

## CHAPTER 7

# Experimental and Nonexperimental Research Designs

### Learning Objectives:

- What are the three forms of empirical control in scientific research (physical, assignment, and statistical), and how are they related to the distinction between experimental and nonexperimental research?
- What does the distinction between laboratory and field settings mean?
- What is the difference between within-case and between-case research designs, and what are its implications for research?
- What are developmental research designs, and how do longitudinal and cross-sectional designs differ?
- What are some implications of the distinction between single- and multiple-case designs?
- How does computational modeling compare to traditional experimental research, in terms of how variables are controlled and how data are generated and analyzed?

In any research study, variables are created or measured in ways that allow particular comparisons to be made and not others. The structure of possible comparisons in a study that results from the way variables are created or measured is the issue of **research design**. Research design largely determines which questions can be asked of data and greatly influences the validity with which those questions can be answered, as we discuss below and in Chapter 11.

## Empirical Control in Research

---

The distinction between variables that are “created” versus those that are “measured” suggests what may be the central distinction in research design, that between **experimental** and **nonexperimental studies**. As we discuss in detail below, at least one variable is created in experimental studies, whereas variables are only measured in nonexperimental studies. To understand this central distinction, let’s return to our list of the goals of a science from Chapter 1. One of those goals was explanation, which we described as the explication of causal relations among entities and events, an explication that accounts for patterns within data. Scientific researchers want to be able to explain patterns in their data, and they want to be able to do it validly. In order to increase their ability to draw valid causal conclusions, researchers exercise empirical control over the phenomena they study. **Empirical control** is any method of increasing the ability to infer causality from empirical data<sup>1</sup> (not to be confused with the fourth scientific goal from Chapter 1 of exercising practical or material “control” over phenomena).

Empirical control is exercised in one of three ways, or a combination thereof. The first is **physical control** (mentioned in Chapter 4). This is physically controlling the data collection situation in order to reduce the potentially distracting influence on our data patterns of factors that are not of interest to us. Some examples of physical control include isolating cases, simplifying or purifying tasks or procedures, and increasing the consistency of the research situation across data collection episodes. Although some social and behavioral scientists employ physical control in their research, it is more common in the biophysical sciences, particularly laboratory physical sciences like chemistry; that’s the reason they wash chemical beakers so carefully.

The second type of control is **assignment control**. Researchers using assignment control *create* at least one of the variables in the study in order to test ideas about its effect on some other measured variable. The created variable is usually at the nominal level, so that each value or level of the created variable is a discrete experimental **condition** to which cases are exposed that are assigned to it. Researchers incorporate assignment control in their studies when they determine which cases to assign to which conditions of the study. For example, a researcher studying soil erosion could create five values of a variable called *Solvent pH*, which concerns how aqueous solutions differing in pH cause different rates of erosion in particular soil types. Values of this variable could be “pH 5,” “pH 6,” “pH 7,” “pH 8,” and “pH 9” (neutral and a little acidic or basic). The study involves assignment control because the researcher decides which soil material (the case) would be assigned to be dissolved with which solvent solution. Alternatively, assignment control takes place when the same cases are subjected to all of the conditions of the study in an order determined by the researcher, defined below as a “within-case design.” For example,

---

<sup>1</sup>Empirical control often increases statistical power and precision (Chapters 8 and 9) as well.

a researcher studying GIS interface design could expose samples of GIS users to two systems that use either verbal labels or iconic symbols to represent particular GIS operations; all users would be exposed to interfaces with both symbol types but in an order controlled by the researcher. Either way of controlling the assignment of cases to a variable is known as **manipulation**, even though they do not involve physical control, as might be implied by that term. Manipulation is most often done by **random assignment** to conditions, because that is a simple and straightforward technique that leaves the manipulated variable uncorrelated with all the non-manipulated (measured) variables, on average. Variations on random assignment include **matching**, in which two cases that have some characteristic in common are assigned to contrasting conditions.

The third and final way to exercise empirical control is **statistical control**. In studies employing statistical control, the researcher measures, but does not create, variables in the study. But the researcher takes explicit steps to identify, measure, and statistically analyze any variable that he or she thinks might have an effect on the main variables of interest in the study. For example, a geographer studying land-use changes might want to focus on the possible effect of changing family structure, such as the addition of more children, on the transition from one type of agriculture to another. To examine the effect of family structure without being distracted by the possible effect of, say, earnings from off-farm employment, the geographer could make sure to measure earnings and enter them into a regression equation along with the family-structure variable. Any effect of family structure would then be over and above the effect of off-farm earnings.

Having described the three forms of empirical control (physical, assignment, statistical), we can return to the important distinction between experimental and nonexperimental research designs. The term “experiment” is often used colloquially to refer to any scientific research study. Technically, however, it has come to refer specifically to studies that involve the manipulation of one or more variables—assignment control, in other words. This technical use of the term “experiment” is reminiscent of Hume’s third principle of causality we discussed in Chapter 2: *Controlling the cause will control the effect*. Nonexperimental studies may involve physical and/or statistical control, but they do not involve manipulated variables. We absolutely do not mean to imply that nonexperimental studies are necessarily less “scientific,” however. In fact, geographic researchers in many topic areas, including many physical geographers, rarely or never conduct experimental studies. That is simply because some constructs cannot easily be manipulated, if at all. It is impossible for a researcher to assign mountain chains to levels of a variable expressing orientation to dominant wind patterns; it is similarly impossible to assign cities to different base ratios of economic activity. At the end of this chapter, we consider an approach to research design that emerged in the 20th century as a very promising technique to expand the domain of cases and constructs that can be studied experimentally—computational modeling.

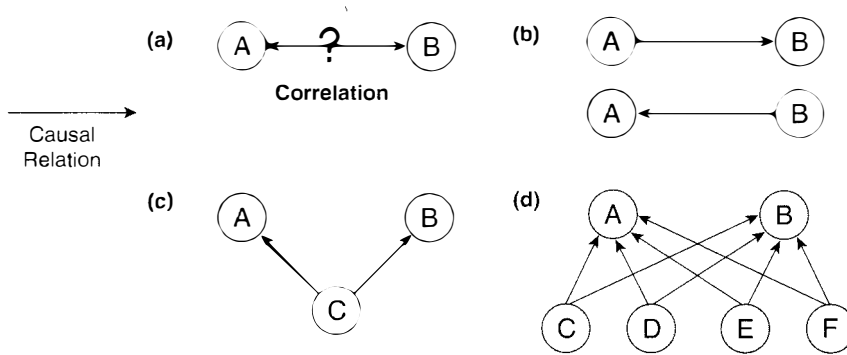
So all true experiments have one or more manipulated variables and one or more nonmanipulated variables. The manipulated variable is “created” by the researcher, who may accurately be called the “experimenter.” The distribution of values of the manipulated variable across cases is independent of what the cases do or the properties they have, so it is called an **independent variable** (IV or “factor”). The nonmanipulated variable is only measured by the researcher. Its distribution of values across cases depends at least partially on what the cases do or the properties they have, so it is called a **dependent variable** (DV). The IV is the potential *causal* variable of interest in an experiment, and the DV is the potential *effect* variable of interest. Strictly speaking, it is misleading to refer to variables in nonexperimental studies as independent or dependent variables (in a sense, all are dependent), even though one or more variables in such a study is often considered a potential cause and others are considered potential effects. It is more correct in such a study to refer instead to “predictor” and “criterion” variables, as we will in Chapter 9 when we discuss *X* and *Y* variables in regression analyses.

