CHAPTER 2

Fundamental Research Concepts

Learning Objectives:

- What are the major idea concepts and the major empirical concepts in science?
- What is causality, and what role does it play in scientific inquiry?
- What are the four levels of measurement, and why are they important?
- What does the concept of scale mean in geography?
- What are some systematic ways to generate research ideas?

In Chapter 1, we discussed the fact that scientists create and evaluate ideas, via logic and empirical observation. In this chapter, we flesh out a variety of basic scientific concepts that are fundamental to the conduct and interpretation of scientific research. We classify these as **idea concepts** and **empirical concepts**.

Idea Concepts

Given that scientific research ultimately concerns developing valid ideas about reality, we first consider the idea concepts: theory, law, hypothesis, causality, model, and construct. Probably the central idea concept is **theory**. Defined narrowly, as we prefer, a theory is an idea or conjecture about a causal relationship in reality. It answers the question of "why" something is the way it is by identifying its antecedent causes. This narrow use that we prefer is in contrast to the term's broad use in everyday speech to refer to a conjecture about any aspect of reality. For example, I could "theorize" that door number 2 has a goat behind it, which of course is not a guess about *why* there is a goat there. Scientists also use the term broadly in some cases; they sometimes speak of "theorizing" about a description or prediction of reality, rather than an explanation. For instance, chemists might speak of the "theory of atomic structure," which posits electrons, protons, and so on. Although the theory does not necessarily posit that the components of the atomic structure have a causal relationship to each other, it does posit that their structure is the cause of patterns of data that result from observing atoms. Similarly, we may understand the "theory of gravity" as a description or prediction of the behavior of bodies with mass; in and of itself, the theory does not explain *why* bodies attract. Mathematical expressions of relationships that we expect to hold precisely in an "ideal" world, like the formula for the force of gravitational attraction between two bodies, are often called **laws** rather than theories. For the most part, laws have been identified only in the physical sciences.

We prefer to use "theory" in the narrow sense of a conjecture about causality in order to recognize explanation as the ultimate goal of basic scientific research, as we discussed in Chapter 1. We also prefer the narrow use of "theory" in order to distinguish it from another common idea concept, that of **hypothesis**. A hypothesis is a conjecture about a pattern of observations of the world. The difference between theory and hypothesis is in fact fairly subtle, and as we have suggested, some people use the terms virtually synonymously. But the term hypothesis tends to refer more to a specific and directly testable idea, often times a specific idea about a particular pattern of data in some population that may or may not imply anything directly about causality. We may think of the relationship between theory and hypothesis in terms of conditional ("if-then") logic: If theory A is true, then one hypothesizes that data pattern B will hold.

Let's consider the idea concept of **causality** a bit more. There is a long history of philosophical analysis of causality, notably by the 18th-century British Empiricist philosophers, particularly David Hume. Causality is the apparent fact that the occurrence of an event A (state A, entity A) brings about or determines the occurrence of event B; A is a reason B occurred, or the occurrence of B depends in some sense on the occurrence of A. A is the **cause** and B is the **effect**. For instance, lightning is a cause of wildfires. Hume offered three principles in his analysis of causality:

a. Covariation between cause and effect (they co-occur)

b. Temporal precedence of cause (the cause comes first)

c. Controlling the cause will control the effect.

The 20th century witnessed a great deal of critical discussion of the concept of causality, especially in the context of quantum and relativistic physics. Is causality in the mind or in the world? Can causality occur simultaneously between two spatially separated entities? Can causality move backward in time? In spite of these discussions, however, it is still true that most everyday scientists who focus on reality above the atomic size (including geographers) accept the importance of causality as a concept and the meaning of theories as attempts to explicate causality. That is, notions like quantum causality apparently have little or no relevance to the concerns of most geographers.

But it is clearly naive to think of causality simply as one perfect pool ball hitting another perfect pool ball on the felt-covered slate of reality. In most areas of geography, both physical and human, causality is complex and multivariate. Furthermore, it is typically **probabilistic (stochastic)** in nature rather than **deterministic**. That is, it is only required that causes *probably* bring about effects, not definitely each and every time the cause occurs. In fact, a few problem areas in geography, particularly physical geography, do employ deterministic rather than probabilistic analyses. We discuss this further in Chapters 7 and 9 but note here that probabilistic processes occur in systems in which causality is complex, and partially outside of our ability or interest to observe or conceptualize.

The notion of probabilistic causality is related to an old distinction between **necessary** and **sufficient causes**. An effect cannot happen if a necessary cause does not occur, but it may not happen just because the necessary cause occurs. Droughts are pretty much necessary causes of wildfires; generally speaking, a wildfire will not take place without dry conditions, but wildfires do not have to take place just because it is dry. An effect happens if a sufficient cause occurs, but it may happen even if the sufficient cause does not occur. Given dry conditions, a lightening strike constitutes a sufficient cause of wildfires; an arsonist can take the place of the lightning, however.

