10
Evaluating and Interpreting Information

Not all information is equally valuable. Not all interpretations are equally valid. For instance, if you really want to know how well the latest innovation in robotic surgery works, you need to check with other sources besides, say, its designer (from whom you could expect an overly optimistic or insufficiently critical assessment).

Whether you work with your own findings or those of other researchers, you need to decide if the findings are valid and reliable. Then you need to decide what your information means. Figure 10.1 outlines this challenge.

EVALUATE THE SOURCES

Not all data sources are equally dependable. A source might offer information that is out of date, inaccurate, incomplete, mistaken, or biased.

Is the Source Up-to-Date?

Even newly published books contain information that can be more than a year old, and journal articles often undergo a lengthy process of peer review.

Certain types of information become outdated more quickly than others. For topics that focus on *technology* (multimedia law, superconductivity, alternative cancer treatments), information more than a few months old may be outdated. But for topics that focus on *people* (business ethics, management practices, workplace gender equality), historical perspectives often help.

NOTE *The most recent information is not always the most reliable—especially in scientific research, a process of ongoing inquiry in which what seems true today may be proven false tomorrow. Consider, for example, the recent discoveries of fatal side effects from some of the latest "miracle" weight-loss drugs.*

Is the Printed Source Dependable?

One way to assess a publication's reputation is to check its copyright page. Is the work published by a university, professional society, museum, or respected news organization? Do members of the editorial and advisory board have distinguished titles and degrees? Is the publication *refereed* (all submissions reviewed by experts before acceptance)?

Does the bibliography or list of references show that the author has extensively researched the issue (Barnes)?

One way to assess an author's reputation is to check citation indexes (page 142) to see what other experts have said about this research. Many periodicals also provide brief biographies or descriptions of authors' earlier publications and other achievements.

Is the Electronic Source Trustworthy?

The Internet offers information that may never appear in other sources, for example from listservs and newsgroups. But much of this material may reflect the bias of the special-interest groups that provide it. Moreover, anyone can publish almost anything on the Internet—including misinformation—without having it verified, edited, or reviewed for accuracy. Don't expect to find everything you need on the Internet. (Pages 170–71 offer suggestions for evaluating sources on the Web.)

Even in a commercial database, such as DIALOG, decisions about what to include and what to leave out depend on the biases, priorities, or interests of those who assemble the database.

NOTE *Because it advocates a particular point of view, a special-interest Web site (as in Figure 10.2) can provide useful clues about the ideas and opinions of its sponsors; however, don't rely on special-interest sites for factual information—unless the facts can be verified elsewhere ("Evaluating Internet-Based Information").*

Is the Information Relatively Unbiased?

Much of today's research is paid for by private companies or special-interest groups that have their own social, political, or economic agendas (Crossen 14, 19). Medical research may be sponsored by drug or tobacco companies; nutritional research, by food manufacturers; environmental research, by oil or chemical companies. Public policy research (on gun control, school prayer, endangered species) may be sponsored by opposing groups (environmentalists versus the logging industry) producing opposing results.

Instead of a neutral and balanced inquiry, this kind of "strategic research" is designed to support one special interest or another (Crossen 132–34). Furthermore, those who pay for strategic research are not likely to publicize findings that contradict their original claims or opinions (profits lower than expected, losses or risks greater than expected). As consumers of research, we should try to determine exactly what the sponsors of a particular study stand to gain or lose from the results (234).

NOTE *Keep in mind that any research ultimately stands on its own merits. Thus, funding by a special interest should not automatically discredit an otherwise valid and reliable study.*

Also, financing from a private company often sets the stage for beneficial research that might otherwise be unaffordable, as when research funded by Quaker Oats led to other studies proving that oats can lower cholesterol (Raloff, "Chocolate Hearts" 189).

How Does This Source Measure Up to Others?

Most studies have some type of flaw (page 189). Therefore, instead of relying on a single source of study, seek a consensus among various respected sources (Cohn 106).

NOTE Some issues (the need for defense spending or the causes of inflation) are always controversial and will never be resolved. Although we can get verifiable data and can reason persuasively on some subjects, no close reasoning by any expert and no supporting statistical analysis will "prove" anything about a controversial subject. Some problems are simply more resistant to solution than others, no matter how dependable the sources.

EVALUATE THE EVIDENCE

Evidence is any finding used to support or refute a particular claim. While evidence can serve the truth, it can also create distortion, misinformation, and deception. For example:

- How much money, material, or energy does recycling really save?
- How well are public schools educating children?
- Which investments or automobiles are safest?

Competing answers to such questions often rest on evidence that has been stacked to support a particular view or agenda.

Is the Evidence Sufficient?

Evidence is sufficient when nothing more is needed to reach an accurate judgment or conclusion. A study of the stress-reducing benefits of low-impact aerobics, for example, would require a broad survey sample: people who have practiced aerobics for a long time; people of both genders, different ages, different occupations, and different lifestyles before they began aerobics; and so on. Even responses from hundreds of practitioners might be insufficient unless those responses were supported by laboratory measurements of metabolic and heart rates, blood pressure, and so on.

NOTE *Although anecdotal evidence ("This worked great for me!") might be a good starting point for an investigation, your personal experience rarely provides enough evidence from which to generalize. No matter how long you might have practiced aerobics, for instance, you cannot tell whether your experience is representative.*

Is the Presentation of Evidence Balanced and Reasonable?

Misuse of evidence in a courtroom often makes headlines. But evidence routinely is misused beyond the courtroom as well, as in the following instances.

OVERSTATEMENT. Consumers are offered a daily menu of cures for ailments ranging from insomnia to cancer. Overzealous researchers might exaggerate their achievements, without mentioning the limitations of their study.

OMISSION OF VITAL FACTS. Aspirin is widely promoted for *decreasing* heart attack and stroke risk caused by clotting, but far less emphasized is its role in *increasing* stroke risk caused by brain bleeding (Lewis 222). Acetaminophen products are advertised as a ''safe'' alternative pain reliever, without aspirin's side effects (stomach irritation, Reye's syndrome), but even small overdoses of acetaminophen have caused liver failure. Moreover, acetaminophen is the leading cause of U.S. drug fatalities (Easton and Herrara 42-44).

DECEPTIVE FRAMING OF THE FACTS. A *frame of reference* is a set of ideas, beliefs, or views that influences our interpretation of other ideas. For example, people with a fundamentalist view of the Bible might reject the concept of evolution. Or consider the well-known optimist/pessimist test: Is the glass half full (a positive frame of reference) or half empty (a negative frame of reference)? In medical terms, is a ''90-percent survival rate'' more acceptable than a ''10-percent mortality rate''? Framing sways our perception (Lang and Secic 239-40). For example, what we now term a ''financial recession'' used to be a ''financial depression,'' which was a euphemism for ''financial panic'' (P. Bernstein 183). For more on euphemisms, see page 268.

The framing of survey questions can manipulate responses. For example, in a survey of attitudes toward abortion, consider how responses might differ depending on the following phrasing: “Do you approve of abortion on demand?” versus “Do you approve of abortion under any circumstances?” (Phillips 192).

Even unintentional framing can have major consequences. Researchers Kahneman and Tversky describe this situation at one hospital:

> [D]octors were concerned that they might be influencing patients who had to choose between the life-or-death risks in different forms of treatment. The choice was between radiation and surgery in the treatment of lung cancer. Medical data ... showed that no patients die during radiation but have a shorter life expectancy than patients who survive the risk of surgery; the overall difference in life expectancy was not great enough to provide a clear choice between the two forms of treatment. When the question was put in terms of risk of death during treatment, more than 40% of the patients favored radiation. When the question was put in terms of life expectancy, only about 20% favored radiation. (cited in P. Bernstein 276)

Whether the language is provocative (“rape of the environment”), euphemistic (“teachable moment” versus “mistake”), or demeaning to opponents (“bureaucrats,” “tree huggers”), deceptive framing obscures the real issues. Common “spin” strategies of politicians, for example, include painting situations as rosier than they are or calling people names.