The experimental-nonexperimental distinction is so important to research design because of its considerable implications for our ability to infer causality from empirical data. Experimental manipulation makes it logically much easier to establish that the IV is the variable responsible for patterns of variation in the DV, as opposed to some other causal variable being responsible. These “alternative” causal variables may not even have been thought about by the researcher, let alone measured or manipulated. To understand this rather fascinating aspect of scientific logic, consider that all empirical attempts to establish causality in science involve finding correlations between putative cause variables and effect variables. That is, variable *A* can be a cause of variable *B* only if the presence of *A* is always or usually accompanied by the presence of *B*, and the absence of *A* is always or usually accompanied by the absence of *B*.<sup>2</sup> However, even though finding patterns of correlation between variables is necessary to establish causality, it is not sufficient. That’s because two variables can be correlated without one directly causing the other, as Figure 7.1 shows. Perhaps a third variable might directly cause both *A* and *B*, which in turn causes *A* and *B* to be correlated. This “third variable” could be a single variable *C* or a larger set of variables of any size or complexity. Even if the correlation between *A* and *B* is due to direct causality, the presence of the correlation itself does not establish whether *A* caused *B* or *B* caused *A*. Thus, the well-known and important adage that “correlation is not causality” is more correctly stated this way: *Correlation is causality, but the specific pattern of that causality is ambiguous.*

So manipulation makes it logically more likely that the variable we think is the cause of correlations in our data really is the cause, although it does not absolutely

---

<sup>2</sup>Philosophers such as Francis Bacon in the 17th century and J. S. Mill in the 19th century referred to the method of establishing causality via manipulation as the “Method of Concomitant Variation.”



**Figure 7.1** Alternative causal patterns that can explain an empirically observed correlation between variables *A* and *B*, shown in (a). The correlation may come about because *A* causes *B* or *B* causes *A* (b), because a third variable *C* causes both *A* and *B* (c), or because a set of any number of third variables that are interrelated to any degree of complexity causes both *A* and *B* (d). Without additional information, the underlying pattern of causal interrelations that explains the observed correlation is completely ambiguous.

ensure this. When one variable consistently changes value in response to changes in another variable that we, the researchers, have brought about, it becomes very unlikely—too much of a coincidence—that some other third variable could really be the cause of those changes. In nonexperimental studies, the chance that some third variable, or **confound**, is the actual cause of correlations in our data is usually quite substantial. In some specific research designs, as we discuss more below, the risk of a confound causing the correlations can be very high. In that case, the observed correlation between the two variables of interest is not the result of a direct causal relation between them; instead, the apparent direct causality is **spurious**. An example of attributing spurious causality to a correlation would be mistaking a proxy measure of past climate for the ultimate cause of the climate (see Chapter 4).

In the minds of many scientists, the likely existence of a confound may be the single most damning flaw in a research study because it sheds such doubt on the validity of causal conclusions (we learn in Chapter 11 that the validity of causal conclusions is called “internal validity”). So if you are interested in drawing causal conclusions when you are planning a study, thinking about possible confounds is critical. And the existence of confounds in other researchers’ studies is one of the central issues you should consider when critically interpreting them, as when you review for a journal. It is vital to note, however, that all possible confounds are *not* necessarily important threats to the validity of causal conclusions. To be an important threat, both of the following circumstances must hold:

- (a) The potential confounding variable must vary differentially across values (conditions, in an experiment) of the causal variable of interest.
- (b) Even if the potential confound does vary in this way, it must also be related to the effect variable of interest.

In fact, every study in geography, experimental or otherwise, even those done very well, involves any number of potential confounds. But potential is not necessarily actual. For example, in our study of soil erosion described above, one could argue that it was temperature rather than the solvent pH that caused differential erosion. This critique is invalid unless (a) the temperature of the solvent solution varied sufficiently across pH conditions *and* (b) solvent temperature actually does correlate with the rate of soil dissolution. In this example, (b) is generally true but (a) would probably not have been true—the researcher would likely have used physical control to make sure the solvents were applied at the same temperature. Our general point is that when you critique the research of others, it behooves you to express concern for potential confounds only if they satisfy circumstances (a) and (b).

This is a good place to discuss a concept most people have heard of but may misunderstand a little—the concept of the **control group** (control condition). This is one or more experimental conditions that are added to an experimental design in order to allow particular comparisons that are logically relevant to the conclusions of an experiment. Many people think of the control group as a special experimental condition in which a certain “treatment” manipulation is withheld from the cases. For instance, most people know that medical researchers cannot establish the effectiveness of a new drug unless its use is compared to a condition where it is not used. This is correct, but we think it is an overly restrictive way to think about control groups. We don’t use the term much ourselves because we know that all experiments require the necessary conditions to support whatever comparisons you want to make. The control group is really no different than any other condition in this respect. As a matter of fact, the comparisons you want to make in an experiment may or may not even involve some type of “empty” control condition. Medical researchers learned decades ago that a necessary comparison condition for drug trials is the “placebo” condition (an inert pill, for example), not the “no-pill” condition.

We’re also in good position now to make sure we clearly understand the distinction between “random sampling” and “random assignment.” They really don’t have much to do with each other. Random sampling, as we discuss in Chapter 8, refers to selecting cases from a sampling frame to be in a sample. It influences the validity of the generalizations we draw from samples to sampling frames, including generalizations about statistical patterns that have nothing directly to do with causality. Random assignment refers to determining which cases in an experiment to expose to which condition or sequence of conditions. It influences the validity of causal conclusions, as we discussed at length above. In other words, the only similarity between random sampling and random assignment is their randomness. It’s not only possible to have either one without the other, it’s actually quite common. Researchers who work in disciplines or topical areas that stress one type of validity tend not to stress the other type. For example, political pollsters put a lot of energy

into successful random sampling, but they usually do not assign cases to conditions at all because their studies are not experiments.

## Laboratory vs. Field (Naturalistic) Settings

In Chapter 4, we mentioned a distinction between lab and field settings for conducting research studies. To many people, an “experiment” conjures up an image of a white-coated scientist working away in a sterile and technical-looking room known as a laboratory. Aside from the fact that even those geographers who *do* conduct experiments in labs usually don’t wear white lab coats, this stereotype is misleading in a more substantial way. It conflates the distinction between experimental and nonexperimental designs with the distinction between laboratory and field settings. A **laboratory**, or lab, is a specially designed setting, nearly always an interior room in a building, that allows researchers to exert physical control while conducting studies on their phenomenon of interest. Both physical and human geographers collect data in laboratories. Temperature, humidity, and chemical contamination can be controlled; so can sounds, sights, and other distractions. The **field** setting, in contrast, is also a setting where geographers conduct studies on their phenomenon of interest, but it is the setting in which the phenomenon normally occurs. Field settings are essentially “naturalistic” settings, by which we mean they are data collection locales that are the places where phenomena of interest go on as they normally do. Given the interests of most geographers, data collection in the field is more common than in the lab.<sup>3</sup> Field settings are usually outdoors, but they need not literally be *fields* (although they are to many biogeographers and agricultural geographers!). Again, both physical and human geographers conduct studies in the field—at forests, seashores, drainage basins, glaciers, cloud formations, and eroded mountains; also at rural villages, central business districts, front yards, manufacturing plants, legislative sessions, and national parks.

