CHAPTER 1

Numeracy

Critical Thinking, Step One: What do the numbers and statistics tell you? When hearing or reading about research results, often the first things you notice are the data. Perhaps the initial step in critically thinking about any claims made is to understand the data and how accurate the numbers, averages, and percentages are. Learning to interpret data, read statistics, and make estimates of what may be reasonable about a set of numerical findings is a good place to begin. Understanding, representing, using, and analyzing numbers are important tools of critical thinking and numeracy.

When asked, most people claim they are better drivers or fairer or smarter than the average person. It is not atypical for people to overestimate their abilities: As Zuckerman and Jost state (2001: 209): "Desires for self-enhancement and positive self-presentation lead people to make self-serving comparisons between themselves and others. Thus most people believe that they are 'better than average' on a wide variety of traits, abilities, and outcomes." How is it possible for most of us all to be better than average? What makes this statement suspicious depends on knowing something about percentages, averages, and their mathematical properties. More important, it requires overcoming a block some people have when dealing with numbers—what has been termed "innumeracy" (Paulos, 1989).

Learning to interpret data and make estimates of what may be reasonable about a set of quantitative findings is essential to thinking critically. Just knowing how to perform basic mathematical functions is not the important thing; it's also crucial to develop a sense of understanding, representing, using, and analyzing numbers. Deciphering and interpreting numbers with confidence are central elements of "numeracy," "quantitative reasoning," or "mathematical literacy." *Numeracy* is a building block of critical thinking. The California Critical Thinking Skills Test-Numeracy (2016) defines it as

> the ability to solve quantitative reasoning problems and to make well-reasoned judgments derived from quantitative information in a variety of contexts. More than being able to compute or calculate a solution to a mathematical equation, numeracy includes understanding how quantitative information is gathered, represented, and correctly interpreted using graphs, charts, tables and diagrams.

According to the Association of American Colleges & Universities (Rhodes, 2010b), *quantitative literacy* involves:

- *interpretation* (ability to explain information presented in mathematical forms—e.g., equations, graphs, diagrams, tables, words),
- *representation* (ability to convert relevant information into various mathematical forms—e.g., equations, graphs, diagrams, tables, words),
- *calculation (*successful and sufficiently comprehensive to solve the problem clearly),
- *application/analysis* (ability to make judgments and draw appropriate conclusions based on the quantitative analysis of data while recognizing the limits of this analysis),
- *assumptions* (ability to make and evaluate important assumptions in estimation, modeling, and data analysis), and
- *communication* (expressing quantitative evidence in support of the argument or purpose of the work).

An element of quantitative reasoning is developing confidence with statistics. The American Statistical Association (GAISE, 2015: 11–12) wants students to believe and understand why:

- data beat anecdotes;
- variability is natural, predictable, and quantifiable;
- random sampling allows results of surveys and experiments to be extended to the population from which the sample was taken;
- random assignment in comparative experiments allows cause-and-effect conclusions to be drawn;
- association is not causation;
- statistical significance does not necessarily imply practical importance, especially for studies with large sample sizes;

- finding no statistically significant difference or relationship does not necessarily mean there is no difference or relationship in the population, especially for studies with small sample sizes.

In addition (GAISE, 2015: 12), students should recognize:

- common sources of bias in surveys and experiments;
- how to determine the population to which the results of statistical inference can be extended, if any, based on how the data were collected;
- how to determine when a cause-and-effect inference can be drawn from an association based on how the data were collected (e.g., the design of the study);
- that words such as "normal," "random," and "correlation" have specific meanings in statistics that may differ from common usage.

These quantitative-reasoning and numeracy goals and skills are addressed in this chapter and subsequent ones, with the emphasis less on calculation and more on learning the basic tools required to interpret, represent, critically analyze, and communicate quantitatively.

Let's consider below several basic concepts relevant to developing a critical ability to evaluate common everyday reports that use or misuse numbers: percentages, margin of error, levels of measurement, central tendency measures (averages), and estimations.

PERCENTAGES

In the summer of 2015, the California Highway Patrol (CHP) "released a new study showing a 39 percent increase in the percentage of California drivers seen using a cell phone while driving." That figure is based on 9.2 percent of drivers on a cell phone in 2015 compared with 6.6 percent in 2014. CHP also said "law enforcement wrote 35 percent more tickets for texting-while-driving compared to 2014" (California Highway Patrol, 2015).