Another distinction is between **mechanistic** and **functional** causality. Mechanistic causality is the classic idea that causes move forward "densely" in space and time—the continuously connected causality we discussed in Chapter 1 as being widely preferred by scientists. Let's consider this preference in greater detail with an example. Consider the scenario of hitting a light switch to cause a bulb to illuminate a room. Starting somewhat arbitrarily at the moment when the metal piece connected to the switch provides contact between the power source and the wire leading to the light bulb (we could start, for example, at the moment we desire to turn on the light), causality moves continuously along the wire, hits the base of the bulb, and is transmitted to the filament, which causes electrical stimulation, making the wire emit electromagnetic photons in the visible range of the spectrum. Then the photons quickly move out continuously over the space of the room, directly impinging on the retina or bouncing off surfaces before hitting the retina, then travel along the optic nerve and tract to the visual cortex.[1] That's mechanistic causality.

Alternatively, many theories in and out of science posit a different type of causal explanation, one that places the cause after the effect by focusing on the cause as functional or purposeful—one that considers the cause as a goal. For example, Darwinian evolution is often thought of as "survival of the fittest," as if the traits of organisms evolve because of their function in improving the organism's future reproductive fitness. In economic geography, individuals and firms are often said to locate themselves in order to "maximize profit," a future state of affairs that apparently motivates the present or past. Functional causality may seem to violate scientists'

[1]The question of how causality is then transmitted continuously to brain areas responsible for the conscious experience of light is somewhat understood, although the conscious part is still largely a scientific mystery. As a *scientific* mystery, however, we assume the ultimate answer would not violate continuously connected causality.

antipathy to disconnected and backward-acting causality. In fact, we can make functional causality compatible with mechanistic causality by rephrasing functional explanations ultimately in mechanistic terms. Evolution does not occur because of an attraction to a future perfect form; genetic changes are random or haphazard, but some traits are simply more likely to lead to survival and reproduction in particular environments, which leads to the creation of more organisms with the new genetic material. Likewise, a firm does not locate because of future profit; it locates because of the current *anticipation* of future profit. We can find functional causes in scientific theories, but we should understand them as useful heuristic devices rather than literal truths.

Closely related to the concept of theory is that of **model**. A model is a simplified representation of a portion of reality, expressed in conceptual, physical, graphical, or computational form. A model is essentially a complex information-bearing assemblage that conveys a set of interrelated theories about structures and processes of a system of interest in the world—it expresses the parts and their causal interactions. There are many examples of models from all areas of systematic geography. For instance, the Huff model, an example of so-called gravity models in economic geography, predicts and explains store choice by consumers according to the distance the consumer must travel to each store and the relative attractiveness of each store compared to the others in the comparison set. Another instance is the Davisian model, a geomorphologic model of land forms, which says that the physical shape of the earth's surface terrain is a function of the geological composition and elevation of the land (due to mountain-building forces), processes of soil and rock denudation (wearing away and transport), and the passage of time. These two examples are relatively simple models; sometimes models reach great levels of complexity in geography. We discuss models and their logic in more detail in Chapter 7, particularly those expressed in mathematical or computational form. But it is worth emphasizing here that when we say models are "simplifications of reality," we want to stress that they are *necessarily* simplified and that their usefulness to researchers stems a great deal from their simplified nature.

Our final idea concept is that of **construct**. Actually, construct is essentially another word for a scientific concept. Just as a theoretical statement is an elementary component within a model, a construct is an elementary component within a theory. Constructs are pieces of the idealized world that compose the subject matter of theories, and they are the hypothetical entities that we attempt to measure when we perform our systematic empirical observations. For example, in John's research discussed in Chapter 1, he attempted to measure the construct of "dissociative institutions" by counting liquor stores in the phone book. Biogeographers use the construct of "plant communities" in some of their theories and attempt to measure them by observing and recording the locations of various plant species, their soil and climate conditions, and such properties as elevation and angle to the sun. Discussing constructs in this way may seem to imply that they are rare, exotic, or the sign of a scientist who doesn't really understand what he or she is studying. On the contrary, constructs are utterly commonplace, even ubiquitous, among all scientists, including physical scientists. For example, the idea that a table has length treats length as a construct; we realize that generating a number by laying a ruler end to end is not the table's *actual* length but the result of one attempt to *measure* its length. Such a measurement is at least slightly inaccurate and could be very inaccurate. The construct of a table's length cannot be in error this way.

Thus, it is critical to note that constructs are abstract idea entities that scientists care about, but they are not observed directly—their effect upon measurements is observed. Early in the 20th century, psychologists attempted to finesse the difficult problem of defining intelligence by saying it was "what intelligence tests measure." But intelligence is *not* what intelligence tests measure—it is what intelligence tests *try* to measure. Thought of this way, a construct may be called a **latent variable**, whereas its expression as a set of observations may be called a **manifest variable**. We try to observe as well as we can; that is, we try to measure accurately and precisely and completely (more below).[2] But in fact, measurements (manifest variables) are always imperfect reflections of the constructs (latent variables) they attempt to capture. In Chapter 11, we discuss issues of validity that arise from the relationship of constructs to their measurement.