We discuss the concept of validity at length in Chapter 11. The issue of experimental versus nonexperimental designs has implications for the validity of causal conclusions, as we discussed above. The issue of laboratory versus field settings has implications for the validity of conclusions we make about how our research findings generalize to other settings, cases, measures, and so on. In Chapter 11, we learn that this is the issue of “ecological validity,” which is the natural verisimilitude of the research setting. Our point here is that the issues of experimental versus nonexperimental and lab versus field are mostly independent from one another. A researcher can do a lab experiment, a lab study, a field experiment, or a field study. Perhaps field experiments are even superior in some research domains. For example, the surveys we discussed in Chapter 6 are very common and flexible ways of collecting

---

<sup>3</sup>Most geographers think of field research as requiring the researcher to actually travel to the setting where data are collected. So even though satellites collect remotely sensed data on phenomena in their naturally occurring setting, most geographers would not consider studies using such data to be field research because the researcher sits in an office at a computer terminal in order to acquire data.

data in nonexperimental research, but they are used in experimental research too; the experimenter may decide which respondent gets which version of the survey. Therefore, it is not really correct to contrast “survey research” with experimental research. For that matter, researchers may administer surveys and other explicit reports in a controlled lab setting *or* in a field setting.

## Basic Research Designs

---

We can now turn to an overview of some of the basic alternatives of research design, the choices we have available for how we design variables and set up measurements to allow us to make certain comparisons. As we noted above, our design choices go a long way toward determining which questions we can ask of data and the nature of the truth we can expect from our answers.

First, we should make sure we are clear about the distinction between variables and **levels of variables**. As we defined them in Chapter 2, variables are the attributes or properties of cases that vary, either across cases at the same time, within cases over time, or both. A study must have one or more variables. A study with only one variable is rather unusual, as it is only able to consider the distribution of that single variable across cases. Studies generally have at least two variables, allowing for the investigation of relationships, and they usually have more than two. In any event, a basic question of research design concerns how many variables to measure and/or manipulate. This still leaves the question of how many values each variable will take in a study. We know that two is a minimum—it wouldn’t be a *variable* otherwise. The different values a variable takes (or is designed to take by an experimenter) are different levels of the variable. In experiments, they are conditions; with any discrete variable, these levels might be called “groups” or “classes” (for example, groups of two types of drainage basins). With nonmanipulated variables, this question of levels is in part a question of measurement resolution (precision).

A second basic issue concerns whether levels of the variables differ between cases or within cases. Above, we contrasted experimental designs where each case gets exposed to one and only one condition of the independent variable with designs where each case gets exposed to all conditions of the independent variable. (Intermediate designs in which cases are exposed to more than one but less than all conditions are possible too, but they lead to complexities of analysis and interpretation and should usually be avoided.) This distinction holds for nonexperimental designs, too; a study of biota in limestone caves versus biota in sandstone caves is different than a study of biota in the same limestone caves over time. Whether experimental or nonexperimental, the first type of design is a **between-case design**. Cases are at different levels of an independent or predictor variable, whether placed there by an experimenter or not, so that analyzing these variables involves comparing data *between* different cases. The second type of design is a **within-case design**, also known as a repeated-measures design. Cases are at all levels of a variable at different times, again in an order determined by an experimenter or not, so that analyzing the variable involves comparing data *within* the same cases.



Both classes of designs have their strengths and weaknesses. Generally, within-case designs

- (a) are more efficient, as fewer cases are necessary to get the same amount of data;
- (b) lead to higher precision of estimation and power of hypothesis testing (Chapter 9), even given the same amount of data, because random noise associated with irrelevant variations among cases is reduced (measurements of a city or mountain made at one time are usually more similar to measurements of that city or mountain made at another time than they are to measurements of another city or mountain at the same time); and
- (c) reduce confounds because comparison groups are more nearly equated, for the same reason as in the previous point.

These are certainly enormous benefits, so enormous that you should always do within-case designs if you can. Unfortunately, you cannot. Within-case designs have some severe limitations. They always employ multiple measurements of the same cases and, in experimental designs, multiple exposures to different conditions of the independent variable (IV). Therefore, they often run the risk of changing the cases somehow as the study goes on, including producing order effects, such as those discussed in Chapter 6 that result from the administration of multiple items in explicit reports (fatigue, practice effects, carryover effects, and so on). These can often but not always be addressed by randomizing or counterbalancing the orders of conditions (see Chapter 6). But within-case designs have a much more serious problem than order effects, really an ultimate limitation. Often they simply cannot be done. Many variables fundamentally reflect inherent properties of different cases that can hardly if at all be made to exist within the same case over time. A city in Europe cannot become a city in Africa. Saltwater fish cannot become freshwater fish. Residents of Chinese descent cannot become residents of Korean descent. What's more, measurement sometimes destroys cases, which absolutely precludes a within-case design. A study of insect migration that involves capturing and killing insects is one example. It is an example that suggests that ethical considerations too can preclude the use of within-case designs. We return to this in Chapter 14, and we leave it to you here to entertain additional ethically relevant scenarios of this type.

## Specific Research Designs

Table 7.1 presents several specific research designs, both nonexperimental and experimental. As we pointed out above, research designs such as these vary in the validity of causal conclusions they support. Some designs are so poor in this respect that they rarely if ever deserve to be considered for use. Others are relatively weak in this respect but deserve consideration because they are less costly to carry out, in all senses of cost. And of course, we made it clear above that experimental designs are simply impossible in some research areas.

Table 7.1 Assorted Research Designs\*

<b>Nonexperimental Designs</b>							
One-Group, Single Measurement	O <sub>1</sub>						
One-Group, Multiple Measurement	O <sub>1</sub>	O <sub>1</sub>	O <sub>1</sub>	O <sub>1</sub>			
One-Group, Posttest-Only	E	O <sub>1</sub>					
One-Group, Pretest-Posttest	O <sub>1</sub>	E	O <sub>1</sub>				
One-Group, Multiple Pretest-Posttest	O <sub>1</sub>	O <sub>1</sub>	E	O <sub>1</sub>	O <sub>1</sub>	O <sub>1</sub>	O <sub>1</sub>
Two-Group Nonmanipulated, Single Measurement	O <sub>1</sub> O <sub>2</sub>						
<b>Experimental Designs</b>							
One-Group Manipulated Within	M <sub>1</sub>	O <sub>1</sub>	M <sub>2</sub>	O <sub>1</sub>			
Two-Group Manipulated Within	M <sub>1</sub>	O <sub>1</sub>	M <sub>2</sub>	O <sub>1</sub>			
	M <sub>2</sub>	O <sub>2</sub>	M <sub>1</sub>	O <sub>2</sub>			
Two-Group Manipulated Between, Posttest-Only	M <sub>1</sub>	O <sub>1</sub>					
	M <sub>2</sub>	O <sub>2</sub>					
Factorial Four-Group	M <sub>1</sub> N <sub>1</sub>	O <sub>1</sub>					
Manipulated, Posttest-Only	M <sub>1</sub> N <sub>2</sub>	O <sub>2</sub>					
	M <sub>2</sub> N <sub>1</sub>	O <sub>3</sub>					
	M <sub>2</sub> N <sub>2</sub>	O <sub>4</sub>					

\*“O” is an observation or measurement, “E” is an event that naturally occurs, and “M” and “N” are applications of manipulations. Subscripts on “O” indicate which independent group of cases is observed; numerical labels on “M” and “N” indicate conditions of the manipulated factors.