The short press release received much attention in the media. Details about the study, however, were not in the brief announcement, thus illustrating a problematic way of reporting the results of research. To find out the relevant details left out of the announcement or news broadcast, you had to make an effort to go online to read the complete report. *What more information would you want in order to make sense out of this too-brief media report of survey findings? How do you decide whether these 39 percent, 9.2 percent, and 6.6 percent figures are meaningful or not?*

In the CHP study (Ewald and Waserman Research Consultants, 2015), we are told that the researchers saw 9.2 percent of drivers on a cell phone in 2015. The first question to ask is, How many drivers were studied? We need to know the denominator. Remember, a percentage is a proportion calculated by dividing a subset of items (numerator) by the total number of items (denominator) and then multiplied by 100 (*per centum* is Latin meaning "per one hundred"). Sometimes you will read findings "per capita" (Latin for "per head": that is, per person) and "per 1,000" as the World Bank (2014) does in reporting the number of physicians for every 1,000 people in countries around the world (such as Australia, 3.3; Canada, 2.1; Kenya, 0.2). Sometimes you may even see "per 10,000" which the World Health Organization (2013: 122, 125) uses in reporting the number of hospital beds globally (like Ethiopia, 63; New Zealand, 23). These kinds of standardization (percent, per 1,000, etc.) allow us to compare findings among samples of different sizes.

Consider how important this ability to standardize is when deciding whether to report data in terms of absolute numbers as opposed to percentages. For example, it would be accurate to report that Californians hold the most passports among all Americans, averaging close to 2.5 million per year. As the most populous state, with nearly 40 million residents, this shouldn't be a surprising finding. But using raw numbers when comparing states of different populations distorts the message. It turns out that California is actually behind six other states, including less populous states like Alaska and Delaware, when taking into account the relative population of the states: that is, when calculating the percentage of citizens with passports (Stabile, 2016).

> CRITICAL THINKING TIP
>
> When comparing numbers taken from samples or populations of different sizes, be sure to report both the absolute numbers and the percentages (per 100 or per 1,000 or per capita) so that comparisons can be made. Be critical when only raw numbers are presented, without the total number in each of the comparison groups being provided.

Understanding how to read or calculate an accurate percentage is one step in learning numeracy and thinking critically. It's also important to understand what is meant when reports discuss a percentage increase or decrease over time. Recall the CHP cell-phone-use-while-driving study: 9.2 percent of drivers on a cell phone is 39 percent higher than the previous year's 6.6 percent. *How can 9.2 be 39 percent higher than 6.6?*

At the simplest level, a percentage point increase or decrease is a difference between two percentages, so you could say that cell phone usage went up 2.6 percent in one year (9.2 minus 6.6 equals 2.6 percentage point increase). Some may erroneously misinterpret the "39 percent higher" to mean that driving while on a cell phone increased from 6.6 percent to 45.6 percent. This isn't the case, yet it's easy to see how some misunderstanding or distorted findings can be communicated when not critically asking what the numbers represent.

To arrive at the 39 percent increase, you take the percentage point difference and then divide it by the original percentage (2.6/6.6) to get 0.394 or 39.4 percent (multiply by 100 to get a percentage). Note that the percentage point difference is 2.6, and the percentage increase over time is 39.4. The fact that those are two different numbers can create confusion in readers of data, and can sometimes purposely be used unethically to inflate or deflate actual change.

For example, imagine a shift from 20 percent of a sample in an election political poll supporting Candidate A to 30 percent a few months later. This ten percentage point increase indicates a modest shift in support for the struggling politician, who still does not have a majority. Simply take the percentage point difference and divide it by the original percentage (10/20 = 0.50). Yet, it would be easy to distort the findings to suggest a surge of interest in the candidate by reporting a "tremendous 50 percent improvement" in popularity. Although the statement would be accurate, those without numeracy skills might mistakenly think the candidate went up 50 points in the polls, from 20 percent to 70 percent, focusing on the number 50 rather than the much smaller 30.