Can the Evidence Be Verified?

Hard evidence consists of factual statements, expert opinion, or statistics that can be verified. *Soft evidence* consists of uninformed opinion or speculation obtained or analyzed unscientifically, and findings that have not been replicated or reviewed by experts. Reputable news organizations employ fact-checkers to verify information before it appears in print.

INTERPRET YOUR FINDINGS

Interpreting means trying to reach the truth of the matter: an overall judgment about what the findings mean and what conclusion or action they suggest.

Unfortunately, research does not always yield answers that are clear or conclusive. Instead of settling for the most *convenient* answer, we pursue the most reasonable answer by critically examining a full range of possible meanings.

What Level of Certainty Is Warranted?

As possible outcomes of research we can identify three distinct and very different levels of certainty:

1. The ultimate truth—*the conclusive answer:*

> Truth is what is so about something, the reality of the matter, as distinguished from what people wish were so, believe to be so, or assert to be so.... In the words of Harvard philosopher Israel Scheffler, truth is the view “which is fated to be ultimately agreed to by all who investigate.” The word *ultimately* is important. Investigation may produce a wrong answer for years, even for centuries. ...
>
> One easy way to spare yourself any further confusion about truth is to reserve the word *truth* for the final answer to an issue. Get in the habit of using the words *belief, theory,* and *present understanding* more often. (Ruggiero, 3rd ed. 21–22)

Conclusive answers are the research outcome we seek, but often we have to settle for answers that are less than certain.

2. The *probable answer:* the answer that stands the best chance of being true or accurate—given the most we can know at this particular time. Probable answers are subject to revision in light of new information. This is especially the case with *emergent* science, which one expert defines as “science whose truth has not yet been settled by consensus” (Hornig-Priest 97). Examples include risks versus benefits of cloning, food irradiation, or genetically modified crops.
3. The *inconclusive answer:* the realization that the truth of the matter is more elusive, ambiguous, or complex than we expected.

To ensure an accurate outcome, we must decide what level of certainty our findings warrant. For example, we are *highly certain* about the perils of smoking or sunburn, *reasonably certain* about the health benefits of fruits and vegetables and moderate exercise, and *less certain* about the perils of coffee drinking or electromagnetic waves, or the benefits of vitamin supplements.

The “truth” never changes; however, our notions about “truthfulness” do, as in these examples:

- “*The earth is the center of the universe.*” Though dead wrong, Ptolemy’s cosmology was based on the best information available in the second century A.D. And this certainty survived thirteen centuries—even after new information had discredited Ptolemy’s theory. When Copernicus and Galileo proposed more truthful views in the fifteenth century, they were labeled heretics.
- *“Brush your teeth and blow your nose with asbestos.”* Considered a “miracle fiber” for two thousand years, asbestos—soft, flexible, and fire resistant—was used in countless products ranging from asbestos handkerchiefs in the first century to tablecloths, toothpaste, and cigarette filters in the twentieth century. Not until the 1970s was the truth about the long-suspected role of asbestos in lung disease publicized (Alleman and Mossman 70–74).
- *“Fat is bad, Carbs are good.”* This was the basic message of The Food Guide Pyramid, introduced in 1992 by the U.S. government. The Pyramid has largely been turned upside-down by research indicating that certain carbohydrates may actually promote chronic disease while certain fats may help prevent it (Willett and Stampfer 66).

Communication expert Katherine Rowan reminds us just how elusive certainty can be, in any context:

> ... because more refined, more precise, better explanation is always possible, no phenomenon, no matter how thoroughly studied, is ever fully explained or understood. In this sense, all scientific knowledge is uncertain. (204–05)

Of course, unethical communicators can downplay certainty, as tobacco companies did for decades. Or they can exaggerate certainty, as in recent claims about the absolute safety of genetically modified foods. Each of these ploys is a type of *argument from ignorance,* in which the absence of evidence to the contrary is offered as “proof” of something: “X is true because it has not been proven false” (or vice versa). For example, “Since no one has yet demonstrated harmful side effects from genetically modified food, it must be safe.”

The vast majority of scientific, technical, and social controversies are open-ended. Therefore, no irrefutable presentation of “the facts” will likely settle the controversy over questions like these:

- How rapidly is global warming progressing?
- What is causing the death and disfigurement of frogs worldwide?

- Can vitamins prevent cancer or heart disease?
- Should Affirmative Action programs be expanded or discontinued?

Does this mean that such questions should be ignored? Of course not. Even though some claims cannot be proven, we can still reach reasonable conclusions on the basis of the evidence.

Are the Underlying Assumptions Sound?

Assumptions are notions we take for granted, things we often accept without proof. The research process rests on assumptions like these: that a sample group accurately represents a larger target group, that survey respondents remember certain facts accurately, that mice and humans share enough biological similarities for meaningful research. For a particular study to be *valid* (page 188), the underlying assumptions have to be accurate.

Consider this example: You are an education consultant evaluating the accuracy of IQ testing as a predictor of academic performance. Reviewing the evidence, you perceive an association between low IQ scores and low achievers. You then verify your statistics by examining a cross-section of reliable sources. Can you then conclude that IQ tests *do* predict performance accurately? This conclusion might be invalid unless you could verify the following assumptions:

1. That neither parents, teachers, nor children had seen individual test scores, which could produce biased expectations.
2. That, regardless of score, each child had completed an identical curriculum, instead of being "tracked" on the basis of his or her score.

NOTE *Assumptions are often easier to identify in someone else's thinking and writing than in our own. During collaborative discussions, ask group members to help you identify your own assumptions (Maeglin).*

To What Extent Has Personal Bias Influenced the Interpretation?

To support a particular version of the truth, our own bias might cause us to overestimate (or deny) the certainty of our findings.

> Unless you are perfectly neutral about the issue, an unlikely circumstance, at the very outset ... you will believe one side of the issue to be right, and that belief will incline you to ...

> present more and better arguments for the side of the issue you prefer. (Ruggiero, 3rd ed. 134)

The following example illustrates how *cognitive bias* (seeing what we expect to see) can blind us to the most compelling evidence. The 1989 spill from the oil tanker *Exxon Valdez* polluted more than 1,000 miles of Alaskan shoreline and led to a massive recovery effort that included using high-pressure hot water to clean oil from the beaches. But a respected study shows that the *uncleaned* beaches are now healthier than those sterilized by the hot water: "Whatever [beach life] survived the oiling did not survive the cure" (Holloway 109). But this finding—that cleaning up is more harmful than helpful—remains highly unpopular with the Alaskan public, who continue to insist on the removal of virtually every drop of oil (109–12).

Because personal bias is hard to transcend, rationalizing often becomes a substitute for reasoning:

> You are reasoning if your belief follows the evidence—that is, if you examine the evidence first and then make up your mind. You are rationalizing if the evidence follows your belief—if you first decide what you'll believe and then select and interpret evidence to justify it. (Ruggiero, 3rd ed. 44)

Personal bias often is subconscious until we examine our own value systems: attitudes long held but never analyzed, assumptions we've inherited from our own backgrounds, and so on. Recognizing our own biases is the crucial first step in managing them.

Are Other Interpretations Possible?

Perhaps other researchers disagree with the meaning of these findings. Settling on a final meaning can be hard. For instance, how should we interpret the reported increase in violent crime on U.S. college campuses—especially in light of statistics that show violent crime decreasing in general (Lederman 5)? Should we conclude (a) that college students are becoming more violent, (b) that some drugs and guns in high schools end up on campuses, (c) that off-campus criminals see students as easy targets, or (d) all of these? Or could these findings mean something else entirely? For example, (a) that increased law enforcement has led to more campus arrests—and thus, greater recognition of the problem or (b) that campus crimes really haven't increased, but more are being reported? Depending on our interpretation, we might conclude that the situation is worsening—or improving!

