, is available in school records as the scores on the district's standardized achievement test. The independent variable, time on-task, could be created from school attendance records.

In order to create this independent variable, the researchers might analyze the attendance records for a specific grade level and, based on the number of days absent, identify three groups for their independent variable. One group would be those students who were rarely absent; a second group would be those students who were occasionally absent; the third group would be those students who were frequently absent. The researchers would probably want to try to define the limits for those groups so there would be approximately the same number in each group.

The "participants" in this study would be all students at that grade level for whom both absentee records and achievement test scores were available. To address their research question, the researchers could then compare the average achievement test scores for each of these three groups and identify whether there was a relationship between number of days absent and scores on the achievement test.

The researchers in this example are free to classify the absentee levels in any way they choose, so long as they clearly report how the categories were formed. These are called operational definitions and are found in all research designs. You might not agree with the way the categories were formed. In some instances, the definitions used for the variables may have significant impact on whether you believe the design can address the research question. The obligation of the researchers is to clearly specify the operational definitions used for both independent and dependent variables in a study.
There are many advantages in the ex post facto design procedures. Extensive amounts of data that could be used to address research questions are gathered and stored each year and often just sit gathering dust (or whatever the equivalent would be for files stored on a computer). With this design the researchers can often include significantly larger number of participants than would be feasible with most other designs.

It is also possible with the ex post facto design to obtain data in situations when ethical considerations would preclude other alternatives. For example, consider a situation in which: a) the dependent variable is symptoms of depression, and b) the research question is whether the symptoms improve because of interventions or instead just because of the passage of time. Withholding treatment from a control group in order to conduct a true experimental design might cross an ethical boundary line. As an alternative, many agencies would have already collected measures of depression on their clients during the intake process. These agency records could be used, after the fact, to compare change in depression scores between those who immediately began receiving therapy and those who had to wait.

The term causal-comparative is sometimes used to identify the ex post facto type of research design. It's the “old” label and is not particularly descriptive of the design. Comparisons are used in the design, but the difficulty in controlling extraneous variables makes it even less likely that cause could be determined (if such a thing is even possible).

**Self-Anchored Comparison**

Implicit in the designs we've just completed is that the groups being compared need to be “alike” on everything other than the independent variable. What if there was a way to be even more sure about whether that this was true? There is a way, and it's found in the self-anchored research approach, often referred to as single-case or single-subject experimental design.
The basic logic of this design strategy has an elegant simplicity. A teacher has a student exhibiting conduct problems in the classroom. The frequency of the conduct problems serves as the dependent variable. The teacher wants to know if intentionally ignoring the problems will change how often the problems are occurring. This would be the independent or treatment variable, intentionally ignoring the problems. So, in simplest form, the teacher identifies how often the problem behaviors are occurring, then introduces the independent variable and observes whether there is any change in the frequency of the problem behaviors. If there was a change, how could the teacher know whether it was the result of ignoring the behaviors and not something else? Again in simplest form, the teacher would then stop ignoring the behaviors and observe whether the frequency went back to what it was before the independent variable was introduced.

This is the classic single-subject reversal design, often referred to in the research literature as an ABA design. How was it before the treatment began? How was it after the treatment began? When the treatment stopped, did it return to the way it was before? The idea, probably obvious, is that if there was change associated with the treatment, and the change went away after the treatment stopped, it would appear reasonable to assume that the treatment may have been causing the change.

The first step in this strategy will be for the teacher to determine how often the problem behaviors are now occurring (an important consideration, by the way, for all behavioral change interventions). This is called the baseline, quantification of the status of the dependent variable, before the independent variable is introduced. The baseline phase is akin to the pre-test in a control group design, but with a very important difference. In the self-anchored research strategy it is assumed that more than one measurement will be taken, more like a videotape than a
snapshot. To have any basis for attributing change to the independent variable, it is crucial to be sure what the status was before the independent variable was introduced. In the example, several periods of observation would be required.

The **treatment** phase in this strategy is when the independent variable is introduced. Again (notice the theme) a series of measurements is required, not just a single observation. Notice also, though, that if change is evident during this period it would not necessarily be reasonable to assume that the treatment was causing the change. The independent variable is not the only thing that is different. Time has passed. Other students in the classroom may have changed. The material may have become more interesting. The list goes on and on.

In attempt to determine if the independent variable was, in fact, the cause of the change in the dependent variable, the classical procedure is **reversal**, a return to the baseline. In the example, the teacher would stop ignoring the misbehaviors and observe whether the student's behavior returned to where it was before the study began.

**Simple Time Series.** The simple time series strategy is easily implemented in many regular professional activities, and although not a particularly good research design, it warrants some attention because it is often good professional practice. To use this design, you need repeated measurements of the dependent variable before you introduce the independent variable, and then to continue the repeated measurement for a period of time after the independent variable is introduced. For example, the dependent variable could be students' scores on spelling tests and the independent variable could be adding an extra review session before each test. If the independent variable is something that is not intended to be a permanent addition, for example extra review sessions or a student receiving counseling, the process continues by stopping the independent variable but continuing the repeated measurement.
If this seems very much like the classic reversal design, you are right. The primary difference is in the desired outcome when the independent variable is stopped. In the classic reversal design, the researcher would hope that the dependent variable returns to the baseline level. In many, perhaps most, situations in educational research, the hope would be that the dependent variable does not return to its initial level.