Empirical Concepts

Empirical concepts include cases, variables, measurement, measurement levels, discrete versus continuous variables, and accuracy versus precision of measurement. A **case** is the thing or entity studied. Synonyms include unit of analysis, entity, element, individual, research subject, and respondent. Cases in physical geography are sometimes called "samples," as in a "soil sample" or a "water sample." This usage is a bit confusing, however, insofar as the entire set of entities we measure in a study is called the sample (Chapter 8)—not just one of the entities. In human geography, the case is often an individual person or group of people, such as a family; a city block; a city, county, state, or country; a census tract or other census region; an industry or corporation; or a society or cultural group. Examples of cases in physical geography include water bodies, such as lakes, rivers, marshes, estuaries, or oceans; mountains or mountain chains; air masses; forests or other vegetation communities; soil profiles; and ecosystems. These examples show that cases in geography, especially physical geography, are often units of time or space carved out of a continuous reality. We will see in Chapters 8 and 9 that this creates some very intriguing issues in geographic data sampling and analysis but also some rather special difficulties.

But we do not study cases directly—we study attributes or properties of cases. We don't study mountains; we study their formation or mineral content. We don't study cities; we study their economic base or percentage of senior citizens. Because they vary from case to case, or within cases over time, these properties are called

[2]Defining a variable in our research by describing the techniques (operations) that are used to measure it is called an **operational definition**. The Methods section of a research report would always describe these operations, so that a reader can understand and evaluate what was done in the research (see Chapter 13).

variables, as opposed to **constants**. To say that variables vary is to say that they take on multiple values when observed across cases; thus, the simplest variable possible is a **dichotomous variable** with two values.

The process by which we observe cases and determine their values on our variables of interest is called **measurement**. Or put more formally, measurement is assigning numbers to cases to reflect their values on a variable; in the case of nominal classification, nonnumerical symbols may be used. To return to our table example, we may measure the length (the variable) of a table (the case) by laying a ruler end-to-end (measurement procedure) to determine that it is 2.27 meters long. The measured numbers are the **data** or data set.[3] Measurement of many different types occurs in geography, as we discuss in greater detail in Chapter 3 and several of the subsequent chapters. A mountain's elevation in meters comes from a global positioning system (GPS). The average number of vehicles that travel through a busy intersection each hour comes from time-stamped digital signals recorded whenever anything of sufficient weight goes over a cable laid across the road. The age of trees of a particular species that grow in an area comes from a count of their rings exposed in a tree core. A resident's attitude about a toxic waste dump, expressed as a numerical value on a rating scale, comes from a survey he or she fills out.

As these examples make evident, measurement varies in terms of what type of case is measured, what variable is measured, and how it is measured. Of central importance, however, is characterizing the quantitative content that results from measuring particular things in particular ways. This is called the **measurement level** of a variable—the degree and nature of quantification implied by a measurement. It's important because of its implications for the way we choose and interpret techniques of data analysis and display, as we discuss in Chapters 9 and 10. There are four levels of measurement, starting with the least quantitative content and ending with the most. Each level expresses the quantitative content of all levels above it:

*1. **Nominal.*** Nominal measurement is not quantitative at all. It is simply assigning numbers (or letters or any other symbols) to distinguish one case's value on a variable from that of another case. Nominal measurement most often expresses *classification*—the placement of a case into a class or category that has qualitatively different properties than other classes: for example, the species of each plant in a particular ecosystem. Sometimes, however, nominal measurements simply name, distinguishing one case from another: for example, the case number assigned to each tree in a database that serves as a distinguishing label. Whether classifying or only naming, numbers used to record nominal measurements have no quantitative meaning, although of course numbers used to indicate classes do express qualitative meaning

[3]"Data" is a plural word in its original Latin, so traditional use dictates we speak of data in the plural, such as "The data show that . . ." rather than "The data shows that . . ." "Data set" is singular for the entire collection of data, and a single measured score is a "data point" or even "datum" (Latin). Recently there have been indications that editors and other gateway tenders will allow "data" to be used singularly, but we maintain the traditional plural use in this book.

(a white oak is different than a black oak). Other examples of nominal measurement in geography include soil type, sex (gender), and type of primary industrial activity.