NOTE *Not all interpretations are equally valid. Never assume that any interpretation that is possible is also allowable—especially in terms of its ethical consequences. Certain interpretations in the college crime example, for instance, might justify an overly casual or overly vigilant response—either of which could have disastrous consequences.*

AVOID ERRORS IN REASONING

Finding the truth, especially in a complex issue or problem, often is a process of elimination, of ruling out or avoiding errors in reasoning. As we interpret, we make *inferences:* We derive conclusions about what we don't know by reasoning from what we do know (Hayakawa 37). For example, we might infer that a drug that boosts immunity in laboratory mice will boost immunity in humans, or that a rise in campus crime statistics is caused by the fact that young people have become more violent. Whether a particular inference is on target or dead wrong depends largely on our answers to one or more of these questions:

- To what extent can these findings be generalized?
- Is *Y* really caused by *X*?
- How much can the numbers be trusted, and what do they mean?

Following are three major reasoning errors that can distort our interpretations.

Faulty Generalization

We engage in faulty generalization when we jump from a limited observation to a sweeping conclusion. Even "proven" facts can invite mistaken conclusions.

1. "Some studies have shown that gingko [an herb] improves mental functioning in people with dementia [mental deterioration caused by maladies such as Alzheimer's Disease]" (Stix 30).
2. "For the period 1992–2005, two thirds of the fastest-growing occupations [called] for no more than a high-school degree" (Harrison 62).
3. "Adult female brains are significantly smaller than male brains—about 8% smaller, on average" (Seligman 74).

1. Gingko is food for the brain!
2. Higher education ... Who needs it?!
3. Women are the less intelligent gender.

When we accept findings uncritically and jump to conclusions about their meaning (as in points 1 and 2, above) we commit the error of *hasty generalization*. When we overestimate the extent to which the findings reveal some larger truth (as in point 3, above) we commit the error of *overstated generalization*.

NOTE *We often need to generalize, and we should. For example, countless studies support the generalization that fruits and vegetables help lower cancer risk. But we ordinarily limit general claims by inserting qualifiers such as "usually," "often," "sometimes," "probably," "possibly," or "some."*

Faulty Causal Reasoning

Causal reasoning tries to explain why something happened or what will happen, often very complex questions. Sometimes a *definite cause is apparent* ("The engine's overheating is caused by a faulty radiator cap"). We reason about definite causes when we explain why the combustion in a car engine causes the wheels to move, or why the moon's orbit makes the tides rise and fall.

But causal reasoning often explores *causes that are not so obvious, but only possible or probable*. In these cases, much analysis is needed to isolate a specific cause.

Suppose you want to answer this question: "Why does our college campus not have daycare facilities?" Brainstorming produces these possible causes:

- lack of need among students
- lack of interest among students, faculty, and staff
- high cost of liability insurance
- lack of space and facilities on campus
- lack of trained personnel
- prohibition by state law
- lack of government funding for such a project

Say you proceed with interviews, questionnaires, and research into state laws, insurance rates, and availability of personnel. First, you rule out some items, and others appear as probable causes. Specifically, you find a need among students, high campus interest, an abundance of qualified people for staffing, and no state laws prohibiting such a project. Three probable causes remain:
lack of funding, high insurance rates, and lack of space. Further inquiry shows that lack of funding and high insurance rates *are* issues. These obstacles,
however, could be eliminated through new sources of

revenue: charging a fee for each child, soliciting donations, diverting funds from other campus organizations, and so on. Finally, after examining available campus space and speaking with school officials, you arrive at one definite cause: lack of space and
facilities.

When you report on your research, be sure readers can draw conclusions identical to your own on the basis of the evidence. The process might be diagrammed like this:

The persuasiveness of your causal argument will depend on the quality of evidence you bring to bear, as well as on your ability to clearly explain the links in the chain. Also, you must convince your audience that you haven't overlooked important alternative causes.

NOTE *Anything but the simplest effect is likely to have multiple causes. You have to make the case that the cause you have isolated is the real one. In the daycare scenario, for example, you might argue that lack of space and facilities somehow is related to funding. And the college's inability to find funds or space might be related to student need or interest, which is not high enough to exert real pressure. Lack of space and facilities, however, appears to be the "immediate" cause.*

Here are common errors that distort or oversimplify cause-effect relationships:

Investment builds wealth. [Ignores the roles of knowledge, wisdom, timing, and luck in successful investing.]

Running improves health. [Ignores the fact that many runners get injured, and that some even drop dead while running.]

Right after buying a rabbit's foot, Felix won the state lottery. [Posits an unwarranted causal relationship merely because one event follows another.]

Women in Scandinavian countries drink a lot of milk. Women in Scandinavian countries have a high incidence of breast cancer. Therefore, milk must be a cause of breast cancer. [The association between these two variables might be mere coincidence and might obscure other possible causes, such as environment, fish diet, and genetic predisposition (Lemonick 85).]

My grades were poor because my exams were unfair. [Denies the real causes of one's failures.]

Media Researcher Robert Griffin identifies three criteria for demonstrating a causal relationship:

Along with showing correlation [say, an association between smoking and cancer], evidence of causality requires that the alleged causal agent occurs prior to the condition it causes (e.g., that smoking precedes the development of cancers)

> and—the most difficult task—that other explanations are discounted or accounted for (240).

For example, epidemiological studies found that people who eat lots of broccoli, cauliflower, and other "cruciferous" vegetables have lower rates of some cancers. But other explanations (say, that big veggie eaters might have many other healthful habits as well) could not be ruled out until lab studies showed how a special protein in these vegetables actually protects human cells (Wang 182). For more examples of causal claims, see Consider This, page 182.

AVOID STATISTICAL FALLACIES

The purpose of statistical analysis is to determine the meaning of a collected set of numbers. In primary research, our surveys and questionnaires often lead to some kind of numerical interpretation ("What percentage of respondents prefer X?" "How often does Y happen?"). In secondary research, we rely on numbers collected by primary researchers.

Numbers seem more precise, more objective, more scientific, and less ambiguous than words. They are easier to summarize, measure, compare, and analyze. But numbers can be totally misleading. For example, radio or television phone-in surveys often produce distorted data: Although "90 percent of callers" might express support for a particular viewpoint, callers tend to be those with the greatest anger or extreme feelings—representing only a fraction of overall attitudes (Fineman 24). Mail-in surveys can produce similar distortion because only people with certain attitudes might choose to respond.

Before relying on any set of numbers, we need to know exactly where they come from, how they were collected, and how they were analyzed.

Common Statistical Fallacies

Faulty statistical reasoning produces conclusions that are unwarranted, inaccurate, or downright deceptive. Here are some typical fallacies.

THE SANITIZED STATISTIC. Numbers can be manipulated (or ''cleaned up'') to obscure the facts. For example, the College Board's 1996 recentering of SAT scores has raised the ''average'' math score from 478 to 500 and the average verbal score from 424 to 500 (a boost of almost 5 and 18 percent, respectively) although actual student performance remains unchanged (Samuelson, ‘‘Merchants’’ 44).

THE MEANINGLESS STATISTIC. Exact numbers can be used to quantify something so inexact, vaguely defined, or hard to count that it should only be approximated (Huff 247; Lavin 278):

- Boston has 3,247,561 rats.
- Zappo detergent makes laundry 10 percent brighter.

An exact number looks impressive, but certain subjects (child abuse, cheating in college, virginity, drug and alcohol abuse on the job, eating habits) cannot be quantified exactly because respondents don’t always tell the truth (because of denial, embarrassment, or merely guessing). Or they respond in ways they think the researcher expects.

