CHAPTER 3

CONSTRUCTING RESEARCH DESIGNS

The design used in your research will be determined, at least in part, by your research question. Some designs are fairly simple and others are extremely complicated. If you have succeeded in narrowing your topic down and if you are able to control extraneous variables, you may be blessed with a fairly simple design. However, for most studies in Applied Linguistics, particularly those related to classroom research, the design may be complex. If you wish to be able to generalize from the results of your classroom experiment to other classrooms, from your students to other students, you will need to choose a design that allows you to share your findings as being relevant to other teachers and other classrooms.

In classroom experimentation we must be especially sensitive to the problems of external and internal validity (discussed in Chapter 1). You will remember that we must be sure, whenever we make a claim about the effectiveness of any instruction, that the students not only would not have made the same gains without the instruction but also that they are really random representatives of language learners. If they are all supertalented learners, the results obtained might never again be replicated. A careful choice of design will help you avoid these problems.

One way of avoiding problems is to use a control group in your experiment. Suppose you want to investigate the effect of grammar correction on the writing skills of ESL students. Your independent variable will be the amount of correction and the way correction is given on composition errors. The dependent variable is the degree of grammatical accuracy in your *S*s' writing samples. If, at the end of the semester (or year, or other time period you select), you notice considerable improvement in grammatical accuracy, you might be willing to conclude that the improvement was related to correction. However, we could claim that everybody would improve over that period of time without correction. Thus, your conclusion may be wrong. To deal with this problem, you need to have a control group for comparison purposes. A *control group* refers to *a group of* Ss *whose selection and experiences are exactly the same as the experimental group except that they do not receive the experimental treatment.* If you selected two similar groups of ESL students and corrected the errors of one group but did not correct those of the control group, you still found improvement in your experimental group that far outweighed the improvement in the control group, then your conclusion would be much more defensible. If it's really only a matter of time, there should be no difference between the two groups. Having a control group contributes to the internal validity of the research and lets us interpret our findings with more confidence.

We have already mentioned random selection of our sample in previous chapters. In classroom research, random assignment refers to the method of selecting and assigning your *S*s to experimental and control groups. The notion of randomization is of crucial importance since it allows the researcher to have two truly comparable groups prior to the start of the experiment. If the experimental and control groups are truly equivalent, then you can feel fairly confident that everything except the treatment is the same. Any difference between the groups after instruction can be associated with the treatment. There are many ways in which random assignment of *S*s can be carried out. The more common ways are rolling dice, flipping a coin, or drawing numbers out of a hat. Another method is to use a table of random numbers which you can find at the back of many statistics books. However, the method is not important so long as each *S* has every chance to be assigned to any one of the groups used in your research. If there are important differences among your *S*s and you wish to be sure that equal numbers of *S*s with those characteristics are in each group, then you still need to randomize assignment of *S*s within those subclassifications.

After you have identified and appropriately selected your *S*s, you must next begin to consider the most appropriate research design. We will discuss five major classes of research design. The purpose of each design is to try to avoid as many research errors as possible so that you can share your findings with others. That is, you need to select the design that will allow you to feel confident in discussing your findings and allow you to generalize over and beyond your limited study. The major classes of design to be discussed here are: pre-experimental, experimental, quasi-experimental, and ex post facto.

PRE-EXPERIMENTAL DESIGNS

Pre-experimental designs are not really considered model experiments because they do not account for extraneous variables which may have influenced the results. The internal validity of such a design is also questionable. However, they are easy, useful ways of getting preliminary information on research questions. (Also they are good examples of what you should *not* do when you carry out certain final research projects.) The three most commonly used pre-experimental designs are the one-shot case study, the one-group pretest posttest, and the intact group comparison design.

