



This is “Defining and Measuring Concepts”, chapter 6 from the book Sociological Inquiry Principles: Qualitative and Quantitative Methods ([index.html](#)) (v. 1.0).

This book is licensed under a Creative Commons [by-nc-sa 3.0](http://creativecommons.org/licenses/by-nc-sa/3.0/) (<http://creativecommons.org/licenses/by-nc-sa/3.0/>) license. See the license for more details, but that basically means you can share this book as long as you credit the author (but see below), don't make money from it, and do make it available to everyone else under the same terms.

This content was accessible as of December 29, 2012, and it was downloaded then by Andy Schmitz (<http://lardbucket.org>) in an effort to preserve the availability of this book.

Normally, the author and publisher would be credited here. However, the publisher has asked for the customary Creative Commons attribution to the original publisher, authors, title, and book URI to be removed. Additionally, per the publisher's request, their name has been removed in some passages. More information is available on this project's attribution page (http://2012books.lardbucket.org/attribution.html?utm_source=header).

For more information on the source of this book, or why it is available for free, please see the project's home page (<http://2012books.lardbucket.org/>). You can browse or download additional books there.

Chapter 6

Defining and Measuring Concepts

Measurement, Conceptualization, and Operationalization

In this chapter we'll discuss measurement, conceptualization, and operationalization. If you're not quite sure what any of those words mean, or even how to pronounce them, no need to worry. By the end of the chapter, you should be able to wow your friends and family with your newfound knowledge of these three difficult to pronounce, but relatively simple to grasp, terms.

6.1 Measurement

LEARNING OBJECTIVES

1. Define measurement.
2. Describe Kaplan's three categories of the things that social scientists measure.
3. Identify the stages at which measurement is important.

Measurement is important. Recognizing that fact, and respecting it, will be of great benefit to you—both in research methods and in other areas of life as well. If, for example, you have ever baked a cake, you know well the importance of measurement. As someone who much prefers rebelling against precise rules over following them, I once learned the hard way that measurement matters. A couple of years ago I attempted to bake my husband a birthday cake without the help of any measuring utensils. I'd baked before, I reasoned, and I had a pretty good sense of the difference between a cup and a tablespoon. How hard could it be? As it turns out, it's not easy guesstimating precise measures. That cake was the lumpiest, most lopsided cake I've ever seen. And it tasted kind of like Play-Doh. [Figure 6.1](#) depicts the monstrosity I created, all because I did not respect the value of measurement.

Figure 6.1



Measurement is important in baking and in research.

Just as measurement is critical to successful baking, it is as important to successfully pulling off a social scientific research project. In sociology, when we use the term **measurement**¹ we mean the process by which we describe and ascribe meaning to the key facts, concepts, or other phenomena that we are investigating. At its core, measurement is about defining one's terms in as clear and precise a way as possible. Of course, measurement in social science isn't quite as simple as using some predetermined or universally agreed-on tool, such as a measuring cup or spoon, but there are some basic tenants on which most social scientists agree when it comes to measurement. We'll explore those as well as some of the ways that measurement might vary depending on your unique approach to the study of your topic.

What Do Social Scientists Measure?

1. The process by which we describe and ascribe meaning to the key facts, concepts, or phenomena that we are investigating.

The question of what social scientists measure can be answered by asking oneself what social scientists study. Think about the topics you've learned about in other sociology classes you've taken or the topics you've considered investigating yourself. Or think about the many examples of research you've read about in this

text. In [Chapter 2 "Linking Methods With Theory"](#) we learned about Melissa Milkie and Catharine Warner's study (2011) Milkie, M. A., & Warner, C. H. (2011). Classroom learning environments and the mental health of first grade children. *Journal of Health and Social Behavior*, 52, 4–22. of first graders' mental health. In order to conduct that study, Milkie and Warner needed to have some idea about how they were going to measure mental health. What does mental health mean, exactly? And how do we know when we're observing someone whose mental health is good and when we see someone whose mental health is compromised? Understanding how measurement works in research methods helps us answer these sorts of questions.

As you might have guessed, social scientists will measure just about anything that they have an interest in investigating. For example, those who are interested in learning something about the correlation between social class and levels of happiness must develop some way to measure both social class and happiness. Those who wish to understand how well immigrants cope in their new locations must measure immigrant status and coping. Those who wish to understand how a person's gender shapes their workplace experiences must measure gender and workplace experiences. You get the idea. Social scientists can and do measure just about anything you can imagine observing or wanting to study. Of course, some things are easier to observe, or measure, than others, and the things we might wish to measure don't necessarily all fall into the same category of measureables.

In 1964, philosopher Abraham Kaplan (1964) Kaplan, A. (1964). *The conduct of inquiry: Methodology for behavioral science*. San Francisco, CA: Chandler Publishing Company. wrote what has since become a classic work in research methodology, *The Conduct of Inquiry* (Babbie, 2010). Earl Babbie offers a more detailed discussion of Kaplan's work in his text. You can read it in [Chapter 5 "Research Design"](#) of the following: Babbie, E. (2010). *The practice of social research* (12th ed.). Belmont, CA: Wadsworth. In his text, Kaplan describes different categories of things that behavioral scientists observe. One of those categories, which Kaplan called "observational terms," is probably the simplest to measure in social science. **Observational terms**² are the sorts of things that we can see with the naked eye simply by looking at them. They are terms that "lend themselves to easy and confident verification" (1964, p. 54). Kaplan, A. (1964). *The conduct of inquiry: Methodology for behavioral science*. San Francisco, CA: Chandler Publishing Company, p. 54. If, for example, we wanted to know how the conditions of playgrounds differ across different neighborhoods, we could directly observe the variety, amount, and condition of equipment at various playgrounds.