Overly precise numbers are also harder to read:

> Computer engineers’ salaries increased from $34,717 to $41,346, while programmers’ salaries increased from $26,807 to $32,112.
>
> Computer engineers’ salaries increased roughly from $35,000 to $41,000, while programmers’ salaries increased from $27,000 to $32,000.

Unless greater precision is required, round your numbers to a maximum of two significant digits (“0.00”) (Lang and Secic 40).

THE UNDEFINED AVERAGE. The mean, median, and mode are confused in determining an average (Huff 244; Lavin 279).

- The *mean* is the result of adding up the value of each item in a set of numbers, and then dividing by the number of items.
- The *median* is the result of ranking all the values from high to low, and then identifying the middle value (or the 50th percentile, as in calculating SAT scores).
- The *mode* is the value that occurs most often in a set of numbers.

Each of these measurements represents some kind of average, but unless we know which average is being presented, we cannot interpret the figures accurately.

Assume, for instance, that we are calculating the “average” salary among female managers at XYZ Corporation:

Manager	**Salary**
A	$90,000
B	90,000
C	80,000
D	65,000
E	60,000
F	55,000
G	50,000

In the above example, the "mean salary" (total salaries divided by number of salaries) equals $70,000; the "median salary" (middle value) equals $65,000; the "mode" (most frequent value) equals $90,000. Each is legitimately an average, and each could be used to support or refute a particular assertion (for example, "Women managers are paid too little" or "Women managers are paid too much").

Research expert Michael Lavin sums up the potential for bias in the reporting of averages:

> Depending on the circumstances, any one of these measurements [*mean*, *median*, or *mode*] may describe a group of numbers better than the other two. . . . [But] people typically choose the value which best presents their case, whether or not it is the most appropriate. (279)

Although the mean is the most commonly computed average, this measurement can be misleading when values on either end of the scale (*outliers*) are extremely high or low. Suppose, for instance, that Manager A (above) was paid a $200,000 salary. Because this figure deviates so far from the normal range of salary figures for B through G, it distorts the average for the whole group—increasing the mean salary by more than 20 percent (Plumb and Spyridakis 636). In this instance, the median figure—unaffected by outlying values—would present a more realistic picture.

One drawback of the "median" figure is that it ignores extreme values. But the "mode" focuses only on the "typical" values, ignoring the data's *range* (distance between highest and lowest values), as well as their variability. In short, the clearest picture of a data set may require two of these measures—if not all three.

NOTE Failure to report "outliers" (the small percentage of the results occurring at a great distance from the mean) in calculating the mean or secretly leaving the outliers out of the calculation is deceptive. An ethical approach is to calculate the results with and without the outliers and to report both figures (Lang and Secic 31).

THE DISTORTED PERCENTAGE FIGURE. Percentages are often reported without explanation of the original numbers used in the calculation (Adams and Schvaneveldt 359; Lavin 280): ''Seventy-five percent of respondents prefer our brand over the competing brand''-—without mention that only four people were surveyed. Three out of four respondents is a far less credible number than, say, 3,000 out of 4,000-—yet each of these ratios equals ''75 percent.''

NOTE *In small samples, percentages can mislead because the percentage size can dwarf the number it represents: "In this experiment, 33% of the rats lived, 33% died, and the third rat got away." (Lang and Secic 41.) When your sample is small, report the actual numbers: "Five out of ten respondents agreed . ..."*

Another fallacy in reporting percentages occurs when the *margin of error* is
ignored. This is the margin within which the true figure lies, based on estimated sampling errors in a survey. For example, a claim that "the majority of people surveyed prefer Brand X" might be based on the fact that 51 percent of respondents expressed this preference; but if the survey carried a 2 percent margin of error, the true figure could be as low as 49 or as high as 53 percent. In a survey with a high margin of error, the true figure might be so uncertain that no definite conclusion can be drawn.

THE BOGUS RANKING. Items are compared on the basis of ill-defined criteria (Adams and Schvaneveldt 212; Lavin 284): ''Last year, the Batmobile was the number-one selling car in Gotham City''-—without mentioning that some competing car makers actually sold *more* cars to private individuals, and that the Batmobile figures were inflated by hefty sales-—at huge discounts-—to rental car companies and corporate fleets. Unless we know how the ranked items were chosen and how they were compared (the *criteria*), a ranking can produce a scientific-seeming number based on a completely unscientific method.

The Limitations of Number Crunching

Computers are great for comparing, synthesizing, and predicting—because of their speed, processing power, and "what if" capabilities. But limitless capacity to crunch numbers cannot guarantee accurate results.

DRAWBACKS OF DATA MINING. Many highly publicized correlations are the product of *data mining:* in this process computers randomly compare one set of variables (say, eating habits) with another set (range of diseases). From these countless comparisons, certain relationships, or associations, are revealed (say, between coffee drinking and pancreatic cancer risk). At one retail company, data mining revealed a correlation between diaper sales and beer sales (presumably because young fathers go out at night to buy diapers). The retailer then displayed the diapers next to the beer and reportedly sold more of both (Rao 128).

Because it detects hidden relationships and has predictive power, data mining is a popular business tool. Companies assemble their own *data warehouses* (databases of information about customers, research and development, market conditions, legal matters). Much of what is "mined," however, can turn out to be trivial or absurd:

- Venereal disease rates correlate with air pollution levels (Stedman, "Data Mining" 28).
- The hourly wages of self-proclaimed "natural blondes" (both male and female) averaged 75 cents more than the wages of nonblonde people in 1993 (Sklaroff and Ash 85).

NOTE *Despite its limitations, data mining is invaluable for "uncovering correlations that require computers to perceive but that thinking humans can evaluate and research further" (Maeglin).*

THE BIASED META-ANALYSIS. In a meta-analysis, researchers look at a whole range of studies that have been done on one topic (say, the role of high-fat diets in cancer risk). The purpose of this ''study of studies'' is to decide on the overall meaning of these collected findings.

As "objective" as it may seem, meta-analysis does carry potential for error (Lang and Secic):

- *Selection bias.* Because results ultimately depend on which studies have been included and which omitted, a meta-analysis can reflect the biases of the researchers who select the material (174).
- *Publication bias.* Small studies have less chance of being published than large ones (175–76).

- *"Head counting."* This questionable method tallies the studies that have positive, negative, or insignificant results, and announces the winning category based on sheer numbers—without accounting for a study's relative size, design, or other differences.

THE FALLIBLE COMPUTER MODEL. Computer models process complex assumptions to produce impressive but often inaccurate statistical estimates about costs, benefits, risks, or probable outcomes.

Computer models to predict global warming levels, for instance, are based on differing assumptions about wind and weather patterns, cloud formations, ozone levels, carbon dioxide concentrations, sea levels, or airborne sediment from volcanic eruptions. Despite their seemingly scientific precision, different global warming models generate 50-year predictions of sea-level rises that range from a few inches to several feet (Barbour 121). Other models suggest that warming effects could be offset by evaporation of ocean water and by clouds reflecting sunlight back to outer space (Monatersky 69). Still other models suggest that the 1-degree Fahrenheit warming over the last 100 years might not be the result of the greenhouse effect at all, but of "random fluctuations in global temperatures" (Stone 38).

Choice of assumptions might be influenced by researcher bias or the sponsors' agenda. For example, a prediction of human fatalities from a nuclear reactor meltdown might rest on assumptions about availability of safe shelter, evacuation routes, time of day, season, wind direction, and structural integrity of the containment unit. But the assumptions could be manipulated to produce an overstated or understated estimate of risk (Barbour 228). For computer-modeled estimates of accident risk (oil spill, plane crash) or of the costs and benefits of a proposed project or policy (a space station, welfare reform), consumers rarely know the assumptions behind the numbers.

According to risk expert Peter Bernstein, the major limitation of computer modeling is that it tries to predict the future on the basis of information from the past (334).

Misleading Terminology

The terms used to interpret statistics sometimes hide their real meaning.