The one-shot case study

In this design, there is no control group and the *S*s are given some experimental instruction or treatment (labeled *X*) for a given period of time. At the end of the period of time, the *S*s receive some sort of test (labeled *T*) on the treatment. The schematic representation of this design is

$$X \quad T$$

Very simple. But this design is open to almost all our questions about research validity. The results of such a study are neither valid nor generalizable. For example, if you used a new technique of teaching pronunciation to your junior high ESL students and at the end of the unit administered a pronunciation test, there is no way to conclude that your technique alone was the reason for the improvement. There are too many uncontrolled factors which could have contributed to your *S*s' scores. The findings may be useful to you in deciding whether or not to pursue this research question further, but that's about all.

One-group pretest posttest design

This design is similar to the one-shot case study. The difference is that a pretest is given before instruction (or treatment) begins. So there are two tests: T_1 = the pretest, and T_2 = the posttest. X is used to symbolize the treatment in the following representation of the design:

$$T_1 \times T_2$$

This design is an improvement over the one-shot case study because you have measured the gains that the subjects have made rather than just looking at how well everyone did at the end. However, without a control group, you still cannot make justified claims about the effect of the instruction.

Suppose you wanted to find out whether speed-writing exercises affect grammatical accuracy as well as writing fluency. You gave a composition test under timed conditions before you began your teaching unit on speed writing. You counted the number of words per minute written as your fluency measure and number of errors per 100 words as your accuracy measure. Every day you began class with an eight-minute speed-writing exercise. At the end of the unit you gave your posttest. The gains (assuming you got gains on the two measures) might, then, appear to be related to your instructional program. But this type of design is also open to the question of internal validity. You can't be sure that the improvement might not have been due to other factors—for example, the students' history teacher might have assigned weekly written homework which contributed substantially to the gains.

Intact group design

This is the design that most classroom researchers use. It is often impossible for us to assign students randomly to language classes. Students are placed in classes on the basis of some criterion (e.g., scores on a placement test, successful completion of the prior course, or even self-selection according to the time the class is offered). However, by selecting two classes for your study, you can use one of them as the control group. Both experimental and control groups will receive a posttest, but the experimental group will receive the treatment while the control group does not. You may toss a coin to see which of the two groups becomes the experimental group and which the control group. The presence of the control group in this design eliminates some of the problems related to internal validity. Yet, we have to be cautious about generalizing the results of the study beyond our experiment because the *S*s in the study have not been randomly assigned to the two groups (so we do not have high external validity). In the following representation, the letter *G* stands for group:

$$\frac{G_1 \times T_1}{G_2 \quad T_1}$$

In addition to the selection problem (which influences external validity) there is a problem of internal validity because the groups may not have been equivalent to start with. The majority of students in one group may have had a very special instructor the semester before and received a fantastic program to improve their communication strategies. If your research is on improvement in communication strategies and all this instructor's students are in the control group, there will probably be no difference in the groups after your treatment, and you might decide that your instruction was a dismal failure (when it wasn't at all). Preexisting differences in the two groups could be a potential factor that would influence your results.

In order to make the difference in pre-experimental design types clear, we have limited the examples almost entirely to evaluation of the effect of instruction; it's easier to explain design in this way. We do not mean to mislead you into thinking that research design applies only to the evaluation of instruction. However, for studies which are concerned with the effectiveness of classroom instruction, pre-experimental designs (whether one-shot case studies, one-group pretest posttest designs, or intact group comparisons) do not give us results which we can claim as completely valid. They are questionable in terms of problems involved with internal and external validity. Yet, even for classroom instruction research, this does not mean they are without value. Sometimes we are not sure whether we want to put a great deal of time and effort into investigating our research questions until we have some preliminary evidence to support our ideas. If the results show a trend in the expected direction, then the researcher can think about carrying out a more rigorous, well-designed study. In that case, the label "pre-experimental" seems appropriate.

