2

LEVELS OF MEASUREMENT AND THEIR IMPORTANCE

One Potato, Two Potato

Difficulty Index ☺ ☺ ☺ (a bit harder than Chapter 1 but easily understood)

LEARNING OBJECTIVES

After reading this chapter, you should be able to

- Define the basic testing terms of *variable* and *measurement*.
- Explain the four levels of measurement.
- Compare and contrast the four levels of measurement.
- Give examples of how researchers choose to measure variables using different levels of measurement.

How things are measured is very important to our study of tests and measurement. And since *measuring things* basically means slapping a score on them, the nuts and bolts of the numbers that are used to score tests and how they come to be is especially important.

What test developers and social science researchers call *levels of measurement* refers to the way that numbers are assigned to represent variables. There are low levels of measurement, which provide very little information about how people differ from each other on some variable, and there are high levels of measurement, which tell a lot about how people differ from each other on some variable.

A single variable can be measured using four different levels of measurement and the higher the level of measurement, the more powerful the types of analysis one can use. For example, in one study, researchers from the University of Kansas examined how the very interesting concepts of self-determination and self-concept have an effect on academic achievement for adolescents with learning disabilities. Using trusted and field-tested assessment tools, they found significant relationships among the three variables of self-determination, self-concept, and academic achievement, with self-determination being a useful predictor of academic achievement for these students. Because the variables were quantified using a high level of measurement, cool statistics like correlation coefficients and a sophisticated approach called multiple regression could be used. Otherwise, these relationships might have been completely missed.

Want to know more? Go to Google Scholar and look up Zheng, C., Erickson, A., Kingston, N., & Noonan, P. (2014). The relationship among self-determination, self-concept, and academic achievement for students with learning disabilities. *Journal of Learning Disabilities*, *47*, 462–474.

FIRST THINGS FIRST

Before we start talking about levels of measurement, let's spend a moment defining a few important terms—specifically, what a *variable* is and what the term *measurement* means.

A **variable** is anything (such as a test score) that can take on more than one value. For example, a score on the SAT can take on more than one value (such as 750 or 420), as can what category of color your hair falls into (such as red or black). Age is another variable (someone can be 2 months old or 87 years old), as is favorite flavor of ice cream (such as rocky road or mint chocolate chip). Notice that the labels we apply to outcomes can be quantitative (such as 87 years) or qualitative (such as rocky road or mint chocolate chip). Variables can be represented by quantities or by different categories (which is what *qualitative* means). Good, that's out of the way.

Now, the term **measurement** means the assignment of labels to (you guessed it) a variable or an outcome. So when we apply the label "blue" to a particular outcome, we are measuring that outcome. We can measure the number of windows in a house, the color of a car, the score on a test of memory, and how fast someone can run 100 yards. In every case, we are assigning a label to an outcome. Sometimes

that label is precise (such as 10.7 seconds for the time it takes to run 100 yards), and sometimes it is less precise (such as "like" for how someone feels about a presidential candidate).

There is a great deal of discussion (and of course controversy) over what some levels mean. For example, the biological sex of a child at birth is designated (measured) as male or female. But the *gender* label of that individual may differ from the biological designation, because a gender label is often socially or psychologically determined (male and female might still be used as categories but not be based on body parts or perhaps gender is on a continuum somewhere between male and female). These are interesting and provocative options for measurement folks that represent different theoretical constructs, yes, but also different levels of measurement.

As the world turns, about 75 years ago, in 1946, S. S. Stevens, a famous experimental psychologist (not a steamship), started to wonder about how different types of variables were measured and whether the level of precision at which those variables were measured made a difference. He wanted to know whether a set of rules could be developed to categorize variables based on the how their scores were assigned.

And coming in at about 10 pounds 11 ounces—the idea of levels of measurement was born. Or to be more precise, 171 ounces. Or to be less precise, about 11 pounds.