- *People treated for cancer have a "50 percent survival rate."* This widely publicized figure is misleading in two ways: (1) "survival," to laypersons, means "staying alive," but to medical experts, staying alive for only five years after diagnosis qualifies as survival; (2) the "50 percent" survival figure covers *all* cancers, including certain skin or thyroid cancers that have extremely high *cure rates*, as well as other cancers (such as lung or ovarian) that rarely are curable and have extremely low *survival rates* ("Are We" 5; *Facts and Figures* 2).

- "*The BODCARE Health Plan is proud of achieving 99 percent customer satisfaction.*" The rate of customer "satisfaction," in this context, provides little real information about the quality of the health plan itself. Because most HMO customers are reasonably healthy, they rarely visit a doctor and so have little cause for dissatisfaction. Only those in poor health would be able to evaluate "quality of service." Also, customer service is not the same as "quality of treatment"—which is much harder for laypersons to evaluate. Finally, are we talking about customers who are "satisfied," "somewhat satisfied," or "very satisfied"—all combined into one result (Spragins 77)?

INTERPRET THE REALITY BEHIND THE NUMBERS

Even the most valid and reliable statistics require that we interpret the reality behind the numbers:

- *"Rates for certain cancers double after prolonged exposure to electromagnetic radiation."* Does the cancer rate actually increase from "1 in 10,000" to "2 in 10,000," or from "1 in 50" to "2 in 50"? Without knowing the *base rate* (original rate) we cannot possibly decide how alarming this "doubling" of risk actually is.
- *"Saabs and Volvos are involved in 75 percent fewer fatal accidents than average."* Is this only because of superior engineering or also because Saab and Volvo owners tend to drive very carefully ("The Safest" 72)?
- *"Eating charbroiled meat may triple the risk of stomach cancer."* The actual added risk of dying from a weekly charbroiled steak is 1 in 4 million—lower than the risk of drowning in the bathtub (Lee 280).

The above numbers may be "technically accurate" and may seem highly persuasive in the interpretations they suggest. But the actual "truth" behind these numbers is far more elusive. Any interpretation of statistical data carries the possibility that other, more accurate interpretations have been overlooked or deliberately excluded (Barnett 45).

ACKNOWLEDGE THE LIMITS OF RESEARCH

Legitimate researchers live with uncertainty. They expect to be wrong far more often than right. Experimentation and exploration often produce confusion, mistakes, and dead ends. Following is a brief list of things that go wrong with research and interpretation.

Obstacles to Validity and Reliability

Validity and *reliability* determine the dependability of any research (Adams and Schvaneveldt 79–97; Burghardt 174–75; Crossen 22–24; Lang and Secic 154–55; Velotta 391). *Valid research* produces correct findings. A survey, for example, is valid when (1) it measures what you want it to measure, (2) it measures accurately and precisely, and (3) its findings can be generalized to the target population. A valid survey question helps each respondent to interpret the question exactly as the researcher intended, and it asks for information respondents are qualified to provide.

Survey validity depends largely on trustworthy responses. Even clear, precise, and neutral questions can produce mistaken, inaccurate, or dishonest answers for these reasons:

- People often see themselves as more informed, responsible, or competent than they really are. For example, in surveys of job skills, respondents consistently rank themselves in the top 25th percentile (Fisher, “Can I Stop” 206). Also, 90 percent of adults rate their driving ability as “above average.” And a mere 2 percent of high school students rate their leadership ability “below average,” while 25 percent rate themselves in the top 1 percent (Baumeister 21).
- Respondents might suppress information that reflects poorly on their behavior, attitudes, or will power when asked questions like: “How often do you take needless sick days?” “Would you lie to get ahead?” “How much TV do you watch?” When asked about going to church, for example, “respondents typically overstate their attendance by 70 percent” (“Sunday” 26).
- Respondents might exaggerate or invent facts or opinions that reveal a more admirable picture when asked questions like: “How much do you give to charity?” “How many books do you read?” “How often do you hug your children?”

- Even when respondents don't know, don't remember, or have no opinion, they tend to guess in ways designed to win the researcher's approval.

Reliable research produces findings that can be replicated. A survey is reliable when its results are consistent, for instance, when a respondent gives identical answers to the same survey given twice or to different versions of the same questions. A reliable survey question can be interpreted identically by all respondents. Factors that compromise survey reliability include the following:

- Each sample group has its own peculiarities (in distribution of age, gender, religious background, educational level, and so on).
- Various observers might interpret identical results differently.
- A single observer might interpret the same results differently at different times.

Much of your communication will be based on the findings of other researchers, so you will need to assess the validity and reliability of their research as well as your own.

Flaws in Study Design

While some types of studies are more reliable than others, each type has limitations (Cohn 106; Harris, Richard 170–72; Lang and Secic 8–9; Murphy 143).

EPIDEMIOLOGIC STUDIES. Observing various populations (human, animal, or plant), epidemiologists search for correlations (say, between computer use and cataracts). Conducted via observations, interviews, surveys, or review of records, these studies involve no controlled experiments. Their major limitations:

- Faulty sampling techniques (page 156) may distort results.
- Observation bias, or cognitive bias, (seeing what one wants to see) may occur.
- Coincidence can easily be mistaken for correlation.
- Confounding factors (other explanations) often affect results.

Even with a correlation that is 99 percent certain, an epidemiological study alone doesn't "prove" anything. (The larger the study, however, the more credible.)

LABORATORY STUDIES. Although laboratories offer controlled conditions, these studies also have certain limitations:

- The reaction of an isolated group of cells does not always predict the reaction of the entire organism.
- The reactions of experimental animals to a treatment or toxin often are not generalizable to humans. For example, massive, short-term doses given to animals differ vastly from the lower, long-term doses taken by humans.
- Faulty lab technique may distort results. For example, a recent study linking Vitamin C pills to genetic damage created panic. Experts have since concluded that "researchers themselves may have created 90 to 99% of the genetic damage when they ground up the cells to examine the DNA" ("Vitamin C" 1).

HUMAN EXPOSURE STUDIES (CLINICAL TRIALS). These studies compare one group of people receiving medication or treatment with an untreated group, the *control group*. Major limitations of human exposure studies:

- The study group may be non-representative or too different from the general population in overall health, age, or ethnic background. (For example, the fact that gingko might slow memory loss in sick people doesn't mean it will boost the memory of healthy people.)
- Anecdotal reports are unreliable. Respondents often invent answers to questions like: "How often do you eat ice cream?"
- Lack of objectivity may distort results. A drug's effectiveness is tested in a *randomized trial*, which compares a treated group of patients with a control group, who are given *placebos* (substitutes containing no actual medicine). These studies are *masked (or blind)*: Specific group assignment is concealed from patients. But doctors sometimes sneak their sickest patients into the treatment group hoping they might possibly benefit from new treatments. Despite good intentions, this practice subverts the random selection vital to such trials, making precise assessment of treatment impossible (Wallich 20+).

It is best to look for some consensus among a combination of studies.

Sources of Measurement Error

All measurements (of length, time, temperature, weight, population characteristics) are prone to error (Taylor 3–4).

The two basic types of measurement error are discussed below.

RANDOM ERROR. Anything that causes variations in values from measure to measure is a random error (Taylor 46, 94):

- Technique can vary from measure to measure: in running a stopwatch, reading markings on a scale, locating two points for measuring distance, changing body position in reading an instrument (say, a thermometer).
- Each observation differs with different observers or with the same observer in repeated observations.
- Each sample in a survey differs. Because of this sampling error, the same computation applied to multiple random samples drawn from the same given population produces varying results.

Researchers help neutralize random errors by repeating the measurements and averaging the range of values.

SYSTEMATIC ERROR. Any consistent bias that causes researchers to overestimate or underestimate a measurement's true value is a systematic error (Taylor 94, 97):

- A measuring device can be faulty: for example, a watch that runs slow or an improperly calibrated instrument.
- The measuring device (for example, a scale) might be positioned improperly.