2. Things that we can see with the naked eye simply by looking at them.

3. Things that we cannot see with the naked eye but that require some more complex assessment.

Indirect observables³, on the other hand, are less straightforward to assess. They are "terms whose application calls for relatively more subtle, complex, or indirect observations, in which inferences play an acknowledged part. Such inferences concern presumed connections, usually causal, between what is directly observed

and what the term signifies” (1964, p. 55). Kaplan, A. (1964). *The conduct of inquiry: Methodology for behavioral science*. San Francisco, CA: Chandler Publishing Company, p. 55. If we conducted a study for which we wished to know a person’s income, we’d probably have to ask them their income, perhaps in an interview or a survey. Thus we have observed income, even if it has only been observed indirectly. Birthplace might be another indirect observable. We can ask study participants where they were born, but chances are good we won’t have directly observed any of those people being born in the locations they report.

Figure 6.2



Observational terms, such as playground equipment and conditions, can be seen with the naked eye.

© Thinkstock

Figure 6.3



Indirect observables, such as birthplace, may require some more complex assessment than simply seeing them with the naked eye.

© Thinkstock

Sometimes the measures that we are interested in are more complex and more abstract than observational terms or indirect observables. Think about some of the concepts you've learned about in other sociology classes—ethnocentrism, for example. What is ethnocentrism? Well, you might know from your intro to sociology class that it has something to do with the way a person judges another's culture. But how would you *measure* it? Here's another construct: bureaucracy. We know this term has something to do with organizations and how they operate, but measuring such a construct is trickier than measuring, say, a person's income. In both cases, ethnocentrism and bureaucracy, these theoretical notions represent ideas whose meaning we have come to agree on. Though we may not be able to observe these abstractions directly, we can observe the confluence of things that they are made up of. Kaplan referred to these more abstract things that behavioral scientists measure as **constructs**⁴. Constructs are “not observational either directly or indirectly” (1964, p. 55), Kaplan, A. (1964). *The conduct of inquiry: Methodology for behavioral science*. San Francisco, CA: Chandler Publishing Company, p. 55. but they can be defined based on observables.

4. Abstractions that cannot be observed directly but that can be defined based on that which is observable.

Thus far we have learned that social scientists measure what Abraham Kaplan called observational terms, indirect observables, and constructs. These terms refer to the different sorts of things that social scientists may be interested in measuring. But *how* do social scientists measure these things? That is the next question we'll tackle.

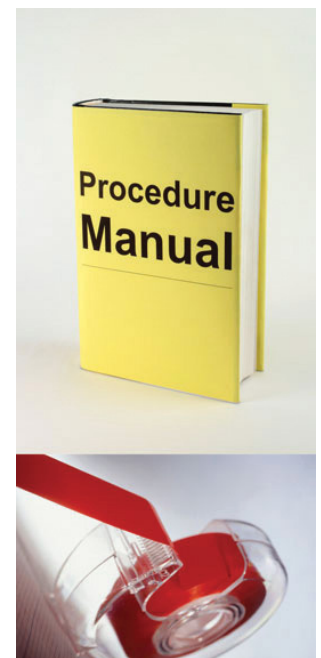
How Do Social Scientists Measure?

Measurement in social science is a process. It occurs at multiple stages of a research project: in the planning stages, in the data collection stage, and sometimes even in the analysis stage. Recall that previously we defined measurement as the process by which we describe and ascribe meaning to the key facts, concepts, or other phenomena that we are investigating. Once we've identified a research question, we begin to think about what some of the key ideas are that we hope to learn from our project. In describing those key ideas, we begin the measurement process.

Let's say that our research question is the following: How do new college students cope with the adjustment to college? In order to answer this question, we'll need to some idea about what coping means. We may come up with an idea about what coping means early in the research process, as we begin to think about what to look for (or observe) in our data-collection phase. Once we've collected data on coping, we also have to decide how to report on the topic. Perhaps, for example, there are different types or dimensions of coping, some of which lead to more successful adjustment than others. However we decide to proceed, and whatever we decide to report, the point is that measurement is important at each of these phases.

As the preceding paragraph demonstrates, measurement is a process in part because it occurs at multiple stages of conducting research. We could also think of measurement as a process because of the fact that measurement in itself involves multiple stages. From identifying one's key terms to defining them to figuring out how to observe them and how to know if our observations are any good, there are multiple steps involved in the measurement process. An additional step in the measurement process involves deciding what elements one's measures contain. A measure's elements might be very straightforward and clear, particularly if they

Figure 6.4



Constructs such as bureaucracy are more abstract than either observational terms or indirect observables, but we can detect them based on the observation of some collection of observables.

© Thinkstock

are directly observable. Other measures are more complex and might require the researcher to account for different themes or types. These sorts of complexities require paying careful attention to a concept's level of measurement and its dimensions. We'll explore these complexities in greater depth at the end of this chapter, but first let's look more closely at the early steps involved in the measurement process.

KEY TAKEAWAYS

- Measurement is the process by which we describe and ascribe meaning to the key facts, concepts, or other phenomena that we are investigating.
- Kaplan identified three categories of things that social scientists measure including observational terms, indirect observables, and constructs.
- Measurement occurs at all stages of research.

EXERCISE

1. See if you can come up with one example of each of the following: an observational term, an indirect observable, and a construct. How might you measure each?

6.2 Conceptualization

LEARNING OBJECTIVES

1. Define concept.
2. Describe why defining our concepts is important.
3. Describe how conceptualization works.
4. Define dimensions in terms of social scientific measurement.
5. Describe reification.