Turning to nonexperimental designs first, we see that the simplest possible design consists of a single measurement on a single group of cases. This design is useful only if you want to infer the absolute level of some variable in some population—no comparison among groups of cases, no evidence of the effects of an event or manipulation, no contrast among different variables. If you repeat the measurement several times, you at least get descriptive evidence for change over time that might suggest something worth investigating with a more sophisticated design. Next, we can take the opportunity to examine the possible effect of a naturally occurring event (that is, not manipulated by the researcher), for example, a volcanic eruption or the enactment of new legislation. A **posttest-only design** is quite deficient, however (“pre-” and “post-” mean before or after the event or manipulation). A single measurement after the event leaves you wondering what the measurement would have been before the event. A **pretest-posttest design** at least gives you a baseline measurement for comparison, although the design still leaves open whether some ongoing trend or some other event, rather than your event of interest, caused the difference from pre- to post-. We get much better

evidence for the effect of a naturally occurring event if we take multiple measurements over time, preferably both before *and* after the event.<sup>4</sup>

If we want to examine the possible effect of a variable that is intrinsically part of our cases, rather than an externally occurring event, we can conduct a between-case study by sampling from two identifiable subpopulations of cases. This allows us to identify relationships between group membership and the measured variable, but as we know from Figure 7.1, that alone leaves causality quite ambiguous. We show only two groups in Table 7.1, but of course any number of groups can be sampled. Finally, note that all designs, whether nonexperimental or experimental, can involve measurement of more than one observed variable. This has the obvious benefits not only of increasing the number of variables you can make conclusions about but allowing comparisons of patterns across variables that may shed light on whether your event or manipulation is the actual cause of changes to any of your measured variables.

Turning to experimental designs, we see that the simplest within-case design exposes a single group of cases to one condition of a factor and a measurement, and then exposes them to a different condition and another measurement. This design is considerably strengthened if we assign our cases into two groups, typically randomly, by exposing each group to one of the two possible orders of the two conditions of the factor. That way, we'll have evidence about the possible effects of being exposed first to one of the two conditions before being exposed to the other. If we want (or need) to conduct a between-case experiment, our simplest option is to assign our cases into two groups, exposing each to different conditions of the manipulated factor, and then measuring them. You can design something a little more complex by manipulating two (or more) factors. In the simplest possible such **factorial design**, the cases would be assigned into one of four conditions formed by crossing the two conditions of one factor against the two conditions of the other (this is a " $2 \times 2$  design"). For example, we could conduct an experiment on the effectiveness of cartographic "in-vehicle navigation systems" in automobiles as a function of their orientation flexibility and information content. Starting with a sample of 120 drivers, we could assign 60 each to receive map displays on a dashboard computer oriented in one of two ways: a fixed orientation with north up or a variable orientation that turns as the car turns so that forward stays up. Within each group of 60, we could assign 30 to receive maps containing labeled landmark features, whereas the other 30 could receive maps without labeled landmarks. The

---

<sup>4</sup>Designs that incorporate a large number of repeated measurements over time, to investigate either the effect of some event or just a normally occurring developmental trend, are called **time-series designs**. These are fairly common in geography, but their statistical analysis is special in ways rather analogous to the analysis of data distributed over space (Chapter 9). The presence of different patterns of **temporal autocorrelation**, wherein values of variables are more or less similar as a function of when they were measured, must be evaluated in the data. For example, many variables show evidence of "seasonality" by taking on characteristics patterns of high or low values at different times of the year (including temperature, precipitation, burglaries, and tourist activity).

value of factorial designs goes well beyond simply being able to test multiple factors all at once. Such designs also allow for the investigation of factorial **interactions**, wherein the effects of one factor depend on the condition of the other factor; more complex interactions involving more than two factors are also possible. In our example, one possible interaction pattern would be if the labeled landmarks proved useful to drivers when the map kept a fixed north-up orientation but not when the map turned as the car turned.

There are an unlimited number of possible research designs that go beyond the basic ones we introduce here. An important class of designs is called **quasi-experiments**, which are studies without manipulated variables that nevertheless attempt to establish causal relations by applying systematic statistical control via more complex research designs. As we suggested above, the number of variables in a study, both manipulated and measured, can be increased *ad infinitum*. The number of levels, or the measurement resolution, of these variables can also be increased to any level you might want. However, not only do research designs of increasing complexity cost more and more (including requiring more cases and more measurements), they also potentially lead to interpretative difficulties that eventually exceed the powers of our limited human minds.

## Developmental Designs (Change over Time)

Geographers and other natural and social scientists are often interested in processes of change over time. How do beaches accrete and erode over time? How do patterns of migration, both human and nonhuman, shift over time? Systematic (nonrandom) processes of change like these are examples of “development.” There are many specific examples of systems in human and physical geography that develop, and many processes responsible for the course of those developments. The planet’s solid surface develops, cultures develop, atmospheric composition develops, cities develop, individual people develop, and ecosystems develop. Broadly speaking, “evolution” is another term for development. Understood in this broad sense, Charles Darwin’s ideas make up just one of many theories of “evolution,” although of course it is an especially noteworthy theory about the genotypic and phenotypic development of species over generations.

**Developmental designs** are studies designed to conduct research on developmental processes. There are two basic approaches to the design of developmental research studies. In the first approach, two or more groups of cases, each at different “ages” or levels of development, are compared at the same time. For example, different forests at different levels of succession or the economies of countries at different stages of economic maturity are compared. This is the **cross-sectional** or “synchronic” approach. Its biggest advantage is its relative ease and efficiency. However, such a design provides no direct evidence on development, only the indirect evidence of comparing the product of development as realized in two or more static groups of cases at two or more developmental ages. These groups of cases of the same age—that is, “born” at the same time—are called **cohorts**. Cross-sectional designs always compare two or more cohorts of cases. A specific cohort sometimes has characteristics that result from the particular generation of which it is a part,

rather than characteristics that have always described or will always describe cases at that age. Such **cohort effects**, wherein properties of cases originating at the same time are mistaken for properties of cases of a certain level of development, are threats to the internal validity (Chapter 11) of cross-sectional designs.

The second approach to developmental studies is the **longitudinal** or “diachronic” design. A group of cases at one level of development is compared to itself over time. By definition, such a design requires repeated measurements. For example, the geographic knowledge of children can be studied by testing a group of children when they are five years of age, the following year when they are six, and a third year when they are seven. As another example, the location and size of a streambed can be compared over a 10-year period by taking measurements every year at the same locations of the same stream. The biggest advantage of longitudinal designs, over and above the other benefits of within-case designs discussed above, is that they provide direct evidence on development—the cases actually change over time during the course of the study. Unfortunately, they are typically quite time-consuming and expensive. Such designs also suffer from various forms of **attrition** or “mortality,” the loss of cases during the course of the study; researchers and research support can also quit or “pass away” during the course of the study. Attrition is especially problematic to internal validity when different types of cases are systematically more likely to drop out of the study than others, a condition known as “differential attrition.” For example, a study of spatial behavior over time in which tracking devices are attached to people’s cars might suffer from the fact that people going to particular types of places rather than others might be more likely to quit the study before it is finished (we return to this in Chapter 14, but please use your imagination for now). Longitudinal designs can also suffer from **history effects**, idiosyncratic events or conditions that hold during the particular time period of the study. Finally, many longitudinal studies that are conducted over several years or, especially, several decades potentially suffer from **instrumentation**, which is a change over time in the way measurements are made. A classic example is the way questions on the U.S. census are modified over the decades, as we discussed in Chapter 6.