However, it is also the case that a pre-experimental design may be a sound one for research. For example, an ethnography or a case study of a child second language learner might include data from early morning grooming activities, from mealtime exchanges, from TV viewing, from the playground, from bathtime, from bedtime, etc. Each of these activities can be considered a "treatment" which might produce a larger or smaller number of instances of some variable you believe is of importance to language acquisition. The design would be a pre-experimental, one-shot case study (several separate treatments). The design, in this case, is sound, for no strong claims about the nature of second language learning are going to be made on the basis of one case study. In such studies, the label "pre-experimental" does not mean that the design is only something one might do before doing a true experiment. For some research questions, it is just as likely that researchers might first do a true experiment and then a one-shot case study. Unfortunately, the labels for the designs seem directional, perhaps even judgmental. We do not regard them as such. Perhaps we ought to flaunt tradition and give them new labels. The appropriateness, the soundness of any design depends on the research question and the kinds of claims to be made about our findings.

TRUE EXPERIMENTAL DESIGNS

True experimental designs have three basic characteristics: (1) a control group (or groups) is present, (2) the *S*s are randomly selected and assigned to the groups, and (3) a pretest is administered to capture the initial differences between the groups.

These three characteristics allow us to avoid almost all the problems associated with internal and external validity. The two most common experimental designs are described below.

Posttest only control group

In this design, there are two groups—an experimental group which receives the special treatment and a control group which does not. The *S*s are randomly assigned to one or the other group, and the decision as to which group will be the experimental group is also decided randomly (e.g., by the flip of a coin). In this design, initial differences between the groups are controlled for by the random selection and random assignment of the *S*s.

$$\frac{G_1\,(\text{random}) \times T_1}{G_2\,(\text{random}) \quad T_1}$$

Pretest posttest control group design

This design is the same as the previous one except that a pretest is administered before the treatment.

$$\frac{G_1\,(\text{random})\; T_1 \times T_2}{G_2\,(\text{random})\; T_1 \quad T_2}$$

You might wonder why anyone who had gone to all the trouble of getting random groups to begin with would not make the effort to give a pretest as well. The reason is concern about the effect of the pretest. As a general rule of thumb, if the time between the pretest and the posttest is not considerable—at least two weeks—you should seriously consider whether or not to give a pretest. If the time interval is sufficient to make you feel fairly confident that it will have little or no effect, then a pretest should be given.

QUASI-EXPERIMENTAL DESIGNS

The concept of experimental design is an idealized abstraction. The ultimate goal of any investigation is to conduct research that will allow us to show the relationship between the variables we have selected. However, in social sciences in general, and in our field in particular, it is not realistic to limit our research to true experimental designs only. The reason is that we are dealing with the most complicated of human behaviors, language learning and language behavior.

In much of our research, it seems quite unlikely that we can have a true experimental design. For example, in observational studies of child second language learners, one is lucky to be able to find one child learner interacting with one native-speaker child, let alone having a group of randomly selected children. In classroom research, it is unreasonable to expect that we can ask a director of courses to randomly assign foreign students to classes for the benefit of our research.

In addition, it is also very difficult to carefully define many of the numerous variables involved in most Applied Linguistics research. One can hardly be certain that the treatment, say reading practice, involves only reading and not any other aspect of language learning or language behavior. Many of the research studies evaluating various methods (audiolingual vs. cognitive code, translation vs. aural-oral, etc.) have been faulted on just these grounds. Can we be sure that the methods are mutually exclusive? Do the treatments really never overlap? The same can be asked about the skill areas: are they really mutually exclusive language areas?

Another problem for us to consider in our research is whether we are really controlling factors when we say we are. Just because we say we've controlled some factor doesn't necessarily mean we have controlled for it. For example, we may have "controlled" for language proficiency by selecting only advanced *S*s in some research project. But those advanced students may have very different advanced abilities—some may be extremely good at grammar but unable to understand or speak English; others may be very good at oral communication but unable to write well. So you may not really have controlled for proficiency in the particular skill areas which may be important in your research.

The impracticalities involved in planning research in our field are sometimes overwhelming. Many can be solved but some simply cannot be changed at will. Even if we have randomly selected our *S*s, we have to remember that individuals are always different, and the ways they approach language learning and language use are also likely to be quite different.