In this section we'll take a look at one of the first steps in the measurement process, conceptualization. This has to do with defining our terms as clearly as possible and also not taking ourselves too seriously in the process. Our definitions mean only what we say they mean—nothing more and nothing less. Let's talk first about how to define our terms, and then we'll examine what I mean about not taking ourselves (or our terms, rather) too seriously.

Concepts and Conceptualization

So far the word *concept* has come up quite a bit, and it would behoove us to make sure we have a shared understanding of that term. A **concept**⁵ is the notion or image that we conjure up when we think of some cluster of related observations or ideas. For example, masculinity is a concept. What do you think of when you hear that word? Presumably you imagine some set of behaviors and perhaps even a particular style of self presentation. Of course, we can't necessarily assume that everyone conjures up the same set of ideas or images when they hear the word *masculinity*. In fact, there are many possible ways to define the term. And while some definitions may be more common or have more support than others, there isn't one true, always-correct-in-all-settings definition. What counts as masculine may shift over time, from culture to culture, and even from individual to individual (Kimmel, 2008). Kimmel, M. (2008). Masculinity. In W. A. Darity Jr. (Ed.), *International encyclopedia of the social sciences* (2nd ed., Vol. 5, pp. 1–5). Detroit, MI: Macmillan Reference USA. This is why defining our concepts is so important.

You might be asking yourself why you should bother defining a term for which there is no single, correct definition. Believe it or not, this is true for any concept you might measure in a sociological study—there is never a single, always-correct definition. When we conduct empirical research, our terms mean only what we say they mean—nothing more and nothing less. There's a *New Yorker* cartoon that aptly

5. The notion or image that we conjure up when we think of some cluster of related observations or ideas.

represents this idea (<http://www.cartoonbank.com/1998/it-all-depends-on-how-you-define-chop/invt/117721>). It depicts a young George Washington holding an ax and standing near a freshly chopped cherry tree. Young George is looking up at a frowning adult who is standing over him, arms crossed. The caption depicts George explaining, “It all depends on how you define ‘chop.’” Young George Washington gets the idea—whether he actually chopped down the cherry tree depends on whether we have a shared understanding of the term *chop*. Without a shared understanding of this term, our understandings of what George has just done may differ. Likewise, without understanding how a researcher has defined her or his key concepts, it would be nearly impossible to understand the meaning of that researcher’s findings and conclusions. Thus any decision we make based on findings from empirical research should be made based on full knowledge not only of how the research was designed, as described in [Chapter 5 "Research Design"](#), but also of how its concepts were defined and measured.

So how do we define our concepts? This is part of the process of measurement, and this portion of the process is called **conceptualization**⁶. Conceptualization involves writing out clear, concise definitions for our key concepts. Sticking with the previously mentioned example of masculinity, think about what comes to mind when you read that term. How do you know masculinity when you see it? Does it have something to do with men? With social norms? If so, perhaps we could define masculinity as the social norms that men are expected to follow. That seems like a reasonable start, and at this early stage of conceptualization, brainstorming about the images conjured up by concepts and playing around with possible definitions is appropriate. But this is just the first step. It would make sense as well to consult other previous research and theory to understand if other scholars have already defined the concepts we’re interested in. This doesn’t necessarily mean we must use their definitions, but understanding how concepts have been defined in the past will give us an idea about how our conceptualizations compare with the predominant ones out there. Understanding prior definitions of our key concepts will also help us decide whether we plan to challenge those conceptualizations or rely on them for our own work.

Figure 6.5



Just as the young George Washington, depicted in the cartoon described previously, makes the point that “it all depends on how you define ‘chop,’” sociological researchers understand that how one defines one’s terms will shape the conclusions one is able to draw.

© Thinkstock

6. The process of defining key terms or concepts.

If we turn to the literature on masculinity, we will surely come across work by Michael Kimmel, one of the preeminent masculinity scholars in the United States. After consulting Kimmel's prior work (2000; 2008), Kimmel, M. (2000). *The gendered society*. New York, NY: Oxford University Press; Kimmel, M. (2008). Masculinity. In W. A. Darity Jr. (Ed.), *International encyclopedia of the social sciences* (2nd ed., Vol. 5, pp. 1–5). Detroit, MI: Macmillan Reference USA. we might tweak our initial definition of masculinity just a bit. Rather than defining masculinity as “the social norms that men are expected to follow,” perhaps instead we’ll define it as “the social roles, behaviors, and meanings prescribed for men in any given society at any one time.” Our revised definition is both more precise and more complex. Rather than simply addressing one aspect of men’s lives (norms), our new definition addresses three aspects: roles, behaviors, and meanings. It also implies that roles, behaviors, and meanings may vary across societies and over time. Thus, to be clear, we’ll also have to specify the particular society and time period we’re investigating as we conceptualize masculinity.

As you can see, conceptualization isn’t quite as simple as merely applying any random definition that we come up with to a term. Sure, it may involve some initial brainstorming, but conceptualization goes beyond that. Once we’ve brainstormed a bit about the images a particular word conjures up for us, we should also consult prior work to understand how others define the term in question. And after we’ve identified a clear definition that we’re happy with, we should make sure that every term used in our definition will make sense to others. Are there terms used within our definition that also need to be defined? If so, our conceptualization is not yet complete. And there is yet another aspect of conceptualization to consider: concept dimensions. We’ll consider that aspect along with an additional word of caution about conceptualization next.