A hybrid approach to developmental designs is known as a **sequential design**. There are many specific variants of the sequential design, but they all combine aspects of cross-sectional and longitudinal designs. Two or more groups of cases differing in their level of development at one time are sampled, as in cross-sectional designs, but these groups are measured repeatedly over time, as in longitudinal designs. The sequential design provides some of the advantages of each approach while mostly avoiding some of their threats to validity, such as cohort and history effects.

A final word about developmental designs concerns temporal scale. As we pointed out in Chapter 2, temporal scale is important to geographers, as is spatial scale. In developmental designs, both time and temporality are critically important with respect to phenomenon scale and analysis scale. What should the duration of the study be, and at what time should data collection begin and end? What should the interval between measurements be, which, when combined with the duration of the study, determines the total number of measurements to be made? What is the temporal form of the developmental process being studied within the time frame

of the study? Is it a straight line going up or down? Is it “monotonic” but nonlinear (see Chapter 9), with most of the change occurring near the beginning of a study’s time frame? Is it a “U-shaped function” that comes down and then goes up, or an “inverted U-shaped function” that goes the other way? These questions must be addressed when designing and conducting developmental research.

## Single-Case and Multiple-Case Designs

Much of what we have said about research studies above and in previous chapters has explicitly or implicitly assumed that a study involves observations of several cases rather than just one. This is not necessarily true. There are research designs, both experimental and nonexperimental, that involve just a single case, such as a single location in a river or a single neighborhood in a city. A **single-case experiment** is a repeated-measures design with one case. This is a particularly effective design when one is interested in demonstrating the practical ability to change the value of a single case on some variable at will. For example, you could measure a case at one time, then expose it to some condition and measure it again. That would not be a very strong design because any change in the measured variable could result from some other event that occurs at about the same time or some preexisting trend that was already ongoing. An improvement would be the **reversal design** (“return to baseline”), wherein, after the second measurement, you would remove the applied condition and measure a third time. More complex designs are possible.

However, *nonexperimental* single-case designs are much more common in geography. A **case study** is an intensive and comprehensive descriptive study of a single case. Case studies can involve any number of data collection types, including any mixture of quantitative and qualitative approaches, although in human geography, they virtually always involve an emphasis on qualitative approaches. Case studies can provide a very rich, holistic, and wide-ranging description of a single case, but they are at best *suggestive* about causality rather than definitive. To some researchers, therefore, they are considered most useful in the early, exploratory stage of a research program, or as adjuncts to multiple-case designs that help make general conclusions more concrete and personal (much the way so many newspaper stories these days are introduced by an anecdote about a single person or other entity).

As compared to multiple-case designs, single-case designs are efficient (only one case to obtain and measure) and provide a rich, more complete picture of the characteristics of a meaningful unit. But multiple-case designs have some very definite and important strengths as compared to single-case designs. Using multiple cases obviously gives you much more ability to explore how findings generalize to various types of cases. It goes a long way toward helping you to avoid spurious conclusions drawn from idiosyncratic cases you might have happened to choose in a single-case design. And given the statistical nature of most phenomena in geography (see Chapter 9), only the multiple-case design allows you to see the effect of a single variable when many actually play a causal role in your system of interest; it allows you to find a “signal” in a background of “noise.”

These benefits of multiple-case designs really matter a lot when researchers accept the goals and characteristic philosophical values of a scientific approach to

research. We went over these in Chapter 1 and introduced the distinction between **nomothetic** and **idiographic** approaches to knowledge. A nomothetic approach attempts to be general over cases, times, places, and so on; nomothetic approaches attempt for a “lawlike” understanding or at least an understanding based on probabilistic general truths. An idiographic approach attempts to be specific to particular cases at particular times and places; idiographic approaches strive for a potentially “idiosyncratic” understanding. We noted in Chapter 1 that scientists strongly prefer a general understanding. An approach that is exclusively idiographic is therefore fundamentally nonscientific, whatever truth it may achieve. Don’t overinterpret this; scientists do not stubbornly refuse all nongeneral or conditional truths just to maintain some status as “real scientists.” In fact, a completely nomothetic, general understanding is apparently not possible—it does not fit the facts of reality as we can appreciate them. A search for “*The Law of Nature*,” a single unifying force, has reputedly occupied physics at its highest conceptual levels for some time, but it seems safe to conclude that no such “über law” will ever characterize any area of research in geography, nor in many other natural or social sciences. Reality seems to lie somewhere between extremely nomothetic and extremely idiographic approaches. In truth, different research approaches tend to take positions somewhere on a continuum between the two extremes. Many hybrid and intermediate approaches attempt to take advantage of strengths of both, for example, a case-study design replicated on several cases. To recognize that some truths generally hold, for example, in certain climates and not others, or with certain cultures and not others, is quite acceptable as a scientific conclusion.

## Computational Modeling

---

In Chapters 2 and 3, we introduced and defined models as simplified representations of portions of reality that include sets of interrelated hypotheses expressing theories about structures and processes of systems of interest in the world, including causal processes. We learned in those chapters that models are pervasive in all areas of geography and in other sciences, especially considering that they come in several forms, including conceptual, physical, graphical, and computational forms. Our interest here is in computational models, defined in Chapter 3 as models of theoretical structures and processes expressed in mathematical form (more precisely, in a **formal language**). Computational models are typically instantiated as sets of equations and other logical/mathematical operations expressed in a computer program or set of programs and attendant databases. However, there is nothing about computational models that intrinsically requires computer representation other than the practical difficulty or impossibility of doing them manually. There are many specific applications of computational modeling in geography. In physical geography, there are models of climate, hydrology, geomorphology, ecology, and other phenomena. In human geography, there are models of population growth, travel and migration, spatial decision-making, retail and manufacturing location, agricultural and urban land use, cultural diffusion, epidemiology, and other phenomena.

Like all models, computational models are necessarily *simplified* representations of reality. That is, they intentionally ignore or distort aspects of the true richness of reality by incorporating **simplifying assumptions**. In fact, in their initial versions, computational models are often blatantly unrealistic in their simplicity. For example, many models in geography start out making the assumption that their natural or human process takes place on a uniform **isotropic plain**. This is a flat, featureless, and unbounded landscape on which all forms of spatial movement and interaction can occur equally easily in any direction. Such an idea may inspire you to wonder what planet these models are supposed to reflect—what about roads, rivers, mountains, and all the other features that serve as barriers to interaction in certain directions and facilitators of interaction in other directions? Another common assumption in geographic models is that of **economic rationality**. This is the idea that economic agents (individuals, families, corporations) will act exclusively to maximize profit and will be informed in this by complete and accurate knowledge of everything that has implications for profitability. How could such blatantly false assumptions be useful to geographers and other scientists?