Furthermore, it is impossible to ask a group of students to serve as a control group if it means depriving them of valued instruction, or if you expect them to waste their time, energy, and tuition fees because of your research. It is also unreasonable to expect students to take tests, fill in questionnaires, or participate in an experiment if it is *only* to give you data for your experiment.

Because of these and many other limitations, constructing a true experimental design may be difficult if not impossible. However, it does not mean that we should abandon research or that our studies need to be invalid. Our goal should be to approximate as closely as possible the standards of true experimental design. The more care we take, the more confident we can be that we have valid results that we can share with others. However, if we reduce our experiments to highly artificial laboratory-type experiments, we must also worry about whether the results can be directly transferred and shared as valid for the classroom. There is certainly a trade-off between the degree of experimental control and the possibility of obtaining generalizable results. Our goal, once again, is to strike a balance which will allow us to collect meaningful data in ways that allow us to share our findings with others.

Quasi-experimental designs are practical compromises between true experimentation and the nature of human language behavior which we wish to investigate. Such designs are susceptible to some of the questions of internal and external validity. However, given the present state of our art, they are the best alternatives available to us.

By using a quasi-experimental design, we control as many variables as we can and also limit the kinds of interpretations we make about cause-effect relationships and hedge the power of our generalization statements. Time-series designs allow us to do this.

Time-series designs

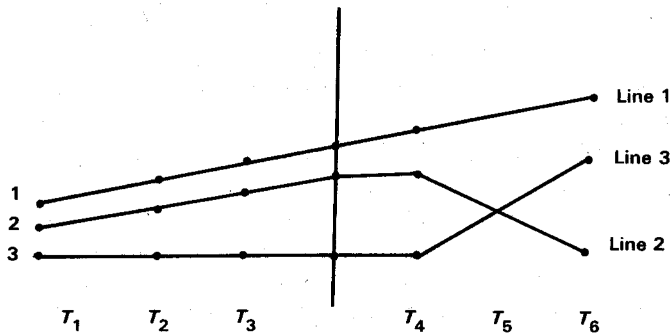
Because of all the limitations mentioned above (and probably more besides), it is sometimes impossible to have a control group for your research project. For instance, you want to test some particular new technique and so you look for another ESL class to serve as your control group. You find one that is at the same level as yours, but they are using a completely different set of teaching objectives than you are. They are not an appropriate, "equivalent" group. You might decide to use one of the designs mentioned earlier—the one-shot case study or the one-group pretest and posttest design. Another way of dealing with the lack of a control group is the time-series design.

In this design, the subjects are administered several pre- and posttests. This design can be represented as follows:

$$T_1\ T_2\ T_3 \times T_4\ T_5\ T_6$$

After the three pretests, the researcher will have an idea about the possible changes in the *S*s' behavior when there is no special treatment. The treatment is then introduced, and finally a few posttests are given to look at the improvement after the treatment. There is no magic number of pre- or posttests that should be given, but you do need enough to give you a learning curve.

As an example, suppose that you have developed a set of programmed materials to teach English articles for your class of Japanese university students. You have not been able to find similar groups to serve as control and experimental groups. You have only your students in your particular ESL class. You start administering tests on articles at the end of each week for the first three weeks of the class. By week four, you can see the improvement that your students have made in article usage. Then you give them the programmed materials to work through. Following this, you start giving posttests on articles at the end of each week following the treatment. A comparison of the curves for the first three weeks with those after the treatment will give you an idea whether your innovative materials worked or not. You might come up with any of the following possibilities:



If you obtained a line similar to Line 1 it would show that there was no effect from your special treatment. The students continued the same developmental pattern that you saw during the first three weeks of the study. If your results were like those in Line 2, it would indicate that your materials had a negative effect, since after the treatment their scores declined consistently. (Maybe the materials confused the *S*s more than they helped them.) Line 3 is the line you would hope to get, since it shows that your treatment was effective, for the improvement is much more dramatic than one would expect given the nonimprovement curve during the first three weeks.

