

1-1-1999

# Using Systematic Thinking to Choose and Evaluate Evidence

Robert Griffin

Marquette University, [robert.griffin@marquette.edu](mailto:robert.griffin@marquette.edu)

---

Published version. "Using Systematic Thinking to Choose and Evaluate Evidence," in *Communicating Uncertainty*. Eds. Sharon M. Friedman, Sharon Dunwoody, and Carol L. Rogers. Mahwah, NJ: Taylor & Francis (Routledge), 1999: 225-248. [DOI](#). © 1999 Taylor & Francis (Routledge).

# Chapter 13

---

## Using Systematic Thinking to Choose and Evaluate Evidence

---

Robert J. Griffin

*Robert J. Griffin is the director of the Center for Mass Media Research at Marquette University, where he is a professor of journalism in the College of Communication and winner of the university's premier award for sustained teaching excellence. He has authored or coauthored articles and chapters on reporting about science, environment, and energy; he is lead editor of the book **Interpreting Public Issues**. His recent research focuses on risk communication and methods of teaching statistical reasoning to journalism students.*

*"God does not play dice with the universe."*

—Albert Einstein, c. 20th century

*"How can I be sure, in a world that's constantly changing?"*

—The Rascals, c. 20th century

*"Be sure of it; give me the ocular proof. . . . No hinge nor loop  
to hang a doubt on."*

—William Shakespeare (Othello III, 3), c. 16th century

Uncertainty will always plague us. You can bet on that. For a journalist, dealing with uncertainty is part of the job. Sometimes we feel uncertain because we sense that we lack enough knowledge of something we could know more about. Sometimes we feel uncertain because we cannot predict the seeming vagaries of real causal forces in the world as they influence card games, football games, volcanic eruptions, elections, human health, and

other aspects of our surroundings ranging from the microscopic to the cosmic (Kahneman & Tversky, 1982). In either case, a journalist's audience usually wants at least enough certainty in the information they are given so that they can deal with their world with some confidence (Eagly & Chaiken, 1993).

Of course, if the audience's world is in fact uncertain, a good journalist accurately points that out. For example, the results of preelection political polls and other sample surveys of the public are couched in sampling error. A journalist who reports such a survey accurately and responsibly will always take into account the survey's margin of error when interpreting the poll for the audience. Similarly, health risk estimates are often posed as a range of probabilities that indicate a person's chance of being victimized by a given injury or illness. For example, the local health department might proclaim that a person who is not vaccinated might have a 1 in 100 to 1 in 500 chance of catching the latest strain of flu.

Due to the nature of scientific inquiry, there is always some degree of uncertainty in scientific findings. In fact, the scientist lives in a world in which absolute proof is virtually impossible. Yet, based on the fruits of scientific inquiry, government officials often have to make decisions affecting public policies, including controversial actions such as limiting the emissions of greenhouse gases or requiring that additives be put into gasoline in certain cities to reduce urban air pollution. Judges and juries often have to determine, based on scientific evidence, whether plaintiffs have been harmed by hazards such as workplace carcinogens, defective products, or medical malpractice. At best, these decisions are made only after carefully weighing the bulk and quality of scientific findings that bear, pro and con, on the policy or judgment.

## CONFIDENCE GAMES

Even careful decisions can be corrupted by those who have a stake in misrepresenting scientific certainty. Basically, there are at least two kinds of practitioners of these scientific "confidence games": those who present scientific results as being more certain than the findings warrant and those who want to dismiss even strong scientific findings because the results are less than absolutely certain.

The former can include some scientists who have a professional stake in presenting findings that are noteworthy and some of the people who allege that they have been victimized by technology, products, or malpractice. The latter

can include some industry representatives whose job is to cast doubt on scientific revelations about deleterious effects of products, medical practices, or manufacturing processes. These two sides often collide in the courts, where judges have been given more and more discretion to act as what the U.S. Supreme Court terms gatekeepers of scientific evidence. For example, judges in product liability cases can exclude expert testimony if "there is simply too great an analytical gap between the data and the opinion offered," according to Chief Justice William Rehnquist (Biskupic, 1997, p. A2). As members of the fourth estate, journalists will continue to have a responsibility to help ensure that governmental and judicial decisions are balanced and based on valid scientific evidence and principles.