The answer is that simplification, even extreme simplification, is necessary in scientific work. It reflects the scientific preference for parsimony we discussed in Chapter 1. Furthermore, it reflects our cognitive limitations as humans—we need to simplify in order to understand complex reality. It might sound a little contradictory, but people must simplify (a form of intentional distortion) if they are to understand anything. The categorical thinking that universally characterizes humans is a manifestation of this cognitive need for simplification. Reality in all its rich complexity, with its contextual factors, probabilistic nature, heterogeneous existence across a huge range of spatial and temporal scales, multivariate and reciprocal causality, emergent patterns, time-lagged and distance-lagged influences, and more, is just too much for our puny little minds to grasp fully and completely. But those who make the scientific bargain we discussed in Chapter 1 accept the compromise of simplicity. They want to know *something* partially and imperfectly rather than *nothing* completely and perfectly. At least they believe, as a matter of faith or personality, that a systematic but imperfect search for truth is better than an irrational and haphazard search, a search based on nothing but invention and fantasy, or no search at all. This aspect of scientific research reminds us a little of the tale about the person who was looking in a parking lot for his lost car keys. His friend asked him why he was searching there, considering he had lost his keys on the other side of the lot. “The light’s better over here,” he replied. Scientists use methods such as simplifying assumptions to shed light on understanding the world; sometimes the methods are bright but do not shine directly on the most relevant portion of reality. Nonetheless, they allow the scientists to see.

In Chapter 3, we listed computational modeling as one of the basic types of data collection in geography. Unlike the other types of data collection, in which portions of reality are directly measured, portions of reality are *simulated* by models. The **model outputs** (usually numerical) are treated conceptually as if they were measurements. These “simulated data” are typically compared to standard empirical measurements made on the portion of reality to which the model refers, as a way



to evaluate the fit of the model to reality (we discuss this more below). However, sometimes models are created and thought about as if they were creations of new realities rather than simulations of existing realities. It is somewhat arguable whether a person who models in order to create new realities is really doing science. You could say they are not because they are not trying to “create and evaluate knowledge about reality” (Chapter 1). Instead, they are doing something like product invention or artifact creation. On the other hand, “reality” certainly includes the intentional products of human activity (buildings, wheat fields), so maybe creating and studying models that do not correspond to the reality that already exists is just the “science of the artificial.”<sup>5</sup>

We discuss computational modeling in this chapter on research design because it can also be thought of as an alternative to traditional experimental designs. As we explained above in this chapter, researchers conduct experimental studies when they manipulate variables—when they control the assignment of cases to levels of independent variables. Computational modeling is such an important 20th-century development because it provides a way to conduct experimental studies when they would otherwise be very difficult or even impossible to conduct. Modeling gives a form of empirical access to events or systems that no longer exist or do not yet exist, that are very rare, or that operate over spatial and temporal scales that are too large to bring into the laboratory. Models afford a form of assignment control when it would otherwise be impossible or ethically unacceptable to manipulate variables. For example, people have often suggested solving particular environmental problems, such as species invasions or the greenhouse effect, by introducing particular chemical or biological agents into the environment; many such introductions are ethically dubious to carry out in reality (see Chapter 14).

Even when traditional empirical approaches *are* viable, modeling has a definite advantage over traditional empirical approaches: the superior ability of models to represent much more of a whole system with many of its complexly interacting parts. Numerous scientific insights of the 20th century involved the realization that causal relations in many systems are not simply binary and unidirectional links, like one pool ball smacking into another. Causality is frequently multivariate, involving many entities or states interacting simultaneously or with complex patterns of **temporally** or **spatially lagged** influences. Causality is frequently **reciprocal**, with one entity or state having causal influence on another that returns an influence back on the first. These complex causal relations are difficult or impossible to study with more traditional means of conceptualizing and empirically evaluating theories.

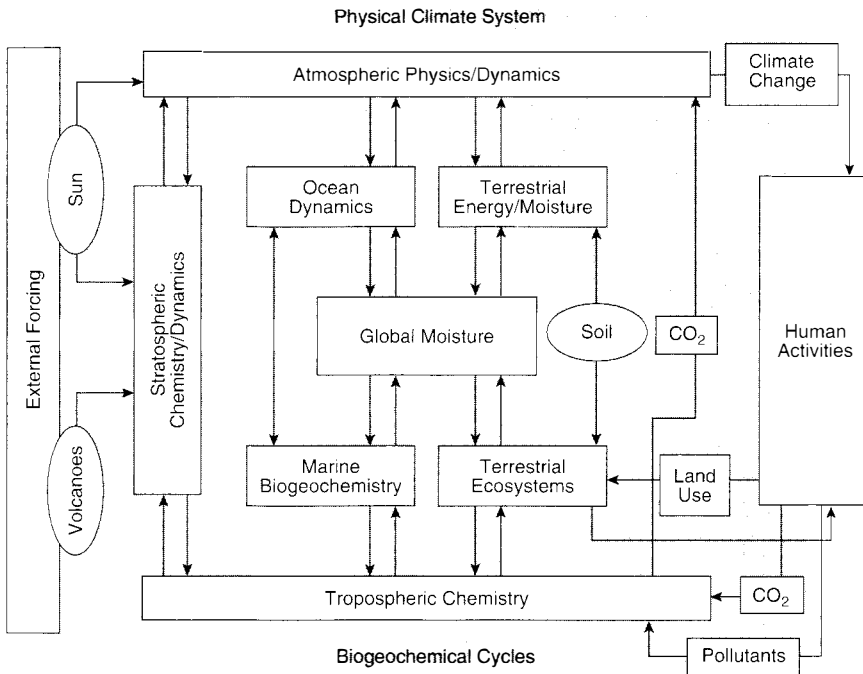
For example, consider Figure 7.2. It depicts a conceptual model of the earth’s systems, showing influences on climate and its change over timescales of decades to centuries.<sup>6</sup> Climate processes operate at multiple spatial and temporal scales,

---

<sup>5</sup>Echoing the title and theme of the well-known book: Simon, H. A. (1969). *The sciences of the artificial*. Cambridge: MIT Press.

including scales too large to represent physically in the lab (although processes that operate over very large or small, or slow or fast, scales often do not need to be explicitly represented in such models). Furthermore, climate reflects complexly interacting states and influences of the atmosphere, ocean, land, ice and snow, and terrestrial and marine biota—including humans and their activities. This is reflected by the climate researchers' use of the term **forcing** to refer to a causal factor in computational models, analogous to an independent variable in true experiments. The term forcing, however, connotes the fact that causal variables in complex systems, such as climate, are inputs of matter and energy that exist within the complicated network of interacting multivariate causality that is the system. The actual observable effect of such a cause often depends on **feedback** influences from other parts of the system. In the case of **positive feedback**, the effect of a forcing can be magnified considerably; for example, increasing water evaporation on earth can lead to increasing surface temperature (water vapor is a major greenhouse gas), which in turn can lead to even more evaporation and water vapor, and so on. In the case of **negative feedback**, a forcing can be dampened so that its influence is negligible; an example is the fact that increasing water in the atmosphere from increasing surface temperature can lead to increasing cloud formation, which in turn would *decrease* insolation on the earth surface and cool it down. In many cases of feedback, changes in a forcing do not necessarily result in much change to outcome variables until the forcing changes reach particular thresholds. Characterizing the magnitude and pattern of changes to outcomes as a result of changes to forcing factors is the subject of **sensitivity analyses**. Finally, we should not forget that climate change requires modeling insofar as most people would find it unethical to carry out experiments to manipulate forcings that they believed might actually change climate in the real world.