A Word of Caution About Conceptualization

So now that we’ve come up with a clear definition for the term *masculinity* and made sure that the terms we use in our definition are equally clear, we’re done, right? Not so fast. If you’ve ever met more than one man in your life, you’ve probably noticed that they are not all exactly the same, even if they live in the same society and at the same historical time period. This could mean that there are dimensions of masculinity. In terms of social scientific measurement, concepts can be said to have **dimensions**⁷ when there are multiple elements that make up a single concept. With respect to the term *masculinity*, dimensions could be regional (Is masculinity defined differently in different regions of the same country?), age based (Is masculinity defined differently for men of different ages?), or perhaps power based (Are some forms of masculinity valued more than others?). In any of these cases, the concept masculinity would be considered to have multiple dimensions. While it isn’t necessarily a must to spell out every possible dimension of the concepts you

7. The multiple elements of a single concept.

wish to measure, it may be important to do so depending on the goals of your research. The point here is to be aware that some concepts have dimensions and to think about whether and when dimensions may be relevant to the concepts you intend to investigate.

Before we move on to the additional steps involved in the measurement process, it would be wise to caution ourselves about one of the dangers associated with conceptualization. While I've suggested that we should consult prior scholarly definitions of our concepts, it would be wrong to assume that just because prior definitions exist that they are any more real than whatever definitions we make up (or, likewise, that our own made-up definitions are any more real than any other definition). It would also be wrong to assume that just because definitions exist for some concept that the concept itself exists beyond some abstract idea in our heads. This idea, assuming that our abstract concepts exist in some concrete, tangible way, is known as **reification**⁸.

To better understand reification, take a moment to think about the concept of social structure. This concept is central to sociological thinking. When we sociologists talk about social structure, we are talking about an abstract concept. Social structures shape our ways of being in the world and of interacting with one another, but they do not exist in any concrete or tangible way. A social structure isn't the same thing as other sorts of structures, such as buildings or bridges. Sure, both types of structures are important to how we live our everyday lives, but one we can touch, and the other is just an idea that shapes our way of living.

Here's another way of thinking about reification: Think about the term *family*. If you were interested in studying this concept, we've learned that it would be good to consult prior theory and research to understand how the term has been conceptualized by others. But we should also question past conceptualizations. Think, for example, about where we'd be today if we used the same definition of family that was used, say, 100 years ago. How have our understandings of this concept changed over time? What role does conceptualization in social scientific research play in our cultural understandings of terms like family? The point is that our terms mean nothing more and nothing less than whatever definition we assign to them. Sure, it makes sense to come to some social agreement about what various concepts mean. Without that agreement, it would be difficult to navigate through everyday living. But at the same time, we should not forget that we have assigned those definitions and that they are no more real than any other, alternative definition we might choose to assign.

8. Assuming that abstract concepts exist in some concrete, tangible way.

KEY TAKEAWAYS

- Conceptualization is a process that involves coming up with clear, concise definitions.
- Some concepts have multiple elements or dimensions.
- Just because definitions for abstract concepts exist does not mean that the concept is tangible or concrete.

EXERCISES

1. Conceptualize the term *discipline* and identify possible dimensions of the term. Have someone who is in the class with you do the same thing (without seeing your conceptualization). Now compare what you each came up with. How do your conceptualizations and dimensions differ, and why?
2. Identify a concept that is important in your area of interest. Challenge yourself to conceptualize the term without first consulting prior literature. Now consult prior work to see how your concept has been conceptualized by others. How and where does your conceptualization differ from others? Are there dimensions of the concept that you or others hadn't considered?

6.3 Operationalization

LEARNING OBJECTIVES

1. Describe how operationalization works.
2. Define and give an example of indicators.

Now that we have figured out how to define, or conceptualize, our terms we'll need to think about operationalizing them. **Operationalization**⁹ is the process by which we spell out precisely how a concept will be measured. It involves identifying the specific research procedures we will use to gather data about our concepts. This of course requires that one know what research method(s) he or she will employ to learn about her or his concepts, and we'll examine specific research methods in [Chapter 8 "Survey Research: A Quantitative Technique"](#) through [Chapter 12 "Other Methods of Data Collection and Analysis"](#). For now, let's take a broad look at how operationalization works. We can then revisit how this process works when we examine specific methods of data collection in later chapters.

Indicators

Operationalization works by identifying specific **indicators**¹⁰ that will be taken to represent the ideas that we are interested in studying. If, for example, we are interested in studying masculinity, indicators for that concept might include some of the social roles prescribed to men in society such as breadwinning or fatherhood. Being a breadwinner or a father might therefore be considered *indicators* of a person's masculinity. The extent to which a man fulfills either, or both, of these roles might be understood as clues (or indicators) about the extent to which he is viewed as masculine.

Let's look at another example of indicators. Each day, Gallup researchers poll 1,000 randomly selected Americans to ask them about their well-being. To measure well-being, Gallup asks these people to respond to questions covering six broad areas: physical health, emotional health, work environment, life evaluation, healthy behaviors, and access to basic necessities. Gallup uses these six factors as indicators of the concept that they are really interested in: well-being (<http://www.gallup.com/poll/123215/Gallup-Healthways-Index.aspx>).