*2. **Ordinal.*** Ordinal measurement is minimally quantitative. It is assigning numbers to distinguish the relative order, or rank, of the value of one case on a variable from that of another case; for example, the oldest tree gets a "1," the next oldest gets a "2," and so on. Notice that ordinal variables do express "more" and "less," which are quantitative properties, but they do not specify *how much* more or less one case is than another. The second-oldest tree might be 10 years or 100 years younger. In geography, examples of ordinal measurement include ranking cities in terms of importance in the urban hierarchy or ranking streams in terms of their position within a watershed. Many textbooks in geography and other disciplines recommend that scores from rating scales of attitudes or preferences be treated as ordinal data; that is, that rating preference for the state of Ohio as a "5" and the state of Maine as a "9" represents four "ranks of liking" rather than an interval of four "units of liking." In Chapter 6, we discuss the use of rating scales in explicit reports and argue that rating-scale data should not be treated as merely ordinal.

*3. **Interval.*** Interval measurement expresses not only the ranks of cases on some variable but the quantitative lengths of intervals between the cases: for example, the relative locations of trees in a stand. Although interval measurements contain information about lengths between data scores, they do not contain information about a *true zero.* That is, an interval variable does not express a value of "nothing"—no amount of the variable. The classic example is temperature expressed in Celsius or Fahrenheit; 0° does not mean no temperature or no heat but simply represents another value like the rest do.[4] In our example, none of the trees can be said to have "no" location. In fact, spatial location is an important example of an interval variable in geography; consider location expressed in a spatial coordinate system like latitude-longitude (0° latitude is not "no" latitude).

*4. **Ratio.*** Ratio measurement expresses not only the lengths of intervals between cases on some variable but also the lengths of intervals relative to a true zero. For example, the widths or heights of trees are ratio variables. Because a ratio variable does express a value of "nothing," comparisons between its score values can be validly conveyed as a ratio. That is, a tree that is 0.8 meters wide is *twice* as wide as a tree that is 0.4 meters wide (a ratio of 2:1). Notice how interval variables cannot validly be placed in a ratio like this; 70° F is *not* twice as hot as 35°, notwithstanding that we have heard weather reporters say this. Examples of ratio variables in geography are very common, and include amounts of rainfall, distances between places, and family incomes. Ratio and interval measurement, taken together, are known as

[4]0° kelvin (−273° C) is a theoretical abstraction that represents absolutely no heat. Of course, 0° C (slightly above 0°, to be precise) does have the special relevance of corresponding to the phase change between liquid water and ice. To many people, 0° F has the special psychological relevance of indicating a temperature below which it is preposterously cold to go outside.

metric, to reflect their important property of expressing quantitative distances between values.[5]

Related to the concept of measurement level is the concept of whether variables are **discrete** or **continuous**. This concept too has implications for appropriate data analysis and display. Discrete variables have a limited set of distinct possible values. For instance, the number of states bordering a given state in the conterminous U.S. is 1, 2, 3, 4, 5, 6, 7, or 8 (can you guess which has 8?[6]). It is not possible for a state to border a fractional number of other states, because any contact, even at a point, is considered one whole contact. Similarly, a city may contain 123,488 or 123,489 people, but not 123,488.3 people (a discrete variable like this is called **countably infinite** because it could take on any arbitrarily large value). Between any two values of a continuous variable, in contrast, there are potentially an infinite number of additional values. Between a snow pack depth of 1.52 and 1.53 meters could be a depth of 1.524 meters. Thus, continuous variables essentially map onto the real number line, or a piece thereof. Only measurement precision (discussed below) limits the number of possible values of a continuous variable. Because measurement must necessarily have finite precision, any actual data always consist of discrete values, although the number of different values may be very large and include values with several digits past the decimal. How does the discrete-continuous distinction relate to levels of measurement? There is a partial overlap. Nominal and ordinal variables are necessarily discrete; interval and ratio variables may be either discrete or continuous. Put conversely, discrete variables may be any of the four levels, but continuous variables must be interval or ratio.

The final empirical concept we introduce here is the distinction between **accuracy** and **precision of measurement**. Accuracy refers to the correctness of a measurement—how close the measured value is to the true value of the thing being measured. Precision refers to the sharpness or resolution of a measurement—how small the units are with which a value is measured. To understand this distinction more clearly, it may help to consider an analogy to a cluster of darts thrown at once toward a target (Figure 2.1). Think of the resulting cluster of five darts as a single measurement, with the bull's-eye as the true value of the thing being measured. The distance from the spatial center (centroid) of the darts to the bull's-eye is accuracy; the spread of the five darts around their centroid is precision. As Figure 2.1 shows, a tight cluster of darts may be centered near the bull's-eye or far from it. A wide cluster may be centered right over the bull's-eye or some distance away; of course, a wide cluster would likely have some darts off the board if it were inaccurate in this manner. Similarly, a digital bathroom scale may measure in a precise manner

[5]Many writers reserve the term "measurement" exclusively for ordinal and metric measurement, or even just metric. We find it useful to use the term to refer to all situations in which numbers or other symbols are assigned to cases to represent their value on variables, even if that variable is a set of classes. In other words, all empirical studies in geography involve some form of measurement, even if it is "qualitative" measurement.