Look at this paragraph from a *New York Times* article (Saulny, 2011) about an increase in multiracial people in the 2010 U.S. Census: "In North Carolina, the mixed-race population doubled. In Georgia, it expanded by more than 80 percent, and by nearly as much in Kentucky and Tennessee. In Indiana, Iowa and South Dakota, the multiracial population increased by about 70 percent." Note the word "doubled" and the large numbers, 80 and 70. Sounds impressive. A few paragraphs later the article reports a possible national multiracial growth rate of 35 percent, maybe even a 50 percent increase from the last census, in 2000. Again, these are large and impressive numbers. *What kinds of missing information do you need to better understand these percentages?*

Then the article states that in 2000 only 2.4 percent of Americans selected more than one race on the census form. It's one thing to claim that the multiracial population may increase 50 percent, but when the original figure is only 2.4 percent of Americans, a 50 percent increase simply means that the 2010 multiracial population could end up around 3.6 percent of the population. (50% of 2.4 = 1.2, and 2.4 + 1.2 = 3.6.) The number 50 surely sounds more impressive than the smaller figure, 3.6. Manipulating these numbers can create misleading impressions, sometimes unethically done with intention.

By the way, in the 2010 United States Census, 2.9 percent identified as two or more races, while the Pew Research Center (2015a) estimates that 6.9 percent of the U.S. population could be considered multiracial when including "how adults describe their own race as well as the racial backgrounds of their parents and grandparents." Understanding how researchers calculated the number and how it was measured are important questions to ask when interpreting data. *Which percentage makes more sense to you, 2.9 or 6.9? What questions would you ask to determine if respondents are multiracial?*

Being able to interpret numbers correctly often comes down to reporting all the relevant data. In the *New York Times* article, we really do not have enough information to know exactly what the multiracial population is in the states listed. That they doubled in North Carolina and increased 80 percent in Georgia tells us little about the actual 2000 or 2010 census figures in those states.

> CRITICAL THINKING TIP
>
> Always identify clearly the items being compared (when, where, and what) and express differences as percentages of the initial reference values. (See Miller, 2004.) After all, an increase of 75 cents when purchasing a cappuccino is more annoying (and a much higher percentage change) than the same amount tacked on to the price of a new car!

MARGIN OF ERROR

When political polls show that Candidate A has 46 percent of the popular vote and Candidate B has 42 percent, we cannot conclude that Candidate A is winning, despite announcements that A is ahead. Before making this statement, critical thinkers need to ask what the *margin of error* for the poll is. If the margin of error is plus or minus 4 percent, then it would suggest that Candidate A has between 42 percent and 50 percent support in the total population from which the sample was selected, and Candidate B has between 38 percent and 46 percent. There is a lot of overlap in support, and it could well be that A's true support percentage is, for example, 43 percent, whereas B's is 46 percent. Both those figures are within each candidate's margin of error.

The margin of error is based on sampling issues (see chapter 2) assuming that randomly selected respondents reflect the characteristics of a population with some, but not perfect, accuracy. Usually, the margin of error specifies how confident the researcher is in generalizing a statistical finding to the total population based on a random sample of a particular size drawn from that population. Here is the Pew Research Center's (2016a) statement about their survey practices:

> The sampling error for a typical Pew Research Center national survey of 1,500 completed interviews is plus or minus approximately 3 percentage points with a 95% confidence interval. This means that in 95 out of every 100 samples of the same size and type, the results we would obtain will vary by no more than plus or minus 3 percentage points from the result we would get if we could interview every member of the population. Thus, the chances are very high (95 out of 100) that any sample we draw will be within 3 points of the true population value.

CRITICAL THINKING TIP

When reading political polls or results from a survey, be sure to look for the margin of error. Then add and subtract that number from the results presented. Only then will you be able to critically analyze the findings and understand the range of possible results.

Consider the margin of error in Pew Research Center's (2015b) study about teenagers' friendships and romantic relationships in the digital age. As table 1 indicates, any results from the parents surveyed should

TABLE 1 MARGINS OF ERROR

Category	Sample size	Margin of error (percentage points)
All parents	1,060	±3.4
All teens	1,060	±3.7
Girls	537	±5.2
Boys	523	±5.3
White, non-Hispanic	614	±4.5
Black, non-Hispanic	101	±13.3
Hispanic	236	±8.1
Teen cellphone owners	929	±3.9
Teen smartphone owners	759	±4.4
Teen social media users	789	±4.3

SOURCE: Pew Research Center (2015b).