Systematic errors are harder to neutralize than random errors because *all* systematically flawed measurements are inaccurate.

Sources of Deception

One problem in reviewing scientific findings is "getting the story straight." Deliberately or not, consumers are often given a distorted picture.

UNDERREPORTED HAZARDS. Although twice as many people in the United States are killed by medications than by auto accidents-—and countless others harmed-—doctors rarely report adverse drug reactions. One Rhode Island study showed some 26,000 such reactions noted in doctors' files, of which only 11 had been reported to the Food and Drug Administration (Freundlich 14).

THE UNTOUCHABLE RESEARCH TOPIC. Paranormal phenomena and extraterrestrials are rarely the topics of respectable research. Neither is alternative medicine—despite the fact that one U.S. citizen in three uses some type of alternative treatment (chiropractic, acupuncture, acupressure, herbal therapy) at least once a year. Little or no scientific evidence suggests that any or all of these therapies have actual medical benefits. Why? ''Partly because few 'respectable' scientists are willing to risk their reputations to do the testing required, and partly because few firms would be willing to pay for it if they were.'' Drug companies have little interest because ''herbal medicines, not being new inventions, cannot be patented'' (''Any Alternative?'' 83).

A "GOOD STORY" BUT BAD SCIENCE. Spectacular claims that are even remotely possible are more appealing than spectacular claims that have been disproven. This is why we hear plenty about claims like the following—but little about their refutation:

- Giant comet headed for earth!
- Tabletop device achieves cold fusion—producing more energy than it consumes!
- Insects may carry the AIDS virus!

Does all this mean we can't believe anything? Of course not. But we should be very selective about what we do choose to believe.

EXERCISES

1. Assume you are an assistant communications manager for a new organization that prepares research reports for decision makers worldwide. (A sample topic: "What effect has the North American Free Trade Agreement had on the U.S. computer industry?") These clients expect answers based on the best available evidence and reasoning.

   Although your recently hired coworkers are technical specialists, few have experience in the kind of wide-ranging research required by your clients. Training programs in the research process are being developed by your communications division but will not be ready for several weeks.

   Meanwhile, your manager directs you to prepare a one- or two-page memo that introduces employees to major procedural and reasoning errors that affect validity and reliability in the research process. Your manager wants this memo to be comprehensive but not vague.
2. From print or broadcast media, personal experience, or the Internet, identify an example of each of the following sources of distortion or of interpretive error:
   - a study with questionable sponsorship or motives
   - reliance on insufficient evidence
   - unbalanced presentation
   - deceptive framing of facts
   - overestimating the level of certainty
   - biased interpretation
   - rationalizing
   - unexamined assumptions
   - faulty causal reasoning
   - hasty generalization
   - overstated generalization
   - sanitized statistic
   - meaningless statistic
   - undefined average
   - distorted percentage figure
   - bogus ranking
   - fallible computer model
   - misinterpreted statistic
   - deceptive reporting

   Hint: For examples of faulty (as well as correct) statistical reasoning in the news, check out Dartmouth College's *Chance Project* at <www. dartmouth.edu/~chance>.

   Submit your examples to your instructor along with a memo explaining each error, and be prepared to discuss your material in class.
3. Referring to the list in Exercise 2, identify the specific distortion or interpretive error in these examples:

   a. *The federal government excludes from unemployment figures an estimated 5 million people who remain unemployed after one year* (Morgenson 54).

   b. *Only 38.268 percent of college graduates end up working in their specialty*.

   c. *Sixty-six percent of employees we hired this year are women and minorities, compared to the national average of 40 percent.* No mention is made of the fact that only three people have been hired this year, by a company that employs 300 (mostly white males).

   d. *Are you pro-life (or pro-choice)?*

4. Identify confounding factors (page 182) that might have been overlooked in the following interpretations and conclusions:

   a. *The overall cancer rate today is higher than in 1910* (“Are We” 4). Does this mean the actual incidence of disease has increased or are other explanations for this finding possible?
   b. *One out of every five patients admitted to Central Hospital dies* (Sowell 120). Does this mean that the hospital is bad?
   c. *In a recent survey, rates of emotional depression differed widely among different countries—far lower in Asian than in Western countries* (Horgan 24+). Are these differences due to culturally specific genetic factors, as many scientists might conclude? Or is this conclusion *confounded* by other variables?
   d. *“Among 20-year-olds in 1979, those who said that they smoked marijuana 11 to 50 times in the past year had an average IQ 15 percentile points higher than those who said they’d only smoked once”* (Sklaroff and Ash 85). Does this indicate that pot increases brain power or could it mean something else?
   e. *Teachers are mostly to blame for low test scores and poor discipline in public schools.* How is our assessment of this claim affected by the following information? *From age 2 to 17, children in the U.S. average 12,000 hours in school, and 15,000 to 18,000 hours watching TV* (“Wellness Facts” 1).

5. Uninformed opinions are usually based on assumptions we’ve never really examined. Examples of popular assumptions that are largely unexamined:

   - “Bottled water is safer and better for us than tap water.”
   - “Forest fires should always be prevented or suppressed immediately.”
   - “The fewer germs in their environment, the healthier the children.”
   - “The more soy we eat, the better.”

   Identify and examine one popular assumption for accuracy. For example, you might tackle the bottled water assumption by visiting the FDA Web site <www.FDA.gov> and the Sierra Club site <www.sierra.org>, for starters. (Unless you get stuck, work with an assumption not listed above.) Trace the sites and links you followed to get your information, and write up your findings in a memo to be shared with the class.

COLLABORATIVE PROJECTS

1. Exercises from the previous section may be done as collaborative projects.
2. Assess the findings below and then describe how you would rework the comparison to arrive at a meaningful conclusion.

   > A 1998 article used colorful graphics to underscore the “gap” between teachers’ wages and the “average wages” of other workers. State-by-state, teachers’ wages exceeded the “average wage” by a figure ranging from 2.9 percent (in the District of Columbia) to 65.2 percent (in Pennsylvania)—the gap in most states ranging from 20 to 60 percent. These figures were offered as evidence for the claim that teachers are indeed handsomely paid. (Brimelow 51)

SERVICE-LEARNING PROJECT

Divide into groups and identify a controversial environmental or technology issue (for example, the need to drill for oil and natural gas in the Alaskan Wildlife Refuge or the feasibility of the Star Wars Defense Initiative). Assume that your public-interest group publishes a monthly newsletter designed to give readers an accurate assessment of opposing claims about such issues. Using this chapter as a guide, review and evaluate the main arguments and counterarguments about the issue you've chosen. Prepare the text for a 1,500-word article that will appear in the newsletter, pointing out specific examples of questionable sources, interpretive error, or distorted reasoning.

FIGURE 10.1 Decisions in Evaluating and Interpreting Information Collecting information is often the easiest part of the research process. Your larger challenge is in getting the exact information you need, making sure it's accurate, and figuring out what it means. Throughout this process, there is much room for the types of errors covered in this chapter.

"How current is the information?"

"What is the source's reputation?"

"Can the source be trusted?"

"Who sponsored the study, and why?"

FIGURE 10.2 A Web Site That Advocates a Particular Viewpoint What concerns or issues are reflected in the content of this page? What public attitudes does this government agency appear to be advocating?

Source: Home page from Nuclear Energy Institute, <www.nei.org>. Reprinted by permission of Nuclear Energy Institute.

FIGURE 10.2 A Web Site That Advocates a Particular Viewpoint *(continued)*

"What are similar sources saying?"

Questions that invite distorted evidence

"Is there enough evidence?"