[6]Both Tennessee and Missouri border eight states. Only Maine borders one state.

Box 2.1 Zeno's Paradoxes: Space, Time, and Theme as Discrete or Continuous

The distinction between discrete and continuous is not as straightforward or as mundane as it might seem. Yes, it has implications for data collection, analysis, and display. More than this, however, the distinction is in fact a major intellectual enigma that can be fascinating to ponder. In Chapter 8 and again in Chapter 12, we learn that geographers conceptualize phenomena as being continuous fields (for example, atmospheric temperature) or discrete objects (for example, lakes). But this is not always easy to decide. We noted in this chapter that measurement precision is necessarily finite, which ultimately forces all data to be discrete. Arguably, however, all real phenomena are actually continuous. Even nominal variables like soil type and sex are categorical simplifications of multivariate and continuously valued possibilities of reality (for example, mixed alphisols-mollisols, hermaphrodites). Seemingly discrete entities like clouds, lakes, and mountains actually have vague boundaries that are difficult to identify precisely and change over time; their very existence as objects is debated by trained scientists, never mind lay people. (We once attended an entertaining talk at a national geography conference that concerned itself with the question of whether the "mountains" of West Virginia are really tall enough to deserve that name; the speaker concluded they are not.) But then again, one can also make a good argument that all real phenomena are actually discrete. As the physical sciences have apparently shown us over the centuries, all reality is really composed of multitudes of tiny discrete entities (atoms, electrons, photons, quarks . . . ?). So which is reality—continuous or discrete?

The Greek philosopher Zeno, in his famous paradoxes about space and time, touched upon this mystery more than two millennia ago. Zeno of Elea was a contemporary of Socrates and Plato who believed, with Parmenides, "*that there is only one thing, and it does not change or move, never came into existence and will not cease to exist.*" (We don't even want to think about the implications of this for our lives.) Zeno presented four logical arguments in support of this that can largely be understood to rest on the nature of space and time as discrete or continuous substrates for reality. These four "paradoxes" attempt to prove that change and motion are impossible. A paradox is a seemingly contradictory statement that may nonetheless be true; ultimately, the contradiction of a paradox is only apparent. Zeno's four paradoxes are as follows:

1. **The motionless runner.** The first paradox concerns a runner trying to get from a start location at A to a finish at B. The paradox argues that motion is impossible because in order to get from A to B, the runner must first get halfway between A and B. Before the halfway point, the runner must get one fourth of the way, and so on. This paradox rests on discretizing space into an infinite number of points, all of which must be reached in an infinite number of moments of time. Zeno argues that this is forever.

2. **Achilles and the Tortoise.** The second paradox posits that Achilles can never catch, let alone pass, a tortoise that is given a head start in a race. Here Zeno discretizes time and space in a similar manner to the runner paradox by arguing that *when* Achilles reaches the place where the tortoise started, the tortoise will have moved on some distance; Achilles will then have to catch up to the new location of the tortoise. Thus, Achilles will have to exist in an infinite number of points in time, all of which take place while Achilles is behind the tortoise.

3. **The arrow.** Imagine an arrow in flight. At any given *instant* of time the arrow rests at a specific location in space. An arrow cannot move in an instant, so how does it change its spatial location during an infinite sum of instants?

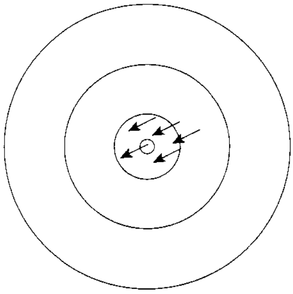
4. The stadium. Zeno's fourth paradox describes a person standing still in a stadium at point A. Two other people are running at the same speed toward A from opposite sides, west or east. To each other, the runners appear to be traveling at twice the speed they appear to the stationary person (which is "impossible").

These paradoxes, especially the first three, rest in part on the apparent incommensurability of discrete and continuous reality (the fourth actually anticipates Einstein's 20th-century arguments about the intrinsic dependence of space and time). We find them intellectually enriching and entertaining to ponder. They point to some of the deep conceptual and philosophical questions inherent in the subject matter of geography, even subject matter that most nongeographers probably think is obvious and not at all controversial. Just how many lakes are there in Minnesota?

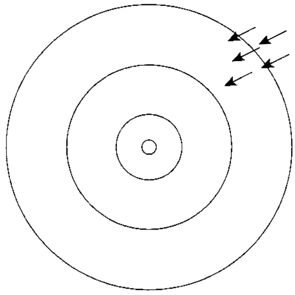
because it reads off weight to the nearest tenth of a pound but measure inaccurately because, unknown to you, it reads off weight 30 pounds too heavy (for some of us, such inaccurate scales are apparently to be found everywhere).