NOTE: These margins of error are based on the individual sizes of each subsample category and are used to interpret each subsample's answers to the survey questions (not reported here).

be viewed as accurate 95 percent of the time within plus or minus 3.4 percentage points of the figure reported. Notice how size of sample impacts the margin of error: With only 101 black teenagers surveyed, any responses to individual survey questions should be interpreted as being within plus or minus 13.3 percentage points of the reported result.

In the survey, 79 percent of all teens said that they instant-messaged their friends. With the margin of error for all teens at 3.7 percent, this means that if Pew were to complete 100 surveys with a similar random sample of 1,060 teens, 95 of the surveys would report a finding between 75.3 percent and 82.7 percent of teens texting their friends (79 – 3.7 and 79 + 3.7). That is, there is a high probability (95%) that the Pew Research Center sample finding of 79 percent is within plus or minus 3.7 percent of the true value in the total population.

Here's a situation where margin of error resulted in some arbitrary decisions with important political ramifications. For the Fox News Channel's (FNC's) first Republican Party U.S. presidential candidates' debate, in August 2015, FNC decided to use an average of several polls to determine the ten most popular participants. What's evident in table 2 is that the margins of error for each candidate in the average of five opinion polls made deciding who should be in and who should be out a difficult choice among those in the tenth through fourteenth positions (Kurtzleben, 2015). *How would you interpret these findings using the concept of the margin of error?*

TABLE 2 POLL DATA USED TO DETERMINE REPUBLICAN-PARTY PRESIDENTIAL-NOMINEE DEBATE PARTICIPANTS, 2015–16

Candidate	Average percent[a]	Margin of error (percentage points)
Donald Trump	23.4	±2.19
Jeb Bush	12.0	±1.66
Scott Walker	10.2	±1.51
Marco Rubio	5.4	±0.70
Rand Paul	4.8	±0.66
Mike Huckabee	6.6	±0.70
Ben Carson	5.8	±0.66
Ted Cruz	5.4	±0.70
Chris Christie	3.4	±0.43
John Kasich	3.2	±1.16
Rick Perry	1.8	±0.35
Bobby Jindal	1.4	±0.43
Rick Santorum	1.4	±0.43
Carly Fiorina	1.2	±0.66

SOURCE: Based on Kurtzleben, 2015.

[a]Average percent from multiple polls indicating voters' preferences for the Republican presidential nomination.

LEVELS OF MEASUREMENT

Calculating percentages and margins of error, however, depends on the level of measurement used in a study. *Level of measurement* refers to the type of values that characterize the elements of a variable. For example, imagine a survey asking, "What type of music do you listen to?" In this case, "music type" is a *variable* (a concept that varies); "folk," rap," "classical," and "show tunes" could be the *values* used to measure that variable. The method used to determine the values is called the level of measurement.

Consider for a moment the *Billboard* Top 200 music chart. *How do you create a list of the top best-selling music when the concept of an album has changed?* For years, the Top 200 entailed counting sales of physical vinyl records and CDs by scanning bar codes. Then along came purchasing digital music online (such as from iTunes and Amazon Prime) and streaming music (such as with Spotify and Pandora). *Billboard* (2014), based on Nielsen Entertainment's measurements, counts "10 digital track sales from an album to one equivalent album sale, and 1,500 song streams from an album to one equivalent album sale."

Why not 1,600 streams or nine digital tracks? As you can see, writing a question for a survey and measuring the answer is not necessarily a simple or objective process. Assessing how questions are posed and how concepts are measured are essential steps in critical thinking and very important in determining which levels of measurement to select.

Look back at how percentages and margins of error are calculated. The numbers must be such that they can be added, subtracted, divided, and multiplied. These basic mathematical operations cannot be performed on numerals, figures that look like numbers but really aren't. *When is a number not a number?*

Imagine a survey asking the following question:

What is your favorite type of movie?