GUIDELINES for Evaluating Sources on the Web*

1. *Consider the site's domain type and sponsor.* In this typical address, <http://www.umassd.edu>, the site or domain information follows the *www*. The *.edu* signifies the type of organization from which the site originates. Standard domain types in the United States:

   .com = business/commercial organization
   .edu = educational institution
   .gov = government organization
   .mil = military organization
   .net = any group or individual with simple software and Internet access
   .org = nonprofit organization

   The domain type might signal a certain bias or agenda that could skew the data. For example, at a .com site, you might find accurate information, but also some type of sales pitch. At a .org site, you might find a political or ideological bias (say, The Heritage Foundation's conservative ideology versus the Brookings Institution's more liberal slant). A tilde (~) in the address usually signifies a personal home page. Knowing about a site's sponsor can help you evaluate the credibility of its postings.

2. *Identify the purpose of the page or message.* Decide whether the message is intended to merely relay information, to sell something, or to promote a particular ideology or agenda.

3. *Look beyond the style of a site.* Fancy graphics, video, and sound do not always translate into dependable information. Sometimes the most reliable material resides in the less attractive, text-only sites. People can design flashy-looking pages without necessarily knowing what they are talking about. Even something written in clear, plain English instead of in difficult scientific terms (as in medical information) might be inaccurate.

4. *Assess the site's material's currency.* An up-to-date site should indicate when the material was created or published, and when it was posted and updated.
5. *Assess the author's credentials.* Learn all you can about the author's reputation, level of expertise on this topic, and institutional affiliation (a university, a Fortune 500 company, a reputable environmental group). Do this by following links to other sites that mention the author, by using search engines to track the author's name, or by consulting a citation index (page 142). Newsgroup postings often contain "a *signature file* that includes the author's name, location, institutional or organizational affiliation, and often a quote that suggests something of the writer's personality, political leanings, or sense of humor" (Goubil-Gambrell 229–30).

*Guidelines adapted from Barnes; Busiel and Maeglin 39; Elliot; Fackelmann 397; Grassian; Hall 60–61; Hammett; Harris, Robert; Kapoun 4; Stemmer.

NOTE *Don't confuse an author (the person who wrote the material) with a Webmaster (the person who created and maintains the Web site).*

6. *Decide whether the assertions/claims make sense.* Based on what you know about the issue, decide where, on the spectrum of informed opinion and accepted theory, this author's position resides. How well is each assertion supported? Never accept any claim that seems extreme without verifying it through other sources, such as a professor, a librarian, or a specialist in the field. Also, consider whether your own biases might predispose you to automatically accept certain ideas.
7. *Compare the site with other sources.* Check related sites and publications to compare the quality of information and to discover what others might have said about this site, author, or topic. Comparing many similar sites helps you create a benchmark, a standard for evaluating any particular site based on the criteria in these guidelines). Ask a librarian for help.
8. *Look for other indicators of quality.*
   - *Worthwhile content:* The material is technically accurate. All sources of data presented as "factual" are fully documented. (See Appendix A.)
   - *Sensible organization:* The material is organized for the user's understanding, with a clear line of reasoning.
   - *Readable style:* The material is well written (clear, concise, easy to understand) and free of typos, misspellings, and other mechanical errors.
   - *Relatively objective coverage:* Debatable topics are addressed in a balanced and impartial way, with fair, accurate representation of opposing views. The tone is reasonable, with no "sounding off" or name-calling ("radicals," "extremists," "fringe groups," "fear mongers").
   - *Expertise:* The author refers to related theory and other work in the field and uses specialized terminology accurately and appropriately.
   - *Peer review:* The material has been evaluated and verified by related experts.
   - *Up-to-date links to reputable sites:* The site offers a gateway to related sites that meet quality criteria.
   - *Follow-up option:* The material includes a signature block or a link for contacting the author or organization for clarification or verification.

For more on evaluating Web-based sources, go to <www.webcredibility.org> and <www.library.cornell.edu/okuref/research/webeval.html>.

10.1
For more on evaluating questionable web sources visit
<www.ablongman.com/
lannonweb>
How evidence can be misused
“Is this claim too good to be true?”
“Is there a downside?”
Is the glass “half full” or “half empty”?
How framing of a survey question can affect responses
How framing can influence a life-and-death decision
“Is the evidence hard or soft?”
Source: © The New Yorker Collection 1998 Robert Mankoff from cartoonbank.com. All Rights Reserved.
10.2
What is truth? Can there be more than one truth? More at
<www.ablongman.com/
lannonweb>
A practical definition of “truth:
“Exactly how certain are we?”
Some changing notions about “Truth”
In science, uncertainty is a fact of life
Ethical
violations in communicating uncertainty
Questions that invite controversy
How underlying assumptions affect research validity
10.3
Where does bias come from? Learn more at
<www.ablongman.com/
lannonweb>
Personal bias is a fact of life
How bias can outweigh evidence
Rationalizing versus reasoning
“What else could this mean?”

CONSIDER THIS Standards of Proof Vary for Different Audiences and Cultural Settings

How much evidence is enough to “prove” a particular claim? The answer often depends on whether the inquiry occurs in the science lab, the courtroom, or the boardroom, as well as on the specific cultural setting:

- *The scientist demands evidence that indicates at least 95 percent certainty. A scientific finding must be evaluated and replicated by other experts. Good science looks at the entire picture. Findings are reviewed before they are reported. Inquiries and answers in science are never “final,” but open-ended and ongoing: What seems probable today may be shown improbable by tomorrow’s research.*
- *The juror demands evidence that indicates only 51 percent certainty (a “preponderance of the evidence”). Jurors are not scientists. Instead of the entire picture, jurors get only the information made available by lawyers and witnesses. A jury bases its opinion on evidence that exceeds “reasonable doubt” (Monastersky, “Courting” 249; Powell 32+). Based on such evidence, courts have to make decisions that are final.*
- *The corporate executive demands immediate (even if insufficient) evidence. In a global business climate of overnight develop-ments (in world markets, political strife, military conflicts, natural disasters), important business decisions are often made on the spur of the moment. On the basis of incomplete or unverified information—or even hunches—executives must make quick decisions to react to crises and capitalize on opportunities (Seglin 54).*
- *Specific cultures may have their own standards for authentic, reliable, and persuasive evidence. “For example, African cultures rely on storytelling for authenticity. Arabic persuasion is dependent on universally accepted truths. And Chinese value ancient authorities over recent empiricism” (Byrd and Reid 109).*

Questions for testing inferences
Factual observations
Invalid conclusions
"How much can we generalize from these findings?"
10.4
What are some consequences of relativism? Learn more at
<www.ablongman.com/
lannonweb>
"Did *X* possibly, probably, or definitely cause *Y*?"
Ignoring other causes
Ignoring other effects
Inventing a causal sequence
Confusing correlation with causation
Rationalizing
How numbers can mislead

CONSIDER THIS Correlation Does not Equal Causation

Causal reasoning often relies on statistical analysis. The following definitions help us evaluate a statistical analysis of causal relationships.

- *Correlation:* a numerical measure of the strength of the relationship between two variables—say, smoking and lung cancer risk, or education and income (Black 513).

  We depict the strength of a correlation by plotting data in *scatter diagrams* (Lang and Secic 94):

  As correlation increases, all data points approach the pattern of a perfect slanting line. In a *positive correlation,* as one variable (say, education) increases, so does the other (say, income). In a *negative correlation,* one variable (say, exercise) increases, while the other (say, weight) decreases.

  But instead of "proving" that one thing "causes" another, correlation merely signals a possible relationship.
- *Causation:* the demonstrable production of a specific effect (smoking causes lung cancer). Correlations between smoking and lung cancer or between education and income signal a causal relationship that has been demonstrated by many studies.

Causation Actually Demonstrated

Multiple studies (especially in labs) have demonstrated how specific causes produce specific effects: say, how asbestos fibers invade lung tissue to cause a cancer known as mesothelioma, or how ionizing radiation alters cellular DNA (Harris, Richard 172).