Of course, everyday people also have to make decisions based in part on scientific findings. These decisions can include which products to purchase, which scientific, technical, or environmental policies to support as citizens, and which changes in lifestyles or habits might affect their personal health and safety. Confidence hucksters can muddy those choices as well. So, it is also part of a reporter's job to help audiences sort empirical fact from junk science when people consider various consumer, political, and health-related options. In fact, people tend to rely a lot on the mass media as sources of information about health risks in particular (Freimuth, Edgar, & Hammond, 1987; Singer & Endreny, 1987).

One of the best ways for a journalist to discover whether scientific-sounding claims are valid is to subject them to the rigors of systematic thinking, that is, to the rules of evidence and reasoning scientists routinely apply to their investigations. Systematic thinking does not require journalists to be experts in research methodology and mathematics, although some basic knowledge of scientific procedures and statistics is certainly helpful (see, e.g., Cohn, 1989; Meyer, 1991). In fact, contrary to what many may believe, journalists are not inherently math dummies. Freshman college students going into journalism are just as adept at math as the average college freshman (Becker & Graf, 1994). "It is time to let the secret out," as Paulos (1995) wrote in his book, *A Mathematician Reads the Newspaper*. "Mathematics is not primarily a matter of plugging numbers into formulas and performing rote computations. It is a way of thinking and questioning that may be unfamiliar to many of us, but is available to almost all of us" (p. 3).

So, most journalists should be able to handle the kinds of reasoning processes required to think systematically about the news and, because science and statistics underlie many news stories, they have an increasing responsibility to their audiences to do so. However, one of the most common practices of the

journalistic craft—relying on anecdotal information in gathering and presenting the news—can markedly interfere with a reporter's attempts to apply systematic thinking.

## THE STORYTELLERS

Basically, the journalist's job is to tell a story. So it is quite natural for a journalist to think in terms of anecdotes to present the news. If there is a new treatment for diabetes, interview a few diabetics and find out how that will make life easier for them. If the state health department warns people about health dangers from eating fish that might contain mercury or PCBs, tap some anglers on the shoulder and ask them if they are worried. If a new study shows that TV viewing influences the academic performance of preteens, ask some local grade school teachers what they have noticed. If a federal report shows that urbanization is seriously encroaching on land used to grow food, write a story that follows a farm family through years of economic struggles.

The common wisdom among journalists is that anecdotes draw audience interest, humanize a news story and, because anecdotes tend to be vivid, make the news memorable. Unfortunately, relatively little is known about the actual effects these techniques have on audiences. What is known suggests that journalists should exercise some caution when employing anecdotal information in stories so as not to mislead audiences or themselves.

### Certain Examples, Uncertain Evidence

Anecdotes can be fine examples, but they are usually poor evidence, especially in the news. It is easy for a reporter to assume that the grassroots quotes she just gleaned from a dozen motorists at the local filling station were not just great copy for her story about gasoline taxes but typify a cross section of public opinion as well. To a social scientist, those same interviews are a convenience survey of an unrepresentative sample of 12. In other situations, a reporter might find an example of a prominent local athlete who is struggling to overcome drug dependence to illustrate his story on addiction in the city. To a social scientist, the athlete is strikingly atypical of the problems everyday people face.

Problems multiply when journalists present anecdotal information directly as evidence. For example, *Newsweek* ran an article headlined, "Conspiracy mania feeds our growing national paranoia" (Marin & Gegax, 1996–1997, p.

64). Claiming that “conspiracy paranoia is surrounding us,” the article takes a brief yet critical tour of some popular conspiracy theories. The news peg, as reflected in the headline, is that popular belief in conspiracies is growing. “This great nation has always had its share of conspiracy freaks. . . . But the ranks of the darkly deluded may be growing,” the authors stated (p. 66). “Clearly, something is heating up in the more tropical climes of the American psyche,” Marin and Gegax concluded, based primarily on the following evidence cited in the article:

- Three quarters of Americans believed that the government is somehow involved in conspiracy, according to a survey reported in *George* magazine.
- America Online had begun a channel for fans of the paranormal and the paranoid.
- Mel Gibson starred in a movie called *Conspiracy Theory*.
- The editor of *The Skeptical Inquirer*, a publication that debunks the farout, said that there certainly seems to be a resurgence in sympathy toward conspiracy theory and an increase in paranoia.

Most of the evidence is anecdotal and none of it would support the conclusion that, nationally, paranoia and belief in conspiracies are growing. That conclusion requires evidence that compares representative surveys of Americans at two points in time and asks about their beliefs in a variety of conspiracies. Such evidence might balance the anecdotal information presented in the story and relegate it to what it is: example instead of evidence.