Variables can be measured in different ways, and the way a variable is being measured determines the *level of measurement* being used. For example, we can measure the height of an individual in several different ways. If we say only that Bruce is taller than Neil, then we have chosen to measure this variable at a level that distinguishes one person from another only in a rank order sort of way—one person is less tall than another, but we don't know by how much. On the other hand, if we chose to measure this variable in a way that differentiates people by a certain quantity of inches, that is much more precise and much more informative. It's possible that Bruce is only barely taller than Neil (maybe just an inch), not a lot taller. Those are different interpretations of the same variable because two different levels of measurement were used.

THE FOUR HORSEMEN (OR LEVELS) OF MEASUREMENT

A **level of measurement** represents how much information is being provided by the outcome measure. There are four levels of measurement—nominal, ordinal, interval, and ratio—and here's more about each.

The Nominal Level of Measurement

The **nominal level of measurement** describes a measurement system where numbers are used to identify different categories but not used as quantities. These

are variables that are *categorical* or *discrete* in nature. These "labels" (the scores) are qualitative in nature, and people (or objects or whatever) can be placed in one and only one category, which is why such scores are mutually exclusive. You can't be in one reading group (the Guppies) and another reading group (the Sharks) at the same time.

For example, Max is in the Guppies and Aya is in the Sharks. Imagine we use 1 to mean Guppies and 2 to mean Sharks. Knowing that Max's score is a 1 and Aya's score is a 2 tells us only that Max and Aya are in two different reading groups—not how the groups differ, or who is better, or how well Max or Aya can read—just that they are in two different groups. It counts as measurement because it differentiates people on a variable, but, gosh, it sure doesn't tell us much by itself. It is barely better than no information at all.

We call this level the nominal level of measurement (after the word *nomin* in Latin, meaning name), because the only distinction we can make is that variables differ in the category in which they are placed. Measuring a variable such as reading ability by group assignment (and nothing more) is similar to measuring political attitudes using the labels Republicans and Democrats; measuring income by seeing if people are Honda Civic or Tesla drivers; and measuring preferences for fresh fruit by asking if people shop at Kroger's or the Piggly-Wiggly. They are "measured" only by the nature of the group to which they belong. (When writing this paragraph, Bruce paused to celebrate the first time in his life he had ever typed the phrase "piggly-wiggly." Huzzah.)

Want to know more about how these variables differ? Well, you can't just by knowing in which category a person belongs, and that's the big significant limitation with the nominal level of measurement. If you want more information, then you have to dig more or define (and measure) the variable in a more precise way, which we will do shortly.

Take a second to really think about the limitation of using nominal levels of measurement in social science research. It's a major weakness with this approach, right? But (and you might want to sit down for this) think about how often in science we compare two groups and look for differences! Like comparing an experimental group that received a new drug with a control group that did not on some health outcome (like blood pressure). That variable from which we create the two groups is measured at the nominal level! We might be missing a lot of potential discoveries if we measured all of our research variables at higher levels of measurement than the very imprecise nominal level. (In this example, imagine if everyone in the study got some slightly different amount of the new drug. Not all or nothing. There's a lot more information in that variable.)

Of course, sometimes variables should only be measured at the nominal level because that makes the most sense. An example of a study that used a nominal-level

variable is one conducted by Rik Verhaeghe and his colleagues, detailed in a 2003 article that appeared in *Stress and Health: Journal of the International Society for the Investigation of Stress* (you know, that magazine we all subscribe to). They examined job stress among middle-aged health care workers and its relation to absences due to sickness. One assessment that had to be made was based on the assignment of people to one of two groups—nurses or non-nurses—and that variable was measured at the nominal level. As you can see, a participant can be in only one group at a time (they are mutually exclusive), and for this variable (occupation) you can assign only a label (nurse or non-nurse). People don't vary in the amount of nurse they are.

Want to know more? Take a look at Verhaeghe, R., Mak, R., VanMaele, G., Kornitzer, M., & De Backer, G. (2003). Job stress among middle-aged health care workers and its relation to sickness absence. *Stress and Health: Journal of the International Society for the Investigation of Stress*, *19*(5), 265–274.

There's always a ton of discussion about measurement levels and their utility, starting with the definition of variables and how those variables are measured. And in some cases, there's doubt that the nominal level of measurement is a level of measurement at all rather than only a qualitative description of some outcome. So, if it doesn't make sense to you to treat categorization as measurement just because numbers are used as the names of categories (e.g., 1 = nurse, 0 = not a nurse), that's okay. But here are two things to remember about nominal-level measurement.