Another version of the time-series design is the equivalent time-sample design. It works like this: the treatment is introduced and reintroduced between every other pre- and posttest. That is, after the first pretest, the treatment is introduced followed by a posttest. Then, after the second pretest, an alternative treatment (nontreatment) is introduced and that is followed by a posttest. This procedure is followed for two or three times, and the results following the experimental treatment are compared with the alternative treatment scores. Using X for treatment and a zero (0) for no treatment, this could be represented as:

$$T_1 \text{ X } T_2 \rightarrow T_3 \text{ 0 } T_4 \rightarrow T_5 \text{ X } T_6 \rightarrow T_7 \text{ 0 } T_8, \text{ etc.}$$

There are other alternatives to dealing with the problem of finding a control group. And you should try to find one if at all possible. This could be done by asking the class to participate in the experiment outside of class (in return for additional free English lessons as they participate in your experiment or for a fee if you have research funds). Another possibility is dividing the class into two groups (random assignment) and asking them to come on alternate days, or have one group come a half-hour early and the other group stay a half-hour late. Unless your class is very large, splitting the group in half may give you very few students in each group. In short, having a control group is best, but if it's impossible, then the time-series designs may be an alternative you will want to consider.

EX POST FACTO DESIGNS

When researchers control the threats to internal and external validity, they are trying to find a direct relationship between the independent and dependent variables. In other words, they select the population, sample, treatments, and variables in order to find a cause-and-effect relationship between the variables. As you can see, there are many obstacles which prevent us from designing studies that will allow us to make such claims.

For example, you may have created a series of media lessons on how to say *no* to requests in English. As long as you do not randomly select your *S*s, organize your control and treatment groups, and control for factors aside from the media lessons which might influence the results, you cannot draw causal relationships between your media materials and *S*s' improvement in ability to turn down requests gracefully in English.

When you consider all the factors that you would need to control, you might think that designing a true experimental research project is almost impossible. In some sense that is true. However, it should not mean that we have to give up approximating the ideal as much as possible. Claiming that X causes Y is an extremely difficult thing to do unless the research is carefully designed and as many extraneous factors are controlled as possible.

These problems have led researchers to look for other designs with fewer restrictions on them. The trade-off of conducting a less controlled design is that we have to be very cautious in interpreting our results. When there is no possibility of random selection of *S*s, instead of abandoning the research, we simply have to limit the domain of our claims. We have to avoid making cause-and-effect statements.

Ex post facto designs are often used when the researcher does not have control over the selection and manipulation of the independent variable. This is why researchers look at the type and/or degree of relationship between the two variables rather than at a cause-and-effect relationship.

For example, we can study the relationship between scores on a school-leaving exam in ESL and teachers' ratings (or grades) for the *S*s using an ex post facto design. We will be able to see if there is a certain amount of agreement between the two sets of scores. It does not mean that one is the cause of the other. Suppose we wanted to know about the performance of two groups of students (say, one group is from China and the other from Venezuela) on an entrance exam given to foreign students at our university. Any relationship between the scores of the groups would not be related to any instructional program we had given them before the test. That is why the designs are called ex post facto. The researcher has no control over what has already happened to the *S*s. The treatment, whatever it might be, has been given prior to the research project.

Figure 3-1

Proficiency level	Iranian	French
Advanced		
Intermediate		
Beginner		

Nationality

Correlational designs are the most commonly used subset of ex post facto designs. In correlational designs, a group of *S*s may give us data on two different variables. For example, many students who plan to study in the United States take the TOEFL (Test of English as a Foreign Language). Many universities also have their own entrance exam which they administer to students. We can then look at the relationship of *S*s' scores on one test to their scores on the other. Or, foreign students may be asked to take both the Graduate Record Exam (GRE) and an English placement exam prior to admission to a university. The score for each *S* on one test can then be compared with the score on the other, allowing us to see whether those students who score high on one test also score high on the other.