Accuracy and precision are thus, in an important sense, separate issues, but they are intimately related. Accuracy is the correctness of measurement *at a given level of precision.* It is perfectly accurate to say that the average adult man weighs 200 pounds, as long as you recognize that the statement is precise to the nearest 100 pounds; that is, the measurement can be expressed only in units of 100 pounds. On the other hand, precision is the smallest resolution of measurement *that produces accurate digits in the measured score.* It is not more precise to say that the Nile River is 6,652.327 km long rather than 6,652 km, unless the .327 km can be measured accurately (it cannot be). Such false or **spurious** precision is unfortunately common, probably because of the tendency of computers to output numbers with many more decimal places than are actually warranted by the quality of the data. Perhaps some people also hold a mistaken belief that greater numerical precision is necessarily a sign of more "scientific" work. In any case, one should not report a data value with greater precision than the measurement procedure warrants—the precision it produces that is accurate. When working with summary indices (see Chapter 9), such as the mean or variance, acceptable precision is typically considered to be one digit more precise than the precision of the original data values.[7]

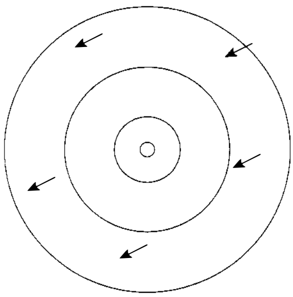
[7]This **rounding** advice is based on the typical situation in which the most precise digit of a measured score is halfway between the lower and upper ranges of possible values. For example, a mountain that is accurately described as 3,528 meters high is actually somewhere between 3,527.5 and 3,528.5 meters high. You should not apply our rounding advice if this situation does not hold. For example, we note that when people report their age in years, they do not follow the usual measurement convention. A person who accurately says he or she is "46" really means "between 46.0 and 47.0" not "between 45.5 and 46.5." You could treat a 46 as a 46.5 for averaging, or better yet, request date of birth.



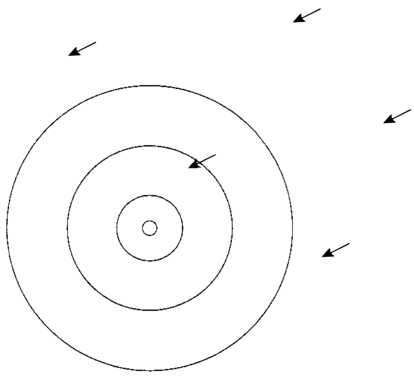
a. High accuracy and high precision.



b. Low accuracy and high precision.



c. High accuracy and low precision.



d. Low accuracy and low precision.

Figure 2.1 Measurement accuracy and precision depicted by analogy to darts thrown at a target. The cluster of five darts thrown at each target should be thought of as one measurement, and the bull's-eye is the actual value of the construct being measured. The distance from the spatial center (centroid) of the darts to the bull's-eye is accuracy; the spread of the five darts around their centroid is precision. The game of darts is but one of many situations where high accuracy and high precision together make a winning combination.

The Concept of Scale in Geography

There is one concept that is especially central to geographic inquiry—the concept of **scale**. Although scale is relevant to all natural and social science disciplines, perhaps no discipline is more sensitive to its implications than geography is. Scale is both an idea concept and an empirical concept. Scale has many implications for research in geography, and we discuss specific aspects of it in several chapters of this book. But it is so fundamental to geography that it deserves to be introduced here.

Scale is about size, either relative or absolute. Scale is relevant not just to space in geography, but also to time and theme (**themes** are the nonspatial and nontemporal characteristics of human and natural phenomena that geographers measure and map as variables). Scale has several meanings in geography, which can be confusing. We group these meanings into three categories: **phenomenon scale**, **analysis**

scale, and **cartographic scale**. Phenomenon scale refers to the size at which human or physical earth structures or processes actually exist, regardless of how they are studied or represented. A lake is larger than a pond (in the English language anyway), and a city is larger than a city block. Analysis scale refers to the size of the unit at which some problem is analyzed. Data at the state level are at a larger scale than data at the county level. Finally, cartographic scale refers to the depicted size of a feature on a map relative to its actual size in the world. (Because cartographic scale is expressed in terms of map size relative to earth size, a "small-scale" map shows a large earth area; most people who are not cartographers find this a little confusing.) Although the three meanings of scale are frequently treated independently, they are in fact interrelated in important ways that are relevant to all geographers and the focus of research for some geographers. We examine several of these interrelations throughout the rest of this book.

It is widely recognized that various scales of geographic phenomena interact, or that phenomena at one scale emerge from smaller- or larger-scale phenomena. This is captured by the notion of a **hierarchy of scales**, in which smaller phenomena are nested within larger phenomena. For example, local economies are nested within regional economies, which are in turn nested within the global economy. Conceptualizing and modeling such scale hierarchies and "couplings" across scales can be quite difficult; for this reason, much geographic work continues the practice of focusing on a single scale.