1. The categories in which measures can be placed on the nominal scale are always mutually exclusive. You can't be in the red preschool room and the blue preschool room at the same time.

2. Nominal-level measures are always qualitative (the values are categories and not quantities). Being in the red room is neither here nor there, as is being in the blue room. You're a preschooler in either room, and that's it; the room assignment says nothing about anything other than that—just which room you are in. If, in your data, a 2 means blue and a 1 means red, blue isn't twice as much as red just because a 2 is twice as much as 1 mathematically. (Note to editor: Please check our math here.) The numbers are only names.

The Ordinal Level of Measurement

Next, we have the **ordinal level of measurement**, which describes how variables can be ordered along some type of continuum. (Get it? Ordinal, as in ordering a set of things.) So outcomes are placed in categories (like the nominal scale), but they also have an order or rank to them as well, like stronger and weaker, taller and shorter, faster and slower, and so on.

For example, let's take Max and Aya again. As it turns out, Max is a better reader than Aya. Right there is the one and only necessary criterion for a measure to be

at the ordinal level of measurement. It's the "better than" or "worse than" thing—some expression of the relationship between categories.

However, from better or worse, we cannot tell anything about how good a reader either Max or Aya is, because ordinal levels of measurement do not include this information. If we are measuring reading ability, and scores are assigned based on rank ordering all students in a classroom from best reader to worst reader (which is an ordinal-level approach), we can know if Max scores higher than Aya or lower than some other student, but we don't know how good a reader he is or Aya is or any student is, just how they *compare* to each other. Max, Aya, and, heck, the whole class might be great readers (or might be not-so-great), but we don't know by looking at scores using the ordinal level of measurement.

Our real-world example of ordinal-level measurement is a study by Kathe Burkhardt and her colleagues that appeared in the journal *Behavior Change*. They examined common childhood fears in 9- to 13-year-old South African children, and one of the ways they assessed fears was by having children rank them. In fact, the researchers found that the children's rankings of fears differed from rankings derived using a scale that attached an actual value to the fear.

Want to know more? Take a look at Burkhardt, K., Loxton, H., & Muris, P. (2003). Fears and fearfulness in South African children. *Behaviour Change*, *20*(2), 94–102.

The Interval Level of Measurement

That's two levels of measurement down and two to go.

The **interval level of measurement** gives us a nice jump in the amount of information we obtain from a new level of measurement. (And, SPOILER ALERT, in real-world research the interval level of measurement is the most popular approach.) You already know that we can assign names (nominal level) and rank (ordinal level), but it is with the interval level of measurement that we can assign a value to an outcome that is based on some underlying continuum that has equal "intervals." And if there is an underlying continuum, then we can make very definite statements about someone's position (their score) along that continuum and their position relative to another person's position, including statements about such things as differences. Wow, that's a lot more complex than the earlier two levels of measurement and provides a lot more information as well. Scores at this level tell us that each score is different from others in some way (like nominal level does), the scores represent more and less of some quantity (like ordinal level does), but now the quantities are precise enough that the distance between any two adjacent scores on the scale are equal. For instance, on a Celsius thermometer, the difference between 48 degrees and 47 degrees is one degree of heat (whatever that means) and the difference between 34 degrees and 33 degrees is also one degree of heat. Everywhere along the scale, there is an equal amount of difference in the variable (heat) between any two scores that are side by side. In other words, a mathematical

difference between two scores is the same as the quantitative difference between units of the variable.

To understand more about interval-level measurement, we look at the architectural origins of the word "interval." Interval, or *inter val*, means between walls. It describes the top of towers and such on old castles. Notice those little stone protective barriers at the top of the towers in Figure 2.1? A well-designed fortress like a castle needed protection for the guards at the gate shooting arrows (or throwing rocks or whatever; what are we, medievalists?), but there also had to be openings to shoot those arrows at the attackers. And the best design had equal spacing between those barriers all the way around because you never know where they will be coming from. So those intervals have equal spacing, like an interval-level scale!