The schematic representation of this design would be

$$T_1 \quad T_2$$

Since there is no causal relationship between the two variables, the distinction between independent and dependent variables is not well defined. It is arbitrary to call one or the other the independent variable. However, it is usually the case that the investigator may be more concerned with one than the other and may therefore label the first the independent variable and the second the dependent variable and show this by the labels X and Y.

Another ex post facto design is called a criterion group design. In this design, two groups of *S*s are compared on one measure. With this design, you might, for example, measure the reading speed of Iranian and French students, assuming you want to see how related or different they might be. The design would look like this:

$$G_1 \quad T_1$$

$$G_2 \quad T_1$$

You can change the design into a two-criterion design by considering level of language proficiency as well as their native language. In this case the criterion group design forms a factorial design (Figure 3-1). So far, we have talked about four major types of research design: pre-experimental, true experimental, quasi-experimental, and ex post facto. In each of our examples we have been concerned with the relationship between only two variables—one independent variable and one dependent variable. However, it is possible that your research will be concerned with more than one independent variable (you may have moderator variables as well). In this case the design is factorial.

FACTORIAL DESIGNS

Factorial design is not really a design type in itself. It is simply the addition of more variables to the other designs. There will be more than one independent variable (i.e., moderator variables) considered and the variables may have one or many levels.

Suppose you believe that massive reading practice will transfer over to improvement in another skill, listening comprehension. You decide to do an experiment to test the influence of reading practice on listening comprehension. The dependent variable is performance on listening comprehension tests. The independent variable is amount of reading practice. You randomly select some *S*s to represent second language learners and randomly assign them to either the control or experimental group. After a period of time during which the experimental group receives practice reading and the control group receives some irrelevant practice on some aspect of language, you administer a listening comprehension test and compare the performance of the two groups. So far we have a posttest-only control group design:

$$\begin{array}{lll} G_1\text{(random)} & X\text{(reading practice)} & T_1 \\ \hline G_2\text{(random)} & & T_1 \end{array}$$

The study (assuming the treatments are mutually exclusive) is valid and sound in design. However, suppose you wanted to expand your research design and ask questions about the language proficiency of your *S*s.

Suppose you have a hunch that advanced learners get more benefit out of reading practice than elementary students—or even the reverse, that elementary students get more benefit from reading in terms of understanding oral discourse. To investigate this possibility, you include language proficiency as a moderator variable with two levels—advanced and elementary. Your design would now look like this:

$$\left.\begin{array}{lll} G_1\text{ (random)Advanced} & X & T_1 \\ G_2\text{ (random)Elementary} & X & T_1 \end{array}\right\}\text{ Experimental groups}$$
$$\left.\begin{array}{lll} G_3\text{(random) Advanced} & & T_1 \\ G_4\text{ (random)Elementary} & & T_1 \end{array}\right\}\text{ Control groups}$$

The design will now allow you not only to talk about the effect of reading on listening comprehension but also to show the differences (if any) for learners at two different levels of proficiency.

If we consider the possible outcomes, there are many. Let's represent a few of them in graphic form. (You may want to try to sketch out some of the other possibilities which we have not included here.)

1. Listening comprehension scores either increased or decreased following massive reading practice vs. no treatment: ($G_1 + G_2$ vs. $G_3 + G_4$). (See Figure 3-2.)

2. Listening comprehension scores increased equally for both proficiency levels in the experimental groups (G_1 vs. G_2). (Figure 3-3.)

3. Listening scores improved for one experimental group more than the other following reading practice. That is, proficiency level moderated the effect of reading practice. There is an *interaction* of treatment and proficiency levels which can be related to improvement. (G_1 vs. G_2). (Figure 3-4.)