Causation Strongly Indicated

Consider this: "7 percent of Americans eat at McDonald's on any given day, and the average child watches 10,000 food commercials on television a year" (In Brief 28). Can we speculate about the relationship between these eating and viewing habits and the fact that the U.S. obesity rate has risen by one-third since 1980 ("Fat Chance" 3)?

Possible Causation

Between 1971 and 2001, forearm fractures among adolescents during sports doubled. One possible cause, according to Mayo Clinic researchers: the increasing consumption of calcium-free soft drinks instead of milk (Seppa 221).

Improbable Causation

A recently discovered correlation between moderate alcohol consumption and decreased heart disease risk offers no sufficient proof that moderate drinking "causes" less heart disease. Only detailed research might answer this question.

No Demonstrable Causation

A 2001 study found that people who win Academy Awards live an average of 4 years longer than people who don't (Herper 12). But is Oscar a "cause" of longer life? Of course not.

- *Coincidence:* the random but simultaneous occurrence of two or more events.
- *Confounding (or confusing) factors:* other reasons or explanations for a particular outcome. For instance, studies indicating that regular exercise

improves health might be overlooking this confounding factor: healthy people tend to exercise more than those who are unhealthy (“Walking” 3–4).

“Exactly how well are we doing?”

“How many rats was that?”

Needlessly precise and hard to read

Easier to decipher and compare

Three ways of reporting an “average”

“Why is everyone griping?”

Unethical uses of “averages”

What the “mean” doesn’t tell us

What the “median” and the ”mode” don’t tell us

“How much of a majority is 51 percent?”

“Which car should we buy?”

Bizarre correlations uncovered by data mining

Potential errors in meta-analysis

How assumptions influence a computer model

“Do we all agree on what these terms mean?”

“Is this news good, bad, or insignificant?”

10.5

What are some consequences of relativism? Learn more at
<www.ablongman.com/
lannonweb>

What makes a survey valid
Why survey responses can’t always be trusted
What makes a survey reliable
Why surveys aren’t always reliable
Common flaws in epidemiologic studies
Common flaws in laboratory studies
Common flaws in clinical trials
How scientific measurement can go wrong
Causes of random error in measurement
Causes of systematic error in measurement
“Has bad or embarrassing news been suppressed?”
“Is the topic ‘too weird’ for researchers?”

10.6

Why are fringe science and bad logic thriving? Learn more at
<www.ablongman.com/
lannonweb>

“Does this report get the story straight?”
Even bad science makes good news

GUIDELINES for Evaluating and Interpreting Information

Evaluate the Sources

1. *Check the source’s date of posting or publication.* Although the latest information is not always the best, it’s important to keep up with recent developments.
2. *Assess the reputation of each printed source.* Check the copyright page, for background on the publisher; the bibliography, for the quality and extent of research; and (if available) the author’s brief biography, for credentials.

3. *Assess the quality of each electronic source.* See page 170 for evaluating Internet sources. Don't expect comprehensive sources on any single database.
4. *Identify the study's sponsor.* If the study acclaims the crash-worthiness of the Batmobile, but is sponsored by the Batmobile Auto Company, be skeptical.
5. *Look for corroborating sources.* Usually, no single study produces dependable findings. Learn what other sources say, why they might agree or disagree with your source, and where most experts stand on this topic.

Evaluate the Evidence

1. *Decide whether the evidence is sufficient.* Evidence should surpass personal experience, anecdote, or news reports. It should be substantial enough for reasonable and informed observers to agree on its value, relevance, and accuracy.
2. *Look for a reasonable and balanced presentation of evidence.* Suspect any claims about "breakthroughs," or "miracle cures," as well as loaded words that invite emotional response or anything beyond accepted views on a topic. Expect a discussion of drawbacks as well as benefits.
3. *Do your best to verify the evidence.* Examine the facts that support the claims. Look for replication of findings. Go beyond the study to determine the direction in which the collective evidence seems to be leaning.

Interpret Your Findings

1. *Don't expect "certainty."* Most complex questions are open-ended and a mere accumulation of "facts" doesn't "prove" anything. Even so, the weight of solid evidence usually points toward some reasonable conclusion.
2. *Examine the underlying assumptions.* As opinions taken for granted, assumptions are easily mistaken for facts.
3. *Identify your personal biases.* Examine your own assumptions. Don't ignore evidence simply because it contradicts your way of seeing, and don't focus only on evidence that supports your assumptions.
4. *Consider alternate interpretations.* Consider what else this evidence might mean.

Check for Weak Spots

1. *Scrutinize all generalizations.* Decide whether the "facts" are indeed facts or assumptions, and whether the evidence supports the generalization. Suspect any general claim not limited by some qualifier ("often," "sometimes," "rarely," or the like).
2. *Treat causal claims skeptically.* Differentiate correlation from causation, as well as possible from probable or definite causes. Consider confounding factors (other explanations for the reported outcome).
3. *Look for statistical fallacies.* Determine where the numbers come from, and how they were collected and analyzed—information that legitimate researchers routinely provide. Note the margin of error.
4. *Consider the limits of computer analysis.* Data mining often produces intriguing but random correlations; meta-analysis might oversimplify

relationships among various studies; a computer model is only as accurate as the assumptions and data programmed into it.

5. *Look for misleading terminology.* Examine terms that beg for precise definition in their specific context: “survival rate,” “success rate,” “customer satisfaction,” “average increase,” “risk factor,” and so on.
6. *Interpret the reality behind the numbers.* Consider the possibility of alternative, more accurate, interpretations of these numbers.
7. *Consider the study’s possible limitations.* Small, brief studies are less reliable than large, extended ones; epidemiologic studies are less reliable than laboratory studies (which also carry flaws); animal or human exposure studies are often not generalizable to larger human populations; “masked” (or blind) studies are not always as objective as they seem; measurements are prone to error.
8. *Look for the whole story.* Consider whether bad news may be underreported; good news, exaggerated; bad science, camouflaged and sensationalized; or research on promising but unconventional topics (say, alternative energy sources) ignored.

CHECKLIST for the Research Process

(Use this checklist to assess your research process. Numbers in parentheses refer to the first page of discussion.)

Methods

- Did I ask the right questions? (120)
- Are the sources appropriately up-to-date? (165)
- Is each source reputable, trustworthy, relatively unbiased, and borne out by other, similar sources? (166)
- Does the evidence clearly support all of the conclusions? (124)
- Is a fair balance of viewpoints represented? (121)
- Can all of the evidence be verified? (173)
- Has the research achieved adequate depth? (122)
- Has the entire research process been valid and reliable? (188)

Reasoning

- Am I reasonably certain about the meaning of these findings? (174)
- Can I discern assumption from fact? (176)
- Am I reasoning instead of rationalizing? (177)
- Can I discern correlation from causation? (183)
- Is this the most reasonable conclusion (or merely the most convenient)? (174)
- Can I rule out other possible interpretations or conclusions? (177)
- Have I accounted for all sources of bias, including my own? (176)
- Are my generalizations warranted by the evidence? (178)

- Am I confident that my causal reasoning is accurate? (179)
- Can I rule out confounding factors? (182)
- Can all of the numbers, statistics, and interpretations be trusted? (181).
- Have I resolved (or at least acknowledged) any conflicts among my findings? (124)
- Can I rule out any possible error or distortion? (189)
- Am I getting the whole story, and getting it straight? (124)
- Should the evidence be reconsidered? (124)

Documentation

- Is the documentation consistent, complete, and correct? (Appendix A)
- Is all quoted material clearly marked throughout the text? (682)
- Are direct quotations used sparingly and appropriately? (682)
- Are all quotations accurate and integrated grammatically? (683)
- Are all paraphrases accurate and clear? (684)
- Have I documented all sources not considered common knowledge? (685)

For more exercises, visit
<www.ablongman.com/lannon>