Using our Max and Ava reading-ability example, not only do we know that Max and Ava fall into two different categories of readers and that Max is a better reader than Ava, but we can also now know *how much* better Max actually is. Let's assume that reading ability was measured using a reading comprehension test. Imagine Max got 82 points out of 100 possible on the test and Ava got 42 points. Because one of the assumptions of this level of measurement is that it is based on a scale that has equally appearing intervals, we can say not only that Max got 40 points more than Ava (which is true mathematically) but also that Max has 40 points more of reading ability. Now, what that means in a theoretical sense is determined by the test developer or the researcher.

Although an interval-level scale provides much more information than an ordinal- or nominal-level scale, you have to be careful in how you interpret these scores. For example, scoring 50% higher on a history test does not mean that score represents 50% more knowledge (unless the test is a perfect, perfect, perfect representative of all the questions that could be asked). Rather, it means only that 50% more of the questions were answered correctly. We can conclude that the more questions correct, the better one is in history, but don't carry it too far and overgeneralize from a test score to an entire area of knowledge or a construct such as intelligence.

FIGURE 2.1 ●   "Intervals" on a Castle Gate
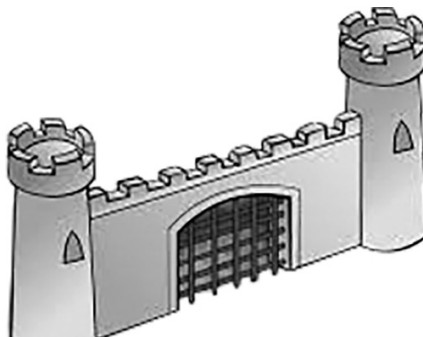
What's the big advantage of the interval level of measurement over the nominal and ordinal besides increased precision? In one word, *information*—there's much more of it when we know what a score actually is and what it means. Remember, Max could be ranked number 1 in his class but get only 50% of the questions correct on the reading comprehension test. On the other hand, knowing what his exact score is relative to some type of underlying continuum provides us with an abundance of information when it comes time to make a judgment about his performance. And, as we will see later in Chapter 5, once you are at the interval level, you can start to use all sorts of fancy statistics to understand the scores, like means, percentile ranks, and such.

Shintaro Saro at Montclair State University and colleagues were interested in tourists who visit sports venues and the relationship between the perceived value of the experience and their loyalty. Measuring variables at the interval level allowed them to use advanced statistical analyses involving correlations. (Unless you are at the interval level of measurement, correlational analyses are much less powerful.) The researchers found that for experienced travelers, the emotional response related to the experience was the best predictor of planning to return again, but for novice travelers, it was perceptions of quality that was most important.

Want to know more? Sato, S., Gipson, C., Todd, S., & Harada, M. (2018). The relationship between sport tourists' perceived value and destination loyalty: An experience-use history segmentation approach. *Journal of Sport & Tourism*, *22*(2), 173–186.

The Ratio Level of Measurement

This is by far the most informative level of measurement, yet the one that is least likely to be seen in the social or behavioral sciences. Why? Because the **ratio level of measurement** is characteristic of all the other scales we have already talked about, but it also includes a very important assumption—an *absolute zero* corresponding to an absence of the trait or characteristic being measured. Physical measurements such as amount of rainfall, weight, and height fall under the ratio level of measurement. Or any time you count things—like the number of dogs on a porch.

The scale and its use become interesting when we begin to look at nonphysical attributes or behaviors. For example, it is possible to receive a score of zero on a reading comprehension test, right? But here's the big question: Does getting such a score indicate that one has no reading ability? They cannot read at all, even with a little accuracy? Of course not. It means only that on this test, they scored really, really low. (Maybe you are the sort of student who is about to raise your hand and say something like, "But what about a one-day-old baby?! They have zero reading ability!!" Yes, but the test isn't designed for one-day-old babies. Nice try.)

That's the challenge, then. Is there any trait or characteristic in the behavioral or social sciences that an individual can have a complete absence of? If there is not,

then a ratio level of measurement is impossible. In fact, this is one reason why, when this category of tests and measurement is taught, the interval and ratio levels often are combined into one. We're not doing that here, because we think they are important enough to keep separated.