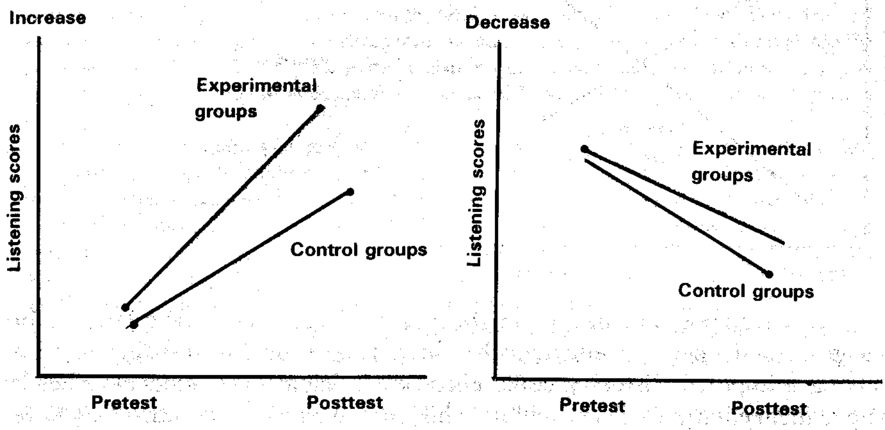


Figure 3-2

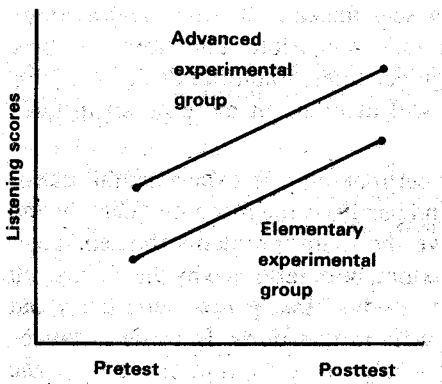


Figure 3-3

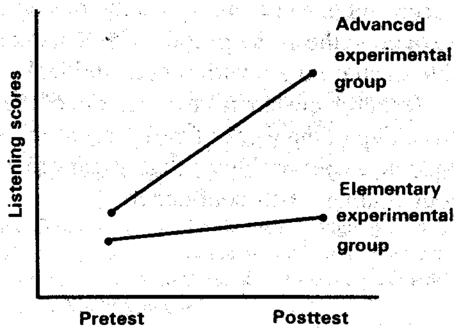


Figure 3-4

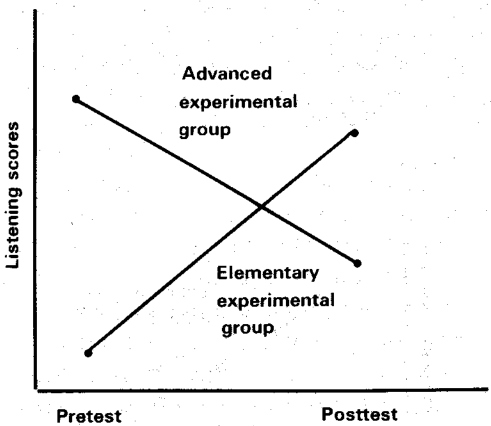


Figure 3-5

4. It is even possible that you might get a positive effect of the treatment for one group and a negative effect for the other. That is, massive reading practice may have improved listening comprehension for elementary students since it helped them enlarge their vocabulary while massive reading practice may have made advanced students feel that only reading was important so that they largely ignored oral language (or some other less farfetched reasons you may think of). In this case there would also be an interaction between the two variables (see Figure 3-5). As you can see, the introduction of moderator variables makes a wider range of interpretations possible. The more variables you add, the more complicated the design becomes. Factorial designs require fairly sophisticated analyses which we will examine in more detail in later chapters.

The choice of design, whether pre-experimental, true experimental, quasi-experimental, or ex post facto, will depend partly on the research question and partly on your ability to meet and solve the many problems that endanger validity of research. It will also, to some extent, be determined by the claims you hope to make at the conclusion of your study. Most people who carry out research do so in order to arrive at answers to questions. In some instances researchers wish to be able to arrive at a careful description of the language learning of one person. In others, they hope to generalize from a sample group of learners to the entire population. Whatever the findings, researchers want to be able to share them with others and feel some degree of confidence in doing so.