1. Foreign language
2. Animation
3. Drama
4. Comedy

Note that a discrete number has been assigned to each answer category. Just as easily, 1 could have been assigned to animation, or 3 to comedy. That's because the numbers here are actually arbitrary numerals, with no intrinsic order (is drama "higher than" animation?) and no intensity (is comedy twice as meaningful as animation?). Letters could also have been selected, (a) Foreign Language; (b) Animation, and so forth. In such cases, we refer to this type of level of measurement as *Nominal* or *Categorical.* Any numbers assigned to the categories (values of a variable) are simply numerals, without any mathematical properties.

CRITICAL THINKING TIP

Do a quick, basic mathematical operation on the numbers you are reading to see if they make sense. Take your friends' phone numbers and add, subtract, multiply, and divide them. What do you get? Is there an average phone number? Maybe they're not numbers but numerals after all, and we should really ask for someone's "phone numeral" instead!

Sometimes the categories are in order, such as shirt sizes: 1, small; 2, medium; 3, large; 4, extra large. You could call 1 extra large; 2, large;

3, medium; 4, small; but you certainly cannot assign numbers this way: 1, medium; 2, small; 3, extra large; 4, large. Once the categories have been ordered, any discrete numeral must also be assigned in order from low to high or high to low. Hence, these are called *ordinal* measures. Yet, these numerals do not have any mathematical properties allowing them to be added, subtracted, multiplied, or divided. A small-sized shirt is not 25 percent (1/4 = 0.25) the size of an extra large one just because small is designated 1 and extra large is numbered 4. It is possible to have ordered numerals with equal-appearing intervals, called *Likert-type scales,* where 1 indicates strongly agree; 2, agree; 3, disagree; and 4, strongly disagree. Occasionally, these can be used as if they were actual numbers. (See the next section, on measurements of central tendency.)

Let's now consider measurements that have mathematical properties. In these cases, the numbers not only represent order but assume equal intervals between them. Some numbers are discrete (such as the number of books in a library, children in a family, or students enrolled in a class—with no fractions or numbers between the counting units), and others are continuous (like time, length, temperature, or age). These types of measurements are called *interval* or *ratio* measures. If there is an absolute zero—that is, no negative numbers, like age or weight—then you can calculate a ratio using multiplication and division. Given their mostly similar properties, these two measures are often labeled singly as *interval/ratio*.

> **CRITICAL THINKING TIP**
>
> Knowing the level of measurement is important when critically evaluating the correct use of graphs and charts, interpreting averages, and deciding whether appropriate statistics are being used. Ask how the variables in a report and questions in a survey are measured, and then decide whether the relevant statistics or charts are being employed.

CENTRAL TENDENCY MEASURES (AVERAGES)

Media reports often present survey results and other data in summary form. One common way is using averages, or more technically, *measures of central tendency*. Critically interpreting these measures requires evaluating what they are telling us about the summarized data and assessing whether the report is employing the best possible measure.

The media reported these numbers from the United States Census (2015):

Mean travel time to work (minutes), workers age 16+	25.7
Persons per household	2.63
Median household income	$53,482
Modal grade for 15-year-olds	10th grade

Where do we begin in understanding what is being reported? The first step in critically thinking about averages is to understand how a particular item is being measured. "Travel time" is the number of minutes it takes workers over the age of sixteen to get to work. Time is usually measured as a continuous-ratio variable, so mathematical operations can be performed, like adding them up and dividing by the total number to arrive at a calculation summary called the *mean.* All the people in the United States Census who completed that question about travel time gave a figure in minutes; these figures were summed and then divided by the number of people who answered the question. In popular jargon, when people refer to an average, it's usually this mathematical mean, as for example a GPA (grade-point average).

But take a look at the second reported average: 2.63 persons per U.S. household. Again, people responding to this question gave numbers that were all added together and then divided by the total number of respondents. Did some people say they were living with 0.63 of a person? *How can that be?* Well, the mathematical calculations on discrete interval/ratio numbers often end up with fractions of a person. Although a mean is correctly applied to interval/ratio data, when the numbers are not continuous, the resulting mean can sound pretty silly. The *World Factbook* (2015), for example, presents Total Fertility Rates (number of children born per woman) that sound peculiar, such as Afghanistan (5.33), Belgium (1.78), Canada (1.59), and Zimbabwe (3.53).