Now, in the physical and biological sciences, it's not as much of an issue or challenge. Consider temperature as measured on the Kelvin scale. This scale starts at 0, which means no heat at all (somewhere in deep space maybe) and goes up, up, up. The sun is almost 6,000 degrees on the Kelvin scale. There are no negative values among those possible scores. Compare that to Celsius or Fahrenheit scoring systems for temperature. It can be negative degrees on those scales, like –20 degrees Fahrenheit at the North Pole or somewhere else. (Notice with the variable of heat, one can *choose* to measure it at the interval level, as with the Fahrenheit system, or choose to measure it at the ratio level. Researchers and test developers have to decide which makes the most sense theoretically.)

How about finger tapping as a measure of responsiveness? It is entirely possible to have no finger taps. Both the number of finger taps and the Kelvin scale truly have a true zero. But even if someone doesn't get anything correct on an intelligence test (perhaps one taken in a language from Mars), does that mean they have no intelligence? Of course not.

By the way, we call it *ratio* level, because proportions and fractions have meaning. You can tap your fingers twice as much as your classmate when you are nervous, so comparing your finger tapping score of 120 (taps per minute) to their score of 60 by saying you tap twice as often is reasonable and makes sense. But if it is 30 degrees Fahrenheit today and it was 15 degrees yesterday, we don't say it is twice as hot today. Because scores on the Fahrenheit scale go below zero, we aren't allowed mathematically to make ratio comparisons like that. But on the Kelvin scale we can!

For social and behavioral scientists (like us, and like you, probably), we will rarely (if ever) see a ratio-level scale in the journal articles we review and read. The scale of measurement simply depends on how the variable is being defined and measured.

A SUMMARY: HOW LEVELS OF MEASUREMENT DIFFER

We just discussed four different levels of measurement and what some of their characteristics are. You also know by now that a more precise level of measurement has all the characteristics of an earlier level and provides more information as well.

In Table 2.1, you can see a summary that addresses the following questions:

1. Are you measuring most of the available information?

2. Can you assign a name to the variable being measured?

TABLE 2.1 ⬡ A Summary of Levels of Measurement and the Characteristics That Define Them					
Level of Measurement	How Much Information About the Variable Do You Get?	Can You Assign Names to the Different Scores or Outcomes?	Can You Assign an Order to the Different Scores or Outcomes?	Can You Assign an Underlying Quantitative Scale to the Different Scores or Outcomes?	Can You Assign an Absolute Zero to the Scale You Are Using?
Ratio	A great amount	☺	☺	☺	☺
Interval	A good amount	☺	☺	☺	☹
Ordinal	Some	☺	☺	☹	☹
Nominal	A little	☺	☹	☹	☹

3. Can you assign an order to the variable being measured?

4. Can you assign an underlying quantitative scale to the variable being measured?

Remember, the more precise your level of measurement, the more information is conveyed.

This table shows us that the ratio level of measurement allows us to answer yes (☺) to these four questions, whereas the nominal level of measurement allows us to answer yes to only one.

OKAY, SO WHAT'S THE LESSON HERE?

The lesson here is that, when you can, try to select a technique for measuring a variable that allows you to use the highest level of measurement possible (most often the interval level). We want to access the most information available with the most precision while understanding that as the scale of measurement changes in precision, the way the variables are being measured will probably change in level of complexity as well, as you can see in Figure 2.2. As variables are measured in more sophisticated ways and become more complex in their definition and nature, they lend themselves better to higher **scales of measurement** (such as interval or ratio) than do variables that are less complex.

For example, when testing the effectiveness of strength training in senior citizens, don't classify them as weak or strong after the intervention is over and after they have been tested. Rather, try to get a ranking of how strong they are, and even better, try to get an actual number associated with strength, like the amount of weight