Many geographers have claimed that we can partially define the discipline of geography—the study of the earth as the home of humanity—by its focus on phenomena at certain scales, such as cities or continents, and not other scales. The range of scales of interest to geographers is often summarized by the use of terminological continua such as "local/global" or "micro-, meso-, and macroscale." Not everyone shares the view that geographers must restrict their focus to particular ranges of scales, however, and advances have and will continue to occur when geographers stretch the boundaries of their subject matter. Nonetheless, few would argue that subatomic or interplanetary scales are properly of concern for geography.

Generating Research Ideas

We finish this chapter by discussing some approaches to generating research ideas. A beginning researcher may wonder where scientists get their ideas. The short answer to this question is . . . *anywhere!* Although we offer some tips here that can help, generating research ideas is a creative component of scientific activity that cannot and should not be entirely formularized. You can get ideas from your intuition, your dreams, from movies or books, or from your personal experiences. You can get them from your neighbor, or your aunt or uncle. You can get them from eating too many chili peppers. However, we can identify a few more systematic approaches[8] that may help you generate research ideas, especially if you are new to designing research:

[8]Adapted from McGuire, W. J. (1973). The yin and yang of progress in social psychology: Seven koan. *Journal of Personality and Social Psychology, 26*, 446–456.

1. Intensive case study. Look closely at a particular marsh, including its shape, size, and depth; its water temperature, clarity, and chemical makeup; and its flora and fauna. This could lead to research on the functioning of aquatic ecosystems.

2. Paradoxical incident. Notice that families often return to hazardous areas after a disaster such as a flood or fire. This could lead to research on the variables that influence residential choice or responses to risky events.

3. Analogical extension. Identify an analogy between people's tendency to shop at closer stores over more distant stores and the greater pull that close planets have over more distant planets (of the same mass). This could lead to research on the "social gravity model" of spatial interaction.

4. Practitioner's rule of thumb. Examine the choices professional mapmakers make when designing topographic maps. This could lead to research on which cartographic variables are more or less effective at communicating relief.

5. Account for conflicting results. Observe that plants of a certain species grow on the sunny face of mountains in one part of the world but on the shady side in another. This could lead to research on the factors that affect plant growth other than insolation, like soil or wind.

6. Reduce complexity to simpler components. Break down a person's daily activity patterns into components like work, shopping, recreation, and so on. This could lead to a model of how commuters organize their travel at different times of the day.

7. Account for exceptions to general findings. This is useful to pursue whenever you find it.

Of course, we don't just want to get a research idea, we want to get a good one—interesting, novel, relevant, and feasible. Identifying ideas like this requires the kind of expertise that hopefully develops over time with experience as a research scientist. But we can help get you started by offering a plan of action for generating and pursuing *good* research ideas, and implementing them as research:

1. Find a research area; focus on what interests you.
2. Generate research ideas, first on your own; avoid groupthink and staleness by not referring to literature or experts right away.
3. Link with other knowledge you already have; is your idea plausible?
4. Check existing literature; ask experts.
5. Formulate your idea as one or more specific hypotheses.
6. Design research to address your hypotheses.

Review Questions

Idea Concepts

- What are the idea concepts of theory, hypothesis, causality, model, and construct?
- What are some historical and contemporary ideas within the philosophy of causality?
- What are the distinctions between probabilistic and deterministic causality; necessary and sufficient causality; mechanistic and functional causality?

Empirical Concepts

- What are the empirical concepts of case, variable, measurement, data, accuracy, and precision?
- What are the four measurement levels for variables, and why is the measurement level of a variable important?
- What is the distinction between discrete and continuous variables, and what are some ways the distinction is important?
- What is spurious precision, and what is a general rule for reporting data with appropriate precision?

The Concept of Scale in Geography

- To what does the concept of scale in geography refer?
- What are phenomenon, analysis, and cartographic scale?

Generating Research Ideas

- What are some general strategies for generating research ideas?
- What are steps of a plan of action for developing good research ideas?

Key Terms

accuracy: the correctness of values resulting from a particular measurement process, at a particular level of precision

analysis scale: the size of the unit at which some problem is analyzed

cartographic scale: the depicted size of a feature on a map relative to its actual size in the world; unlike other meanings of scale, a "small-scale" map shows a large earth area, and a "large-scale" map shows a small earth area

case: the thing or entity a scientist studies; synonyms include unit of analysis, entity, element, individual, research subject, respondent, and (in physical geography) sample

causality: the concept that the occurrence of one state or event can bring about another state or event

cause: antecedent state or event that brings about an effect

classification: grouping entities into classes or categories based on some type of similarity of class members to each other or to a standard, and some type of dissimilarity between class members and nonmembers

constant: attributes or properties of cases that do not vary from case to case but take on a single value; in contrast to variables, which vary across cases

construct: concept that is a piece of the idealized world comprising the subject matter of theories; the hypothetical entities that we attempt to measure when we perform our systematic empirical observations