Related to the mathematical mean is an important concept called the *standard deviation.* Just knowing what the average is doesn't tell you much about what was measured. For example, someone may have a 3.0 (B) grade-point average because all her grades were B's. Yet, someone else could have a 3.0 (B) average with half his grades A and half C.

The mean does not tell the entire story. What is needed is a sense of the distribution of values that lead to a particular mathematical calculation.

A critical thinker is not satisfied simply with knowing the mean but inquires about the range of values obtained by a study. A standard deviation provides a more comprehensive description of what is found. A standard deviation of zero indicates that all the scores were the same; there was no variation from the mean. The student with a 3.0 GPA who had all B grades would have a standard deviation of zero. The larger the standard deviation, the more the scores are spread out (deviate) around the mean.

In some ways, the standard deviation is conceptually similar to the margin of error, in that it tells you what the probability is of finding a value close to or further away from the average. Recall how a margin of error indicates what range of values contains the true finding 95 percent of the time. Similarly, 95 percent of all values in a normal (bell-shaped) curve distribution are within two standard deviations below and above the mean. For example, consider a local high school's reading-comprehension test scores that are normally distributed with a mean score of 80 and a standard deviation of 7. Since it's greater than zero, we know that not all the students scored exactly 80, and so the grades must vary among the test takers. Based on the concept of the normal distribution, 95 percent of the students scored within two standard deviations above (7 + 7 = 14) and two standard deviations below the mean (–14), that is, between 94 (80 + 14) and 66 (80 – 14). Around 68 percent of them scored within one standard deviation above and below the mean. *What would be that range of scores?*

Sometimes an average can be a *median,* which is more appropriate for ordinal-level data or when there are extreme numbers (outliers) in the set of interval/ratio responses. The median (like the strip of land dividing a highway in half) is the point around which half the responses are higher and half are lower (the 50th percentile). Income is usually a good example, where the median is more relevant than the mathematical mean, as in the earlier United States Census example, since extreme incomes can distort the calculation and skew it in the direction of the few outlier scores.

Imagine you are calculating the mean age for kids hanging out in the local playground. To keep it small, let's say there are four kids, aged 1, 3, 4, and 6. The age of 3.5 would be the mathematical mean, and 3.5 would be the median point (halfway between 3 and 4) around which half the kids are older (the 4- and 6-year-olds) and half are younger than 3.5 (the two kids aged 1 and 3). Then all of a sudden a 16-year-old arrives in the park. The mean age now almost doubles, becoming 6 whereas the median increases incrementally to 4. When extreme scores appear in a set of data, the summary measure of central tendency should be the median. After all, if Mark Zuckerberg walked into a bar, everyone there would be a millionaire on average!

> CRITICAL THINKING TIP
>
> Whenever you read or hear about "an average," or something is "on average," first ask which kind of average is being used: mathematical mean, median, or mode. Then figure out how the item was measured (nominal, ordinal, interval/ratio) in order to make sure that the appropriate average has been selected. This is especially the case when there are extreme values, which require a median rather than a mean.

Take this example from the Pew Research Center's (2014) Internet Project survey: "Among adult Facebook users, the average (mean) number of friends is 338, and the median (midpoint) number of friends is 200." With the mean being much higher than the median, this suggests that among 960 Internet users sampled for the study, there are some people with much larger numbers of friends than the rest. The median is typically the best measure of central tendency to use when you have extreme numbers (either very high or very low) reported by some respondents. Note also how Pew correctly labeled the central tendency measures reported in their survey.

Finally, what to do with nominal data that have no order or numerical properties? At best, you can report only the most typical answer, called the *mode;* it's the response given in the greatest number of answers, though not always in the majority of them. The most common age for 10th graders is 15, as the census stated. Remember, a majority requires over 50 percent; and if one choice is over 50 percent, then it's certainly also the mode. But if one response is the one most often selected and it's under 50 percent, it's still the modal response even if not the majority answer. Modes can be used for nominal, ordinal, or interval/ratio data.