Researchers typically distinguish between models whose parameters are based on physical laws (law-based or **numerical models**) and models whose parameters are based on estimates made from data (**empirical models**). Empirical modeling is common in both physical and human geography, but numerical modeling is much more common in physical geography (some population models do incorporate simple population-growth laws). The laws instantiated in numerical models in geography and other earth sciences come from physics, chemistry, biology, and so on. The model in our climate example would utilize the laws of thermodynamics, Newton's laws of motion, and so on. In Chapter 2, we described the difference between probabilistic and deterministic processes, in the specific context of causal processes. Empirical models are always probabilistic or stochastic in nature, because the data upon which they are constructed are understood to derive from samples rather than populations, and because empirical measurement is understood to contain error or noise. Numerical models may be deterministic or stochastic. **Deterministic numerical models** incorporate mathematical expressions of laws to derive parameters without any random input or noise. Given an initial state and set of boundary conditions, a well-constructed deterministic model gives a single precise output. **Stochastic numerical models** ("Monte Carlo models") incorporate a simulation of random processes in them. The randomness is typically supplied by



**Figure 7.2** Conceptual model of earth's systems depicting influences on climate and its change over timescales of decades to centuries.<sup>6</sup>

a random number generator function available in whatever computer software is used to implement the model in the computer (see Box 8.1). Each “run” of a stochastic model will give a somewhat different output for the model. Usually, researchers who use stochastic models focus on patterns that emerge over many runs of the model. A fascinating scientific insight of the 20th century is that such stochastic models can be “chaotic”—their outputs can vary widely as a function of small changes to input values.

## Steps of Computational Modeling

We describe the basic logic of conducting computational modeling as a series of steps in Table 7.2. The table presents an idealized view of computational modeling; depending on the purpose and topical domain of the model, these steps are certainly applied in somewhat different ways and are not fully realized in all instances. First, you must identify and express the conceptual underpinnings of your model.

<sup>6</sup>Adapted from Earth System Science (1986). *A program for global change*. Washington, DC: National Aeronautics and Space Administration.

**Table 7.2** Steps of Computational Modeling

---

1. Create conceptual model
<ul style="list-style-type: none"><li>• derive it from existing theories and concepts, and any other sources</li><li>• visualize it in appropriate graphic form</li></ul>
2. Create computational model
<ul style="list-style-type: none"><li>• identify relevant parameters for structures and processes, their boundary conditions, and so on</li><li>• base it on scientific laws (numerical models) and/or empirical data (empirical models)</li><li>• express model in formal language (mathematical equations, computer operations)</li><li>• write computer program, programming from scratch or using appropriate commercial modeling software</li></ul>
3. Run the computer program
<ul style="list-style-type: none"><li>• produce output as form of “data collection”</li><li>• debug as necessary</li></ul>
4. Compare model output to empirically obtained data
<ul style="list-style-type: none"><li>• base data on measurements of past or current events or states</li><li>• evaluate model both qualitatively and quantitatively</li><li>• interpret discrepancies between model output and empirical observations by refining the model, returning to Step 1, 2, or 3</li></ul>
5. Accept, use, and communicate model
<ul style="list-style-type: none"><li>• eventually, accept model and use to forecast future, explain phenomena, and so on</li><li>• communicate model and its uncertainties to researchers and other interested parties</li></ul>

---

All computational models have a conceptual model underneath them, at least implicitly; it is highly desirable that you make the conceptual model explicit at this stage. You can derive the entities and their interrelating processes that you consider relevant to include in the model from existing theories and concepts, or any other source of ideas; often enough, the entire conceptual model already exists in published scientific literature. It is quite helpful in most modeling situations to visualize the conceptual model as a flow diagram, network structure, or some other graphic; the climate model in Figure 7.2 is an example.

Then you translate the conceptual model into computational form. You must identify and quantify specific **model parameters**, quantify processes that express causal influences, and so on. The identification and **calibration** of model parameters is based on applicable physical laws, prior empirical data, or both, as we described above. You calibrate a model by adjusting its parameters to fit the context

of your particular modeling situation; for example, you have to tune parameters to the particular spatial and temporal scale of your model. You express the computational model formally as mathematical equations, computer operations, and so on. Then you instantiate this formal expression on a computer as a program or set of programs, either programming from scratch or using appropriate commercial modeling software.

Proceed with the “data collection” phase by running the model. Generate model output by providing inputs that are processed through the computational steps of the model. The inputs are based on measured or simulated states of the relevant parameters in the model. You must typically **debug** the model at this step, often extensively. That is, you determine that the program is computing as intended, and locate and fix any errors in it.

You then evaluate and interpret the model output. Do this by comparing “data” output from the model to empirical measurements of the actual system that is being modeled. These measurements are based on past or current events or states, if they exist at all; it is not uncommon for appropriate data to be lacking. Evaluation occurs both qualitatively and quantitatively. Qualitative evaluation involves such things as identifying whether the influence of variables operates in the right direction; for example, does increasing one variable cause another variable to increase or decrease? Quantitative evaluation typically involves statistically testing numerical discrepancies between the model output and the empirical measurements using some version of a statistical **goodness-of-fit** test.<sup>7</sup> Interpret discrepancies between the model output and the empirical observations you used to evaluate it. This will usually lead you to refine the model by returning to steps 1, 2, or 3. You can identify missing parameters, recalibrate parameter boundary conditions and process influences, and so on. You might conduct the sensitivity analyses we described above at this step.

Eventually, you accept the model and use it to describe complex systems, forecast the future, explain phenomena, or achieve practical control over them. One of the most intriguing aspects of model use can be communicating it to other people, including the various uncertainties involved in its outputs and interpretation. This task calls for all of the expertise and cleverness in displaying data you can muster, especially when, like climate change, your model potentially has great implications for society and at the same time is politically contentious.

---

<sup>7</sup>We cover statistical data analysis in Chapter 9. There we discuss statistical “hypothesis tests” that generate probabilities about the correspondence between sample results and possible population truths. The mathematical mechanics of goodness-of-fit tests are essentially those of hypothesis testing, but there is an interesting difference between a researcher’s attitude toward the probabilities of standard hypothesis tests and those of goodness-of-fit tests. When conducting hypothesis tests, as we learn in Chapter 9, one usually wants sample data to *differ* significantly from likely outcomes of the tested null hypothesis (to be improbable if the null were true). When conducting goodness-of-fit tests, in contrast, one usually wants the model output *not* to differ significantly from the measured data.

## Review Questions

---

- What is research design, and why is it so important to conducting and interpreting research?

### Empirical Control in Research

- What are the three forms of empirical control in research?
- What distinguishes experimental from nonexperimental research, and what are the implications of this distinction for the meaning of research?
- Why is statistical correlation ambiguous with respect to causal relationships? What are confounds? What is a control group?
- What is the distinction between laboratory and field settings, and how does this distinction relate to the experimental/nonexperimental distinction?

### Basic Research Designs

- What are between-case and within-case designs, and what are strengths and weaknesses of each?
- What are several specific research designs, and what are some of their strengths and weaknesses?
- What are developmental research designs? What are cross-sectional, longitudinal, and sequential developmental designs?
- What are some implications of the difference between single-case and multiple-case designs?

### Computational Modeling

- Why is computational modeling so useful to geographic researchers?
- Why are simplifying assumptions incorporated into computational models?
- What distinguishes numerical models from empirical models? What distinguishes deterministic models from stochastic models?
- What are the basic steps to conducting research with computational models?