continuous variable: variable that can take on an infinite number of possible values between any two values (assuming unlimited measurement precision)

countably infinite variable: discrete variable that can take on an infinite number of possible values but not between any two values, only toward positive or negative infinity

data: the values obtained by measurement that constitute empirical evidence in a study; data are analyzed, interpreted, and displayed by scientists

deterministic: causal processes that necessarily bring about effects, or relationships that always hold, at every occurrence

dichotomous variable: simplest possible variable, it takes on only two values across cases

discrete variable: variable that can take on only a limited set of distinct possible values (even assuming unlimited measurement precision)

effect: a subsequent state or event brought about by a cause

empirical concepts in science: scientific concepts that directly refer to empirical observations of reality; they include cases, variables, measurement, measurement levels, discrete versus continuous variables, and accuracy versus precision of measurement

functional causality: the idea that causes can follow effects, providing goal states for the effects; often used heuristically by scientists but not literally

hierarchy of scales: the fact that geographic phenomena at different scales often interact, existing in nested and nesting relationships to one another

hypothesis: idea or conjecture about a pattern of observations of the world; similar to theory but more specific and directly testable

idea concepts in science: scientific concepts that directly refer to ideas about reality; they include theory, law, hypothesis, causality, model, and construct

interval measurement: the third of the four levels of measurement; it expresses quantitative distance between scale values, but not an absolute zero point

latent variable: hypothetical entity that we attempt to measure when we perform our systematic empirical observations; synonym of construct

law: mathematical expression describing a quantitative relationship that is expected to hold precisely in an "ideal" world, but is not explanatory in and of itself; the law of gravity is an example

manifest variable: actual entity expressed by our measurements when we perform systematic empirical observations; synonym of measured variable

measurement: assigning numbers or other symbols to cases to reflect their values on a variable

measurement level: typology of types of variables based on the quantitative content they express—the degree and nature of quantification implied by a measurement; from the least to the most quantitative, the four levels are nominal, ordinal, interval, ratio

mechanistic causality: the idea that causes move forward "densely" in space and time, with continuously connected causes and effects

metric measurement: either interval or ratio measurement, which express quantitative distance between scale values

model: simplified representation of a portion of reality, expressed in conceptual, physical, graphical, or computational form

necessary cause: cause that must be in place for the effect to occur, but by itself it may not be enough to make the effect occur

nominal measurement: the first of the four levels of measurement; it expresses no quantitative information at all, only classification or naming

operational definition: defining a variable by describing the techniques (operations) used to measure it

ordinal measurement: the second of the four levels of measurement; it expresses only rank order

phenomenon scale: the size at which some human or physical earth structure or process actually exists, regardless of how it is studied or represented

precision of measurement: the sharpest or highest resolution of accurate values resulting from a particular measurement process

probabilistic: causal processes that sometimes bring about effects, or relationships that sometimes hold, but not at every occurrence; same as "stochastic"

ratio measurement: the fourth of the four levels of measurement; it expresses quantitative distance between scale values and an absolute zero point, which allows for the valid creation of ratios among scale values

rounding: reduction in the precision of measured or calculated values that is done in order to avoid excessive or unnecessary levels of precision

scale: both an idea concept and an empirical concept that concerns size, either relative or absolute; besides spatial scale, temporal and thematic scale are also relevant to geography

spurious precision: precision in measured or calculated values that exceeds the accurate precision actually present

sufficient cause: cause that by itself will make the effect occur, but it may not need to be in place for the effect to occur

themes: the nonspatial and nontemporal characteristics of human and natural phenomena that geographers measure and map as variables

theory: idea or conjecture about a causal relationship that provides an answer to the question of "why" something is the way it is; sometimes used broadly to refer to any conjecture

variable: attributes or properties of cases that researchers measure and study; their value varies from case to case, in contrast to constants, which take a single value across cases

Bibliography

Clugston, M. J. (Ed.) (2004). *The new Penguin dictionary of science* (2nd ed.). London: Penguin.

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

Haines-Young, R. H., & Petch, J. R. (1986). *Physical geography: Its nature and methods.* London: Paul Chapman.

Hudson, J. C. (1992). Scale in space and time. In R. F. Abler, M. G. Marcus, & J. M. Olson (Eds.), *Geography's inner worlds: Pervasive themes in contemporary American geography* (pp. 280–298). New Brunswick, NJ: Rutgers University Press.

McGraw-Hill (1997). *McGraw-Hill encyclopedia of science & technology* (8th ed.). New York: McGraw-Hill.

Montello, D. R. (2001). Scale, in geography. In N. J. Smelser & P. B. Baltes (Eds.), *International encyclopedia of the social & behavioral sciences* (pp. 13501–13504). Oxford: Pergamon.

Rosenthal, R., & Rosnow, R. L. (1991). *Essentials of behavioral research: Methods and data analysis* (2nd ed.). New York: McGraw-Hill.