Using data from a General Social Survey (that asked its almost 2,500 respondents in 2014 how old they were, an interval/ratio measure), we find the mean was 49; median, 48; mode, 53, 56, and 58. These numbers tell us that the mathematical average age is 49, that half the sample is under 48 and half is over 48, and that the most common ages in the sample are 53, 56, 58. Those three ages each account for only 2.4 percent of the total, so clearly the finding is that none of those ages is the majority, only that this sample has multimodal ages (sometimes called bimodal—when there are only two modes).

ESTIMATIONS

During the 2016 U.S. presidential election, Donald Trump claimed Hillary Clinton wanted to let "650 million people pour" into the United States, thereby "tripling the size of our country in one week" (Valverde, 2016). *How realistic is this estimate of many millions of people entering a country in one week?* In 2014, survey respondents in Great Britain estimated that Muslims made up 21 percent of the U.K. population; Canadians thought that immigrants were 35 percent of their population; and Americans believed that 24 percent of girls between 15 and 19 years of age gave birth each year to a child (Nardelli and Arnett, 2014). Does it sound reasonable that so many teenaged girls give birth each year or that Canada had so many immigrants? Critical thinking requires you ask when hearing such claims: *How well do people know the real percentages when filling out a survey? How many of these numbers reflect actual percentages in those countries? How well can you make estimates and roughly judge the accuracy of the data?*

Learning to decipher the accuracy of numbers presented in studies and in the media is an important critical tool. One way, of course, is to get informed by seeking out original studies and relevant data to check claims made. Without immediate access to such information, however, a key numeracy and quantitative literacy skill is the ability to estimate and quickly recognize how close or off any reported number is. When reading survey results or hearing media reports filled with numbers and statistics, we should be able to estimate what a reasonable answer is before interpreting the findings and the accuracy of the reported figures.

Estimation is a close guess, often using some calculation or critical appraisal of the final value. As Steen and the Quantitative Literacy Design Team wrote (2001: 8): "Individuals who are quantitatively confident routinely use mental estimates to quantify, interpret, and check other information." We already make estimates in our everyday routines that are rough approximations of the actual outcome: How long it will take to drive to work at a given hour of the day; how much wrapping paper you need from the roll to fit around the birthday present you bought; or how many feet away you have to be to sink that ball in the basket. Estimating is a skill that improves with practice. It requires some knowledge of a standard or average number, often based on prior experience, in order to come up with some educated guess.

> CRITICAL THINKING TIP
>
> Ask yourself if the numbers that you are hearing or reading about are within a range you may expect. What would be your estimate for the results? If you're not sure, then inquire about where the data come from, what is being measured, who is calculating the findings, and how the researchers arrived at those figures.

Rounding off is a particularly useful trick when making estimates with numbers. If you were asked to figure out how many cups of coffee to have available for the reception following a speech by a well-known author, you would begin by estimating the number of people who could fill the auditorium. Say there are *about* 12 rows with *around* 14 seats in each, so in your head, round off to 10 rows by 15 seats for a quick estimate of at least 150 spaces. The actual number is 168, but 150 is close enough for a simple, quick estimate. If someone were to say that there would be room for 200 or 300 people sitting in on the lecture, you would immediately know that this is an overestimate.

Another method useful to estimating outcomes is comparing findings proportionately to their occurrence in a population. In the 2016 U.S. presidential election, about 50 percent of eligible citizens 18 to 29 years old voted. How do we estimate if this figure is high, low, or average? First we need to know what the overall voter turnout was for eligible voters. For that election, 58 percent of citizens voted. Without doing any statistical analysis, it looks like young adults voted at lower rates than the total population.

Another way of interpreting the percentage is to ask what proportion of the population is in that age range and evaluate whether the percentage of citizens 18–29 years old voting in 2016 is representative of that share of the population. Here we find that people 18–29 years old are around 17 percent of the U.S. citizen population, yet they make up about 19 percent of the votes cast (CNN, 2016). *What would you conclude from these different ways of reporting exit poll findings?*

From a math blog in England (Ellie's Active Maths, 2011):

> One of the best things about estimating is that we use our mathematical reasoning. There are often several ways to come to an answer, none of which is wrong. When a student comes up with an answer, we can ask them to explain, in words, their thought-process/reasoning. It encourages problem solving. Importantly, estimating allows us to check to see if our calculated answers are reasonable.