Research design has evolved out of the need to solve the many problems that turned up in the work of early researchers. If you plan your research carefully, you will avoid problems that might otherwise make it difficult for you to share your findings with confidence.

ACTIVITIES

Note: In the following items, assume *Ss* were chosen and assigned to groups on a random basis.

1. In an introductory linguistics class which you give for TESL teachers, you think *Ss* would be able to identify the articulators better if they use a new method. So you give half the class xeroxed copies of appropriate pages from the *Gray's Anatomy Coloring Book.* The other half gets xeroxed copies of appropriate pages from a regular description in a linguistics text. On the midterm, you give an "identify the parts" question which taps the studied information. If you want to check the results, what design have you used? Draw the schematic representation and identify it.
2. As a needs assessment, you give all the students in your pronunciation class the Prator Diagnostic Passage. In addition to the regular pronunciation lessons, half the students get a teacher-trainee as their friendly tutor outside class. The other half couldn't arrange a time to meet with the tutors. At the end of the course, you give another test using the diagnostic passage again. If you want to say something about whether to continue the tutorial system or not, what design could you use to look at student gains? Draw the schematic representation and name the design. What other factors would you consider in making your recommendations?
3. Let's change question 2. You still have a pronunciation class and you still have tutors. The tutors come to class for half an hour on Thursdays and your class meets on Tuesdays and Thursdays. You give the same lessons on pronunciation as before. You give a short test on each item covered at the end of every session. If you wanted to say something about tutor effectiveness, what design could you use? Remember that this time all students get a tutor on Thursdays for part of the session. Draw the schematic representation and name the design.
4. In our relative clause experiment, assume that we have identified three relative clause types (subject focus, object focus, and possessive focus) and have given a test on these types to two groups of ESL students, Japanese and Arabic students. Give a schematic representation of the design and identify it.
5. You've been teaching a special section of ESL for engineers. At the end of the program, you administer a test to see how well they can identify research hypotheses, experimental design, and statistical procedures in scientific texts. A friend of yours who teaches TEFL courses overseas asks if he can borrow your test to see how his students do on it since they are pre-science students who will receive all their university instruction in English. You agree. If he sends you the results, can you use his results as a control group for your research? (If no, why not?) Assume that you could use them as a control group. What research design would you use? Draw the representation and name the design.
6. You went to the census bureau in your county and got all the listings for non-English-speaking households (English not the major language of the home). From these you randomly select 200. You contacted these, and 80 agreed to answer 10 questions on language maintenance. You employ two research assistants, and between them they speak all the languages in your sample. (They are also absolutely identical twins so we can forget about the interviewer variable.) You randomly assign 40 families to be interviewed in the home language and 40 families to be interviewed in English. If you want to say something about the importance of the language used in interviews, what is your design? Sketch it, please.
7. Many of your adult school ESL students seem to like to use the Language Master for vocabulary study. They see the word, hear the word on the tape, a picture is given, and they have ample opportunity to say the word. You decide you want to present the technique at the next district workshop. What kind of research could you do that would show whether it's effective? What design would you use?
8. You want to find out whether foreign language pronunciation could be helped through hypnosis. You have a group of wealthy tourists who are about to go on a tour of the Soviet Union. They are all taking Russian lessons to prepare themselves for the tour. What kind of research could you do to see whether hypnosis might help? What design would you use?

9. Many people believe that beginning second language learners should be given as much speaking practice as possible. Others believe that it is better to delay oral practice until the learner has developed listening skills. You are teaching English as a Foreign Language to children in two fourth grade classes in Greece. What kind of research could you do to see whether delayed oral practice is better? What design would you use? List two problems that might make others question your results.

Suggested further reading for this chapter: Campbell and Stanley, Isaac and Michael, and Tuckman.