FIGURE 2.2 ⬡ Complexity and Precision and Scales of Measurement
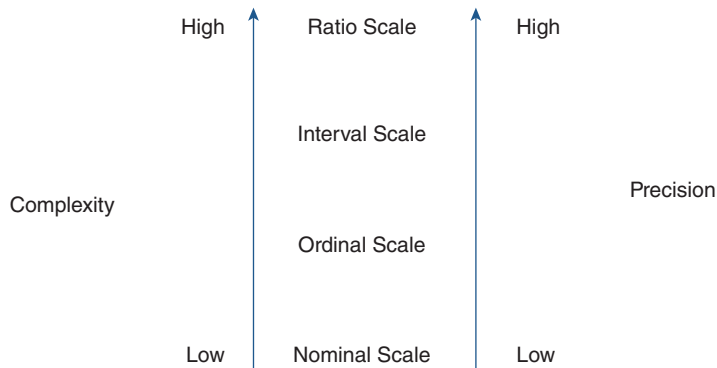
they can lift. That provides much more information and makes your entire quest for knowledge a more powerful one.

But the real world sometimes demands that certain outcomes be measured in certain ways, and that limits us as to the amount of information available. For example, what if you wanted to study prejudice? You may not be able to ascertain anything more than placing participants into ordinal levels called *very prejudiced*, *somewhat prejudiced*, and *not at all prejudiced*. Not as much information as we might like but not bad either. It is what it is.

THE FINAL WORD(S) ON LEVELS OF MEASUREMENT

Okay, so we have the four levels of measurement (three of which are very commonly used)—what can we say about all of them? Here are at least five things:

1. What we measure—be it a score on a test of intelligence, the number correct on a chemistry final exam, or your feelings about peanut M&Ms—belongs to one of these four levels of measurement. The key, of course, is how finely and precisely the variable is being measured.

2. The qualities of one level of measurement (such as nominal) are characteristic of the next level up as well. In other words, variables measured at the ordinal level also contain the qualities of variables measured at the nominal level. Likewise, variables measured at the interval level contain the qualities of variables measured at both the nominal and ordinal levels. For example, if you know that Mateo swims faster than Laurie, that's great information (and it's ordinal in nature). But if you

know that Mateo swims 7.6 seconds faster than Laurie, that's interval-level information, which is even better.

3. The more precise (and higher) the level of measurement (with ratio being highest), the closer you'll get to measuring the true outcome of interest.

4. How you choose to measure an outcome defines the outcome's level of measurement. That decision might be based on theory, the realities of how the variable manifests itself in the world, or just based on a choice you make as the researcher or test developer.

5. Many researchers take some liberty in treating ordinal variables (such as scores on a personality test) as interval-level variables, and that might be fine, especially if the scores distribute themselves in ways similar to how interval-level scores do. See Chapter 5 for a discussion of the statistical issues involved.

To make matters even more complicated, even if scales (such as those of intelligence) are interval-level measures, does one assume that the five-point difference between a score of 100 and a score of 105 is equivalent to a five-point difference between a score of 125 and 130? An interval-level scale would lead us to believe that, but nope—that's not always the case. Moving from a score of 100 to 105 (around average) is not anywhere near the change that is represented by going from a score of 125 to 130 (very high scores).

These four levels of measurement are not carved in stone. We might contend that most measures in the social and behavioral sciences fall at the ordinal or interval level; in practice, however, we surely *act* as though many (if not all) occur at the interval level when they actually probably occur at the ordinal level. There are moderately complex statistical issues when treating ordinal variables as interval level (which are beyond the scope of this book), but it turns out that in most research situations it is probably okay, especially when a total score from a test is used. (Our statistician colleagues cringe a little when we say this sort of thing is "okay," but just a little.) No matter what position one takes, though, all agree that one should try to measure at the highest level possible.

Summary

We just did oodles of good work. (By the way, we measure oodles on an ordinal scale.) Now that we understand what levels of measurement are and how they work, we will turn our attention to the first of two very important topics in the study of tests and measurement—reliability. And that discussion comes in the next chapter.

Time to Practice

1. Why are levels of measurement useful?

2. Provide an example of how a variable can be measured at more than one level of measurement.

3. For the following variables, define at what level of measurement each one occurs and *tell why*.

    a. Hair color—for example, red, brown, or blonde.

    b. IQ score—for example, 110 or 143.

    c. Average number of Volvos owned by each family in Kansas City, Missouri—for example, two or three.

    d. The number correct on a third-grade math test out of 20 possible correct—for example, 17, 19, or 20.

    e. Time running 100 yards—for example, 15 seconds or 12.5 seconds.