So it is with media stories and research studies: Estimating allows us to check to see if the information we are reading or hearing is reasonable. Is it really possible to have a city or school brag that all the children score above average on an intelligence test, now that you know that if it is a median average, half are above and half are below average? Is it reasonable to believe that one out of every four teenage girls in the United States gave birth to babies? (It's actually closer to 3 percent.) Could a little over a third of the Canadian population be immigrants? (It's actually around 20 percent.)

As we encounter various media reports, published survey findings, fake news sites, and politicians' pronouncements using numbers and other data, numeracy becomes a necessary tool in assessing the figures and measurements used. Learn to understand different kinds of measurements and averages. That's step one in critical thinking: interpreting what was found. We also need to assess who was studied and how by evaluating the methods that generated the sample, as chapter 2 shows.

KEY TERMS

ESTIMATIONS Close guesses, based on some rough calculation or by critically thinking what the findings should be.

INTERVAL/RATIO MEASURES Values of a variable in order with equal intervals and actual numbers that can be added, subtracted, multiplied, and divided.

LEVEL OF MEASUREMENT The type of value that characterizes the elements of a variable, such as nominal, ordinal, interval, and ratio.

MARGIN OF ERROR The differences between the true population statistic and the statistic found in a sample from that population, often specified in polls as plus or minus a percentage.

MEAN The mathematical measure of central tendency based on adding the values and dividing by the number of values.
MEDIAN A measure of central tendency that represents the halfway point among a range of ordered numbers.
MODE A measure of central tendency indicating the most frequently occurring value of a variable.
NOMINAL MEASURES Values of a variable using categories or numerals.
NUMERACY A critical thinking tool based on understanding, using, and evaluating numbers with confidence; quantitative literacy.
ORDINAL MEASURES Values of a variable that are in rank order.
PERCENTAGE A mathematical calculation that multiples a ratio by 100.
STANDARD DEVIATION An indication of the dispersion of values around a mathematical mean.

EXERCISES

1. An article on young-adult novels stated: "Thank J.K. Rowling for starting the kid's-book craze with 'Harry Potter' and Stephenie Meyer's 'Twilight' saga for perpetuating the trend that has more adults reading children's titles than ever before. The year 2011 has seen an explosion of books catering to this ever-expanding bimodal audience." (Carpenter, 2011). The typical reader might continue with the article and not raise any questions at this point. But as a *critical thinker,* you would want to raise and answer these questions: What is meant by the phrase "bimodal audience" and how do you think the author arrived at this statement? What data would you need to see? How would you go about conducting a survey to test out this idea; what kinds of questions would you ask to get different kinds of measurements besides the mode?

2. During a political campaign, one candidate tweeted excitedly to her followers that her polling numbers went up from 1 percent to 6 percent in three months. But it's less impressive when you consider that the *margin of error* for the latest poll is 4.7 percent (and was 5.3 percent for the first poll). Explain why the candidate needs to be more cautious about celebrating her poll numbers.

3. The media reports that the *average* American reads 12 books a year. When you first hear or read that headline, as a critical thinker, what questions do you need to ask? Then you see table 3, taken from a Pew Research Center report (2015c). Put into words what it's communicating to us about book reading in the United States.

TABLE 3 AVERAGE (MEAN) AMERICAN READERSHIP OVER THE TWELVE-MONTH PERIOD 2014–15[a]

	Median	*Mean*
Total (all adults ≥18)	*4*	*12*
Gender		
Male	*3*	*9*
Female	*5*	*14*
Ethnicity		
White, Non-Hispanic	*5*	*13*
Black, Non-Hispanic	*3*	*8*
Hispanic	*2*	*8*
Age		
18–29	*4*	*11*
30–49	*4*	*12*
50–64	*3*	*12*
≥65	*3*	*12*
Level of Education		
High school graduate or less	*1*	*7*
Some college	*5*	*13*
College+	*7*	*17*

SOURCE: *Pew Research Center (2015c).*

[a]Among all American adults ages 18+ (including nonreaders), the mean (average) and median (midpoint) number read in whole or in part by each group.