A key to the use of simple time series, and all of the self-anchored comparison designs, is the repeated measures of the dependent variable. During the first phase, baseline, measurement continues until the scores (or class average scores) are essentially flat, not getting better, not getting worse. The second phase, when the independent variable is introduced, is the treatment phase, and measurement continues during this phase long enough to be sure that sufficient time has been provided for effects of the independent variable, if any, to be evident. And, when appropriate, a third phase, follow-up, will continue with repeated measurement of the dependent variable to see if there is a lasting effect of the treatment.

There is not unanimity of opinion among researchers about how many measurements are required in each phase. When the results are just plotted on a graph for visual analysis, a minimum of three measurements or data points is often recommended. One statistical analysis technique for the simple time series assumes a minimum of eight measures in each phase. In either instance, the baseline should be flat before the independent variable is introduced.

**Multiple Baseline.** Multiple baseline designs attempt to control for the effect of extraneous variables by concurrently conducting a study and a replication of it at the same time. Sound complicated? Actually, it's just an elaboration of the simple time series.

To illustrate, assume that you are a teacher whose schedule includes three periods in which you teach reading to different groups of 9th grade students. Daily activities in each class
include an assignment to read a story and then answer questions to determine comprehension. You want to find out if preparing and distributing study guides that direct attention to important elements in the assigned materials (independent variable) will increase the comprehension level (dependent variable). With concern about the impact of extraneous variables, you decide to do this study with students in the first period and the second period classes.

The first step is to gather the baseline data, averaging test scores on the comprehension assignment each day in class. Data gathering continues until there is a stable baseline evident in both groups. Then, the study guides are introduced in one of the classes while the baseline data gathering continues in the other. If the average scores on the comprehension task go up in the class with the study guides, and stay about the same in the other class, you have some preliminary evidence that the study guides were effective. The next step would be to introduce the study guides in the other class and observe if improvement was also evident in that one as well.

It's really a quite logical process. If the group without the study guides made sudden improvement in comprehension at the same time the study guides were introduced in the other class, the conclusion would be that something other than the study guides was causing the difference. If the first group made improvement with the study guides, but the second group did not, you would want to consider other characteristics in the first group that might have caused the improvement (or characteristics in the second group that might have precluded improvement).

**Alternating Treatment.** The third self-anchored design, alternating treatment, provides control of extraneous variables through an interesting application of randomization. The essence of this design is a random presentation of at least two forms of treatment and continuous
monitoring of the effects. It is like the other self-anchored designs in that measurement is ongoing throughout the study. It differs from the other two self-anchored approaches because a baseline period is not an absolute requirement.

To illustrate, think about a special education teacher working with students who have problems with test anxiety. Two forms of treatment will be compared: 1) weekly individual consultation sessions and 2) participation in small group tutoring sessions. All participants in the study will receive the same treatment; each week they will have either an individual or group session. Before treatment begins, the teacher will create the schedule for the sessions using a random order. An anxiety measure will be administered each week.

At the end of the study, the scores on the anxiety measures will be collated for comparison of the reports of anxiety during the weeks that followed an individual session and the reports during the weeks that followed the group sessions. The effect of extraneous variables is controlled in this design by randomizing the order in which the treatment options are provided.

**Correlation-Prediction**

Correlation techniques were introduced before as tools to enhance descriptive studies. This method can also be used with a different purpose in mind—to use the data obtained from one group of participants to predict the outcome for some future group.

To illustrate, go back to the example of a correlation-relationship study. Participants were graduate students. The variables were age, number of credits taken in the current semester, number of study hours per week, number of hours per week in a full or part-time job, and current GPA. To simply describe current status, the researchers would provide the correlation coefficients among the five variables.
Those same data could also be the basis for a correlation-prediction study. Using a statistical technique called regression analysis, you could determine which of the other four variables appeared to be the most effective predictors of grade-point-average. The four “predictor” variables would then become the “independent” variables for the study while grade-point average remains the “dependent” variable.

The statistical regression analysis of these data would give you information about which, if any, of these four independent variables appears to be useful in predicting the dependent variable. After that is determined, you can also calculate a regression equation in which each of the predictor variables is weighted to provide the best possible prediction of the dependent variable, in this case the overall grade-point average. Note: don’t worry when we say “you can calculate a regression equation.” There are software programs that do most of the work for you.

The idea, of course, is not really the prediction of GPA for the participants. After all, their grade point averages were already known. Instead, the intent of this type of study is to predict the dependent variable when you only have information about the other variables. For example, when students take a college admission test, the score report often includes a prediction of first-year grade point average. That prediction came from analysis of information about the relationship between test scores, high school grades, extracurricular activities, and so forth, with grades obtained after the first year of college. In making the predictions, there is an assumption that the persons for whom grades are being predicted are like the persons who provided the original data.

Correlation-prediction analysis can easily be applied in many situations involving use of existing data, providing some advantage over procedures that require group comparison. For example, consider a study in which the research question involved the impact of attendance on
achievement test scores, using data obtained from school records. The researchers could use a comparison approach, using the records to identify three groups: rarely absent, occasionally absent, and frequently absent. In order to create the three groups, lines had to be drawn. For example, the occasionally absent category might have included persons who were absent at least 15 but not more than 25 days.

Logically, a student who was absent 15 days would have more in common with a student absent 14 days than with a student absent 25 days. But, the "14 day" student would contribute to the mean scores in a completely different category, the rarely absent group. When categories have to be created by the researchers, such loss of precision is inevitable. Using correlation analysis avoids this problem. The researchers could instead have used the attendance records as a predictor variable and achievement test scores as the dependent variable, and then analyzed how well the test scores could be predicted by the actual number of absences.