Identifying indicators can be even simpler than the examples described thus far. What are the possible indicators of the concept of gender? Most of us would

9. The process by which we spell out precisely how a concept will be measured.

10. Empirical observations taken to represent the ideas that we are interested in studying.

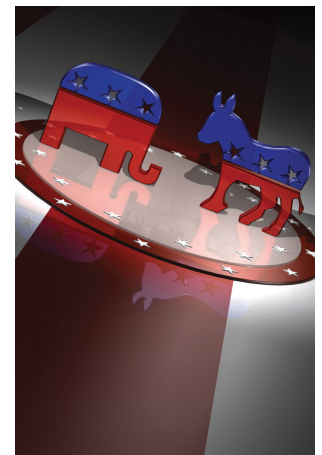
probably agree that “woman” and “man” are both reasonable indicators of gender, and if you’re a sociologist of gender, like me, you might also add an indicator of “other” to the list. Political party is another relatively easy concept for which to identify indicators. In the United States, likely indicators include Democrat and Republican and, depending on your research interest, you may include additional indicators such as Independent, Green, or Libertarian as well. Age and birthplace are additional examples of concepts for which identifying indicators is a relatively simple process. What concepts are of interest to you, and what are the possible indicators of those concepts?

We have now considered a few examples of concepts and their indicators but it is important that we don’t make the process of coming up with indicators *too* arbitrary or casual. One way to avoid taking an overly casual approach in identifying indicators, as described previously, is to turn to prior theoretical and empirical work in your area. Theories will point you in the direction of relevant concepts and possible indicators; empirical work will give you some very specific examples of how the important concepts in an area have been measured in the past and what sorts of indicators have been used. Perhaps it makes sense to use the same indicators as researchers who have come before you. On the other hand, perhaps you notice some possible weaknesses in measures that have been used in the past that your own methodological approach will enable you to overcome. Speaking of your methodological approach, another very important thing to think about when deciding on indicators and how you will measure your key concepts is the strategy you will use for data collection. A survey implies one way of measuring concepts, while field research implies a quite different way of measuring concepts. Your data-collection strategy will play a major role in shaping how you operationalize your concepts.

Putting It All Together

Moving from identifying concepts to conceptualizing them and then to operationalizing them is a matter of increasing specificity. You begin with a general interest, identify a few concepts that are essential for studying that interest, work to define those concepts, and then spell out precisely how you will measure those concepts. Your focus becomes narrower as you move from a general interest to

Figure 6.6

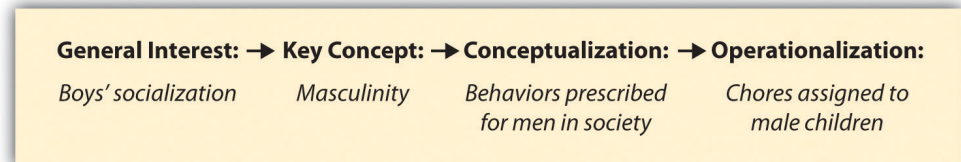


“Republican” and “Democrat” are both indicators of the concept political party.

© Thinkstock

operationalization. The process looks something like that depicted in [Figure 6.7 "The Process of Measurement"](#). Here, the researcher moves from a broader level of focus to a more narrow focus. The example provided in italics in the figure indicates what this process might look like for a researcher interested in studying the socialization of boys into their roles as men.

Figure 6.7 *The Process of Measurement*



One point not yet mentioned is that while the measurement process often works as outlined in [Figure 6.7 "The Process of Measurement"](#), it doesn't necessarily always have to work out that way. What if your interest is in discovering how people define the same concept differently? If that's the case, you probably *begin* the measurement process the same way as outlined earlier, by having some general interest and identifying key concepts related to that interest. You might even have some working definitions of the concepts you wish to measure. And of course you'll have some idea of how you'll go about discovering how your concept is defined by different people. But you may not go so far as to have a clear set of indicators identified before beginning data collection, for that would defeat the purpose if your aim is to *discover* the variety of indicators people rely on.

Let's consider an example of when the measurement process may not work out exactly as depicted in [Figure 6.7 "The Process of Measurement"](#). One of my early research projects (Blackstone, 2003) Blackstone, A. (2003). *Racing for the cure and taking back the night: Constructing gender, politics, and public participation in women's activist/volunteer work*. PhD dissertation, Department of Sociology, University of Minnesota, Minneapolis, MN. was a study of activism in the breast cancer movement compared to activism in the antirape movement. A goal of this study was to understand what "politics" means in the context of social movement participation. I began the study with a rather open-ended understanding of the term. By observing participants to understand how they engaged in politics, I began to gain an understanding of what politics meant for these groups and individuals. I learned from my observations that politics seemed to be about power: "who has it, who wants it, and how it is given, negotiated and taken away" (Blackstone, 2007). Blackstone, A. (2007). Finding politics in the silly and the sacred: Anti-rape activism on campus. *Sociological Spectrum*, 27, 151–163. Specific actions, such as the awareness-raising bicycle event Ride Against Rape, seemed to be political in that

they empowered survivors to see that they were not alone, and they empowered clinics (through funds raised at the event) to provide services to survivors. By taking the time to observe movement participants in action for many months, I was able to learn how politics operated in the day-to-day goings-on of social movements and in the lives of movement participants. While it was not evident at the outset of the study, my observations led me to define politics as linked to *action* and *challenging power*. In this case, I conducted observations before actually coming up with a clear definition for my key term, and certainly before identifying indicators for the term. The measurement process therefore worked more inductively than [Figure 6.7 "The Process of Measurement"](#) implies that it might.

KEY TAKEAWAYS

- Operationalization involves spelling out precisely how a concept will be measured.
- The measurement process generally involves going from a more general focus to a narrower one, but the process does not proceed in exactly the same way for all research projects.