## Key Terms

---

**assignment control:** form of empirical control in which researchers create at least one of the variables in the study—that is, researchers determine which level of the variable cases are exposed to

**attrition:** problem with longitudinal designs of losing cases during the course of the study; also called “mortality”

**between-case design:** research design in which different cases take on different levels of an independent or predictor variable

- calibration:** adjustment of model parameters to fit the context of a particular model, the specific ranges of its input values, the scale of its units of analysis, and so on; calibration may be based on applicable physical laws, empirical data, or both
- case study:** intensive and comprehensive descriptive study of a single case
- cohort:** group of cases of the same age (originating at the same time)
- cohort effect:** threat to the validity of cross-sectional designs wherein a property of cases that arises because they come from the same cohort is mistaken for a property of cases at a certain level of development
- condition:** one of the discrete levels of a manipulated variable in an experiment
- confound:** variable in a study that is not intended to be a causal variable but affects one of the measured variables; also called a “third variable”
- control group:** condition of an experiment that is meant to provide the necessary comparison of “no manipulation”; in fact, all experiments require whatever conditions are necessary to make particular comparisons
- cross-sectional design:** developmental research design in which two or more groups of cases, each at different “ages” or levels of development, are compared at the same time; also called “synchronic”
- debug:** process of determining that the program running a computational model is computing as intended, including locating and fixing errors
- dependent variable (DV):** measured, nonmanipulated variable in an experiment that plays the role of a potential effect
- deterministic numerical model:** numerical model constructed with parameters that assume no random inputs or noise; generates a single precise outcome given a single set of inputs
- developmental design:** research design to study systematic (nonrandom) processes of change over time in human or physical systems
- economic rationality:** common simplifying assumption in geographic models that economic agents act exclusively to maximize profit and are informed by complete and accurate knowledge of everything that has implications for profitability
- empirical control:** any method of increasing the ability to infer causality from empirical data, including physical, assignment, and statistical control
- empirical model:** computational model with parameters based on estimates from data; common in both human and physical geography
- experimental study:** research study in which assignment control (manipulation) is exercised over one or more variables
- factorial design:** experimental research design in which two or more independent variables (factors) are manipulated

**feedback:** reciprocal causal influence in a complex system

**field:** setting where researchers conduct studies on phenomena of interest as they “naturally” occur; usually outdoors in wilderness and inhabited settings

**forcing:** term for a causal factor, an input of matter or energy, in a computational model; connotes the factor’s nonlinear and uncertain influence on effects in a complex system like climate

**formal language:** abstract mathematical or logical symbol system that precisely defines entities and processes that operate on the entities

**goodness-of-fit test:** statistical test to evaluate how well computational model outputs match empirically observed values

**history effect:** threat to the validity of longitudinal designs that arises because of idiosyncratic events or conditions that hold during the particular time period of the study

**idiographic approach:** specific, idiosyncratic approach to creating knowledge in research and scholarship; more characteristic of humanities than sciences

**independent variable (IV):** manipulated variable in an experiment that plays the role of a potential cause; also called a “factor”

**instrumentation:** threat to the validity of longitudinal designs that arises because of changes over time in the way measurements are made

**interaction:** when the effect of one IV on a DV in a factorial design is different at different levels of another IV; more complex interactions involving more than two factors are possible too

**isotropic plain:** common simplifying assumption in geographic models that processes operate on a flat, uniform landscape with no barriers to interaction in any direction

**laboratory (lab):** specially designed setting, usually a room inside a building, that allows researchers to exert physical control while conducting studies

**levels of variable:** the number of different values a variable can take, usually in the context of manipulated variables that take on a modestly sized number of values or “groups”

**longitudinal design:** developmental research design in which one group of cases is compared to itself over time as it develops; also called “diachronic”

**manipulation:** the assignment control of one or more variables in an experiment

**matching:** variation on random assignment to conditions in which two cases that have some characteristic in common are assigned to contrasting conditions in an experiment

**model output:** the simulated data that computational models produce after they are run



- model parameters:** quantitative statements in computational models that mathematically express the characteristics of factors in the model and how they interrelate and change during runs of the model
- negative feedback:** specific pattern of feedback wherein an increase in one forcing causes a change in another forcing that in turn causes the first forcing to decrease; dampens changes in the system
- nomothetic approach:** general, even lawlike approach to creating knowledge in research; more characteristic of sciences than humanities
- nonexperimental study:** research study in which assignment control (manipulation) is *not* exercised over any variables
- numerical model:** computational model with parameters based on prior scientific laws, primarily from physics and chemistry; common only in physical geography
- physical control:** form of empirical control in which researchers physically modify or restrict the data collection situation in order to reduce the potentially distracting influence of confounding factors that are not of interest
- positive feedback:** specific pattern of feedback wherein an increase in one forcing causes a change in another forcing that in turn causes the first forcing to increase more; magnifies changes in the system
- posttest-only design:** weak research design in which a single episode of measurement occurs after an event has taken place
- pretest-posttest design:** research design that is better than the posttest-only design because measurement episodes occur both before and after an event has taken place
- quasi-experiment:** study without manipulated variables that attempts to establish causal relations more validly by applying systematic statistical control over alternative causal variables
- random assignment:** assigning cases to conditions in an experiment according to a random decision rule
- reciprocal causality:** when two entities influence each other
- research design:** the structure of possible comparisons in a study that results from the way variables are created or measured
- reversal design:** single-case experiment wherein the case is measured before the manipulation is applied, after it is applied, and a third time after the manipulation is removed (“return to baseline”)
- sensitivity analysis:** characterizing the magnitude and pattern of changes to outcome variables in a computational model as a result of changes to forcing factors
- sequential design:** developmental research design that is a hybrid of cross-sectional and longitudinal designs

- simplifying assumption:** aspect of the representation of reality in a model that is unrealistically simple but helps researchers understand the complexity they are modeling
- single-case experiment:** experiment conducted with a single case
- spatial lag:** influences in models, including feedbacks, that express their effect over some spatial distance (via some continuously connected mechanism that may not be evident)
- spurious causality:** attributing causality to a variable that correlates with another variable when it is not actually its cause
- statistical control:** form of empirical control in which researchers explicitly identify, measure, and statistically analyze variables they think might have an effect on the main variables of interest
- stochastic numerical model (“Monte Carlo model”):** numerical model constructed with parameters that assume random inputs or noise; generates different outcomes given a single set of inputs
- temporal autocorrelation:** nonindependence among measurements of phenomena as a function of the time of their occurrence relative to other phenomena; it is the temporal analogue of spatial autocorrelation
- temporal lag:** influences in models, including feedbacks, that express their effect after some time delay
- time-series design:** research design that incorporates a large number of repeated measurements over time in order to study developmental phenomena
- within-case design:** research design in which, over time, every case takes on each different level of an independent or predictor variable; also called a “repeated-measures design”

## Bibliography

---

- Bradford, M. G., & Kent, W. A. (1977). *Human geography: Theories and their applications*. Oxford: Oxford University Press.
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Boston: Houghton Mifflin.
- Kenny, D. A. (1979). *Correlation and causality*. New York: Wiley.
- Kirkby, M. J., Naden, P. S., Burt, T. P., & Butcher, D. P. (1993). *Computer simulation in physical geography* (2nd ed.). Chichester, U.K.: Wiley.
- Lane, S. N. (2003). Numerical modeling in physical geography: Understanding, explanation and prediction. In N. Clifford & G. Valentine (Eds.), *Key methods in geography* (pp. 263–290). Thousand Oaks, CA: Sage.
- Schneider, S. H. (1992). Introduction to climate modeling. In K. E. Trenberth (Ed.), *Climate system modelling* (pp. 3–26). Cambridge and New York: Cambridge University Press.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.
- Stake, R. E. (1995). *The art of case study research*. Thousand Oaks, CA: Sage.