4. Access the library database and select three journal articles that include empirical studies. Be sure that you select these articles from your own discipline. Now, for each one, answer the following questions:

    a. What is the primary variable of interest, or what variable is being studied?

    b. How is it being measured?

    c. At what level of measurement is it being measured?

    d. How precise do you think the measurement is?

5. Why does the interval level of measurement provide more information than the nominal level of measurement, and why would you want to use the interval level of measurement if you have a choice.

6. Select five variables that are important to measure in your area of expertise, and identify how they can be measured at the nominal, ordinal, or interval level of measurement.

7. Describe how you could measure temperature with each of the four levels of measurement.

8. Here's a mind bender: Which level of measurement are we using when we say that the nominal level of measurement is the least useful, followed by the ordinal level, the interval level, and then the ratio level as the most useful level of measurement?

9. Come up with a research question of interest to you. What is the highest level of measurement you could use to measure each variable, and how would you do so?

Want to Know More?

Further Readings

*   Laird, S. P., Snyder, C. R., Rapoff, M. A., & Green, S. (2004). Measuring private prayer: Development, validation, and clinical application of the Multidimensional Prayer Inventory. *International Journal for the Psychology of Religion*, *14*(4), 251–272.

How would you measure prayer?? Quantitative and qualitative aspects of prayer were assessed, and results revealed five distinct types of prayer. This study was a very interesting use of measurement.

- McHugh, M. L., & Hudson-Barr, D. (Eds.). (2003). Descriptive statistics, Part II: Most commonly used descriptive statistics. *Journal for Specialists in Pediatric Nursing*, *8*, 111–116.

These authors support what you have already read here—that data from different levels of measurement require different statistical measures. And to use descriptive statistics in the best manner, it is important to know what measurement levels should be used with the statistics and what information the statistics can provide.

And on Some Interesting Websites

- Learn all about counting systems (also a kind of measurement level) at http://galileoandeinstein .physics.virginia.edu/lectures/babylon.html.

- And of course, YouTube to the rescue. The short video from Rice University at https:// www.youtube.com/watch?v=B0ABvLa_u88 provides a nice overview of the four levels of measurement and their usefulness.

- Not the entire article (you can easily get that online through your own institution), but very cool nonetheless—read the first page of S. S. Stevens's cornerstone article on scales of measurement at http://www.jstor.org/stable/1671815?origin=JSTOR-pdf&seq=1#page_scan_tab_contents, and read more to be extra informed on this very critical topic.

And in the Real Testing World

Real World 1

Professor Granberg-Rademacker from the University of Minnesota in Mankato designed a technique that helps make the ordinal level of measurement behave more like an interval- or ratio-level measure, increasing the scale's precision and possibly its accuracy.

Want to know more? Granberg-Rademacker, J. S. (2010). An algorithm for converting ordinal scale measurement data to interval/ratio scale. *Educational and Psychological Measurement*, *70*, 74–90.

Real World 2

This study of news and how it is covered and reported examined the coverage of the pre–Iraq War debate in the *New York Times* and *Washington Post*. Professor Groshek cleverly looked at how coverage declined across all source types and levels of measurement after the congressional resolution to go to war.

Want to know more? Groshek, J. (2008). Coverage of the pre–Iraq War debate as a case study of frame indexing. *Media, War & Conflict*, *1*(3), 315–338.

Real World 3

Everywhere you look today, there's something about health and activity levels and even more about collecting data on our own activities. These researchers examined the relationship between self-reported physical activity and pedometer steps with health-related quality of life in 1,296 older

adults. They found that members of the high-step group had significantly higher scores on mental health, physical health, and global health than did the low-step group. Also, older adults had significantly higher health-related quality-of-life indices. What is so interesting about this study for our purposes is that much of the information relied on the reporting of the participants as far as steps walked (they reported their steps each day).

Want to know more? Vallance, J., Eurich, D., Gardiner, P., Taylor, L., & Johnson, S. (2016). Associations of daily pedometer steps and self-reported physical activity with health-related quality of life. *Journal of Aging and Health*, *28*, 661–674.