EXERCISE

1. Think of a concept that is of interest to you. Now identify some possible indicators of that concept.

6.4 Measurement Quality

LEARNING OBJECTIVES

1. Define reliability.
2. Define validity.

Once we've managed to define our terms and specify the operations for measuring them, how do we know that our measures are any good? Without some assurance of the quality of our measures, we cannot be certain that our findings have any meaning or, at the least, that our findings mean what we think they mean. When social scientists measure concepts, they aim to achieve **reliability**¹¹ and **validity**¹² in their measures. These two aspects of measurement quality are the focus of this section. We'll consider reliability first and then take a look at validity. For both aspects of measurement quality, let's say our interest is in measuring the concepts of alcoholism and alcohol intake. What are some potential problems that could arise when attempting to measure this concept, and how might we work to overcome those problems?

Reliability

First, let's say we've decided to measure alcoholism by asking people to respond to the following question: Have you ever had a problem with alcohol? If we measure alcoholism in this way, it seems likely that anyone who identifies as an alcoholic would respond with a yes to the question. So this must be a good way to identify our group of interest, right? Well, maybe. Think about how you or others you know would respond to this question. Would responses differ after a wild night out from what they would have been the day before? Might a teetotaler's current headache from the single glass of wine he had last night influence how he answers the question this morning? How would that same person respond to the question before consuming the wine? In each of these cases, if the same person would respond differently to the same question at different points, it is possible that our measure of alcoholism has a reliability problem. Reliability in measurement is about consistency. If a measure is reliable, it means that if the same measure is applied consistently to the same person, the result will be the same each time.

11. Exists when the same measure, applied consistently to the same person, yields the same result each time.
12. Exists when there is a shared understanding of the meaning of whatever concept is being measured.

One common problem of reliability with social scientific measures is memory. If we ask research participants to recall some aspect of their own past behavior, we should try to make the recollection process as simple and straightforward for them

as possible. Sticking with the topic of alcohol intake, if we ask respondents how much wine, beer, and liquor they've consumed each day over the course of the past 3 months, how likely are we to get accurate responses? Unless a person keeps a journal documenting their intake, there will very likely be some inaccuracies in their responses. If, on the other hand, we ask a person how many drinks of any kind he or she has consumed in the past week, we might get a more accurate set of responses.

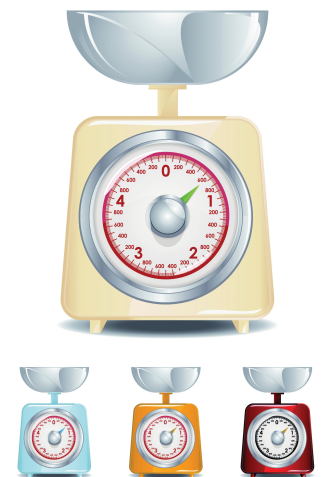
Reliability can be an issue even when we're not reliant on others to accurately report their behaviors. Perhaps a field researcher is interested in observing how alcohol intake influences interactions in public locations. She may decide to conduct observations at a local pub, noting how many drinks patrons consume and how their behavior changes as their intake changes. But what if the researcher has to use the restroom and misses the three shots of tequila that the person next to her downs during the brief period she is away? The reliability of this researcher's measure of alcohol intake, counting numbers of drinks she observes patrons consume, depends on her ability to actually observe every instance of patrons consuming drinks. If she is unlikely to be able to observe every such instance, then perhaps her mechanism for measuring this concept is not reliable.

Validity

While reliability is about consistency, validity is about shared understanding. What image comes to mind for you when you hear the word *alcoholic*? Are you certain that the image you conjure up is similar to the image others have in mind? If not, then we may be facing a problem of validity.

To be valid, we must be certain that our measures accurately get at the meaning of our concepts. Think back to the first possible measure of alcoholism we considered in the subsection "Reliability." There, we initially considered measuring alcoholism by asking research participants the following question: Have you ever had a problem with alcohol? We realized that this might not be the most reliable way of measuring alcoholism because the same person's response might vary dramatically depending on how he or she is feeling that day. Likewise, this measure of alcoholism is not particularly valid. What is "a problem" with alcohol? For some, it

Figure 6.8



Reliability is like a scale: the data you collect is only as dependable as the instrument doing the measuring.

© Thinkstock

might be having had a single regrettable or embarrassing moment that resulted from consuming too much. For others, the threshold for “problem” might be different; perhaps a person has had numerous embarrassing drunken moments but still gets out of bed for work every day so doesn’t perceive himself or herself to have a problem. Because what each respondent considers to be problematic could vary so dramatically, our measure of alcoholism isn’t likely to yield any useful or meaningful results if our aim is to objectively understand, say, how many of our research participants are alcoholics. Of course, if our interest is in how many research participants perceive themselves to have a problem, then our measure may be just fine.

Let’s consider another example. Perhaps we’re interested in learning about a person’s dedication to healthy living. Most of us would probably agree that engaging in regular exercise is a sign of healthy living, so we could measure healthy living by counting the number of times per week that a person visits his local gym. At first this might seem like a reasonable measure, but if this respondent’s gym is anything like some of the gyms I’ve seen, there exists the distinct possibility that his gym visits include activities that are decidedly *not* fitness related. Perhaps he visits the gym to use their tanning beds, not a particularly good indicator of healthy living, or to flirt with potential dates or sit in the sauna. These activities, while potentially relaxing, are probably not the best indicators of healthy living. Therefore, recording the number of times a person visits the gym may not be the most valid way to measure his or her dedication to healthy living. Using this measure wouldn’t *really* give us an indication of a person’s dedication to healthy living. So we wouldn’t really be measuring what we intended to measure.

At its core, validity is about social agreement. One quick and easy way to help ensure that your measures are valid is to discuss them with others. One way to think of validity is to think of it as you would a portrait. Some portraits of people look just like the actual person they are intended to represent. But other representations of people's images, such as caricatures and stick drawings, are not nearly as accurate. While a portrait may not be an exact representation of how a person looks, what's important is the extent to which it approximates the look of the person it is intended to represent. The same goes for validity in measures. No measure is exact, but some measures are more accurate than others.

Figure 6.9



Validity is like a portrait. No measure is exact; what's important is how closely your measure approximates your concept.

© Thinkstock

KEY TAKEAWAYS

- Reliability is a matter of consistency.
- Validity is a matter of social agreement.

EXERCISES

1. Operationalize a concept that is of interest to you. What are some possible problems of reliability or validity that you could run into given your operationalization? How could you tweak your operationalization and overcome those problems?
2. Sticking with the same concept you identified in exercise 1, find out how other sociologists have operationalized this concept. You can do this by revisiting readings from other sociology courses you've taken or by looking up a few articles using Sociological Abstracts. How does your plan for operationalization differ from that used in previous research? What potential problems of reliability or validity do you see? How do the researchers address those problems?

6.5 Complexities in Measurement

LEARNING OBJECTIVES

1. Define and provide examples for each of the four levels of measurement.
2. Define the terms *index* and *typology*, and discuss an example of each.

You should now have some idea about how conceptualization and operationalization work, and you also know a bit about how to assess the quality of your measures. But measurement is sometimes a complex process, and some concepts are more complex than others. Measuring a person's political party affiliation, for example, is less complex than measuring her or his sense of alienation. In this section we'll consider some of these complexities in measurement. First, we'll take a look at the various levels of measurement that exist, and then we'll consider a couple strategies for capturing the complexities of the concepts we wish to measure.

Levels of Measurement

When social scientists measure concepts, they sometimes use the language of variables and attributes. A **variable**¹³ refers to a grouping of several characteristics. **Attributes**¹⁴ are those characteristics. A variable's attributes determine its level of measurement. There are four possible levels of measurement; they are nominal, ordinal, interval, and ratio.

At the **nominal**¹⁵ level of measurement, variable attributes meet the criteria of exhaustiveness and mutual exclusivity. This is the most basic level of measurement. Relationship status, gender, race, political party affiliation, and religious affiliation are all examples of nominal-level variables. For example, to measure relationship status, we might ask respondents to tell us if they are currently partnered or single. These two attributes pretty much exhaust the possibilities for relationship status (i.e., everyone is always one or the other of these), and it is not possible for a person to simultaneously occupy more than one of these statuses (e.g., if you are single, you cannot also be partnered). Thus this measure of relationship status meets the criteria that nominal-level attributes must be exhaustive and mutually exclusive. One unique feature of nominal-level measures is that they cannot be mathematically quantified. We cannot say, for example, that being partnered has more or less quantifiable value than being single (note we're not talking here about

13. A grouping of several characteristics.

14. The characteristics that make up a variable.

15. Level of measurement for which variable attributes meet the criteria of exhaustiveness and mutual exclusivity.

the economic impact of one's relationship status—we're talking only about relationship status on its own, not in relation to other variables).

Unlike nominal-level measures, attributes at the **ordinal**¹⁶ level can be rank ordered, though we cannot calculate a mathematical distance between those attributes. We can simply say that one attribute of an ordinal-level variable is more or less than another attribute. Ordinal-level attributes are also exhaustive and mutually exclusive, as with nominal-level variables. Examples of ordinal-level measures include social class, degree of support for policy initiatives, television program rankings, and prejudice. Thus while we can say that one person's support for some public policy may be more or less than his neighbor's level of support, we cannot say exactly *how much* more or less.

At the **interval**¹⁷ level, measures meet all the criteria of the two preceding levels, plus the distance between attributes is known to be equal. IQ scores are interval level, as are temperatures. Interval-level variables are not particularly common in social science research, but their defining characteristic is that we can say how much more or less one attribute differs from another. We cannot, however, say with certainty what the ratio of one attribute is in comparison to another. For example, it would not make sense to say that 50 degrees is half as hot as 100 degrees.

Figure 6.10



Relationship status is an example of a nominal-level variable.

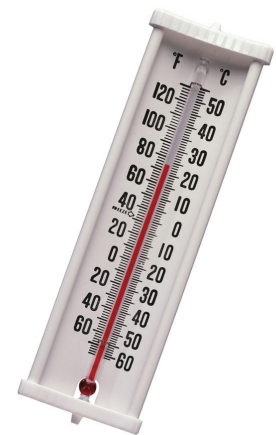
16. Level of measurement for which variable attributes meet the criteria of exhaustiveness and mutual exclusivity and can also be rank ordered.

17. Level of measurement for which variable attributes meet the criteria of exhaustiveness and mutual exclusivity and can be rank ordered, and the distance between attributes is known to be equal.

© Thinkstock

Finally, at the **ratio**¹⁸ level, attributes are mutually exclusive and exhaustive, attributes can be rank ordered, the distance between attributes is equal, and attributes have a true zero point. Thus with these variables, we *can* say what the ratio of one attribute is in comparison to another. Examples of ratio-level variables include age and years of education. We know, for example, that a person who is 12 years old is twice as old as someone who is 6 years old.

Figure 6.11



Temperature is an example of an interval-level variable.

© Thinkstock

Indexes, Scales, and Typologies

Earlier I mentioned that some concepts have dimensions. To account for a concept's dimensions a researcher might rely on indexes, scales, or typologies. An **index**¹⁹ is a type of measure that contains several indicators and is used to summarize some more general concept. The Gallup poll on well-being described earlier in this chapter uses an index to measure well-being. Rather than ask respondents how well they think they are, Gallup has designed an index that includes multiple indicators of the more general concept of well-being (<http://www.gallup.com/poll/123215/Gallup-Healthways-Index.aspx>).

Like an index, a **scale**²⁰ is also a composite measure. But unlike indexes, scales are designed in a way that accounts for the possibility that different items on an index may vary in intensity. Take the Gallup well-being poll as an example and think about Gallup's six dimensions of well-being: physical health, emotional health, work environment, life evaluation, healthy behaviors, and access to basic necessities. Is it possible that one of these dimensions is a more important contributor to overall well-being than the others? For example, it seems odd that a person who lacks access to basic necessities would rank equally in well-being to someone who has access to basic necessities but doesn't regularly engage in healthy behaviors such as exercise. If we agree that this is the case, we may opt to give "access to basic necessities" greater weight in our overall measure of well-being than we give to "healthy behaviors," and if we do so, we will have created a scale.

- 18. Level of measurement for which variable attributes meet the criteria of exhaustiveness and mutual exclusivity and can be rank ordered, the distance between attributes is known to be equal, and attributes have a true zero point.
- 19. A type of measure that contains several indicators and is used to summarize some more general concept.
- 20. A type of measure that contains several indicators that vary in intensity.

A **typology**²¹, on the other hand, is a way of categorizing concepts according to particular themes. For example, in his classic study of suicide, Emile Durkheim (1897) Durkheim, E. (1897 [2006 translation by R. Buss]). *On suicide*. London, UK: Penguin. identified four types of suicide including altruistic, egoistic, anomic, and fatalistic. Each of these types is linked to the concept of suicide, but the typology allows us to classify suicide in ways that make the concept more meaningful and that help simplify the complexities of the concept.

Let's consider another example. Sexual harassment is a concept for which there exist indexes, scales, and typologies. One typology of harassment, used in the US legal system, includes two forms of harassment: quid pro quo and hostile work environment (Blackstone & McLaughlin, 2009). Blackstone, A., & McLaughlin, H. (2009). Sexual harassment. In J. O'Brien & E. L. Shapiro (Eds.), *Encyclopedia of gender and society* (pp. 762–766). Thousand Oaks, CA: Sage. Quid pro quo harassment refers to the sort where sexual demands are made, or threatened to become, a condition of or basis for employment. Hostile work environment harassment, on the other hand, refers to sexual conduct or materials in the workplace that unreasonably interfere with a person's ability to perform her or his job. While both types are sexual harassment, the typology helps us better understand the forms that sexual harassment can take and, in turn, helps us as researchers better identify what it is that we are observing and measuring when we study workplace harassment.

Sexual harassment is a concept for which there are also indexes. A sexual harassment index would use multiple items to measure the singular concept of sexual harassment. For example, you might ask research participants if they have ever experienced any of the following in the workplace: offensive sexual joking, exposure to offensive materials, unwanted touching, sexual threats, or sexual assault. These five indicators all have something to do with workplace sexual harassment. On their own, some of the more benign indicators, such as joking, might not be considered harassment (unless severe or pervasive), but collectively, the experience of these behaviors might add up to an overall experience of sexual harassment. The index allows the researcher in this case to better understand what shape a respondent's harassment experience takes. If the researcher had only asked whether a respondent had ever experienced sexual harassment at work, she wouldn't know what sorts of behaviors actually made up that respondent's experience. Further, if the researcher decides to rank order the various behaviors that make up sexual harassment, perhaps weighting sexual assault more heavily than joking, then she will have created a scale rather than an index.

Let's take a look at one more specific example of an index. In a recent study that I conducted of older workers, I wanted to understand how a worker's sense of financial security might shape whether they leave or stay in positions where they feel underappreciated or harassed. Rather than ask a single question, I created an

21. A way of categorizing concepts according to particular themes.

index to measure financial security. That index can be found in [Figure 6.12](#) "[Example of an Index Measuring Financial Security](#)". On their own, none of the questions in the index is likely to provide as accurate a representation of financial security as the collection of all the questions together.

Figure 6.12 Example of an Index Measuring Financial Security

FINANCES

In general, how financially secure would you say you are?
(Circle one.)

1 2 3 4 5
Not at all secure Moderately Very secure

Since age 62, have you ever received money from family members or friends to help make ends meet?

No Yes

If yes, how many times? 1 or 2 times 3 or 4 times 5 times or more

Since age 62, have you ever received money from a church or other organization to help make ends meet?

No Yes

If yes, how many times? 1 or 2 times 3 or 4 times 5 times or more

Since age 62, have you ever donated money to a church or other organization?

No Yes

Since age 62, have you ever given money to a family member or friend to help them make ends meet?

No Yes

In sum, indexes and typologies are tools that researchers use to condense large amounts of information, to simplify complex concepts, and most generally, to make sense of the concepts that they study.

KEY TAKEAWAYS

- In social science, our variables can be one of four different levels of measurement: nominal, ordinal, interval, or ratio.
- Indexes and typologies allow us to account for and simplify some of the complexities in our measures.

EXERCISES

1. Together with a fellow research methods student, identify six concepts that are of interest to you both. Next, on your own, identify each concept's level of measurement. Share your answers with your peer. Discuss why you chose each level of measurement that you chose and, together, try to come to some agreement about any concepts that you labeled differently.
2. Take a look at Gallup's page on their well-being index:
<http://www.gallup.com/poll/123215/Gallup-Healthways-Index.aspx>.
Read about how various concepts there are operationalized and indexed.