Reliance on anecdotes might affect audiences in other ways as well. With our minds and our worlds filled with uncertainties and our days filled with only 24 hours, we often fall back on judgmental shortcuts, called *heuristics*, to make sense of things (Tversky & Kahneman, 1982). Heuristics are intuitive and can often bias our judgments. For example, people might overestimate the risks of cancer in the population if someone they know has the disease. Reading about a cancer case in the news will probably not have quite the effect on a person’s judgment as would firsthand knowledge. Nonetheless, vivid anecdotes, which are a news staple, could influence a person’s judgments of risks and should be employed carefully by journalists.

### SYSTEMATIC THINKING AND UNCERTAINTY

The use of anecdotes by journalists is certainly not going to disappear, but it is important to gather and present anecdotal information as examples rather than

evidence, to find typical examples instead of extreme cases to illustrate a story, and to couch description within a context that shows how representative or unrepresentative the cases may be. In most situations, this requires the systematic gathering of representative data of some type. However, just because information is in numerical form does not mean that it is necessarily any more representative than a typical verbal anecdote. In fact, undigested statistical information is often nothing more than a quantitative anecdote.

Although we will probably never overcome all of our uncertainty about the world, the techniques of systematic thinking employed by scientists serve to reduce the uncertainty that is brought about by faulty reasoning and improper evidence. The next sections illustrate three steps to systematic thinking that journalists can employ in their daily work.

### As Compared to What?

Journalists often encounter raw statistics, such as the number of people afflicted by heart disease or involved in automobile crashes annually. Some of these raw numbers can be quite astounding and equally misleading. For example, in 1992, *The New York Times* reported that four of America's largest cities—Los Angeles, San Diego, Dallas, and Phoenix—each tallied a record number of killings the previous year. As Arnold Barnett (1994) observed in *Technology Review*:

The implication was that even one all-time high among such cities was unusual, let alone four. The report failed to point out, however, that all four of these cities also reached new highs in population in 1991; thus, even if their per capita murder rates had not changed since Cain slew Abel, their absolute 1991 murder tolls would have set new records. (p. 44)

### First Step: Find the Baseline

Indeed, the first step in systematic thinking is to establish an appropriate baseline. A common means of establishing a baseline is to turn a raw statistic into a rate (e.g., 1 out of every 1,000) or percentage (e.g., .1%) by using an appropriate denominator. If *The New York Times* had used one of these techniques to establish a baseline, a different picture of urban murders might indeed have emerged. Thus, baselines provide essential interpretive context for any raw statistic.

Baselines also give context to an anecdote, especially by helping audiences see how representative the case is. For example, in a feature on the growing proportion of senior citizens who resort to suicide, *Milwaukee Journal Sentinel* reporter Fran Bauer (1996) started the story with this brief account:

Recently, an elderly man parked his car, walked to the top of the High Rise Bridge, stepped over the guard rails and jumped to his death.

Divers tried frantically to rescue the man from the icy waters of the Milwaukee River below. But he died within minutes, a suicide.

For most, the case was quickly forgotten.

Yet statistics tell a far grimmer story. (p. G1)

Bauer then gave readers context for the anecdote by including national data on the disproportionate upturn of suicides among older citizens, especially males, from 1980 to 1992 and showing that, although persons aged 65 and older accounted for only 13% of the population, they made up nearly 20% of all suicides. In doing so, the reporter made it clear that the proportion of suicides in the older age group is not just a simple reflection of their numbers in the population. The rest of the article discussed the factors such as isolation, depression, changing cultural attitudes, and even longevity itself that might contribute to suicides among the aging.

To help bring this message home, the paper accompanied the story with a graphic titled “1994 Suicides by Age Group in Milwaukee County” (Fig. 13.1). Unfortunately, however, that graphic confused the picture the reporter had so carefully presented in text. At first glance, the pie chart seems to show that suicides are decimating the young but are rare among the old. What caused this pie chart to go sour? Instead of depicting suicides in each age group on a per capita basis, as would be appropriate for the story, the chart shows the portion of the total number of suicides that occurred in each age group. In short, the graphic confuses the meaning of the story because it ignores the baseline, specifically, the size of each age group in the local population. To better illustrate Bauer’s trend story, the newspaper might have used a couple of charts—one portraying per capita suicides in each age group in 1994 and another showing the same breakdown for 1980. That way, two essential baselines are used: the population base and the prevalence of suicides at a comparative time in the past.

### **Second Step: Make a Dynamic Comparison**

Despite its flawed graphic, the suicide article showed how comparing rates across groups and across time can reveal dynamic patterns that are otherwise



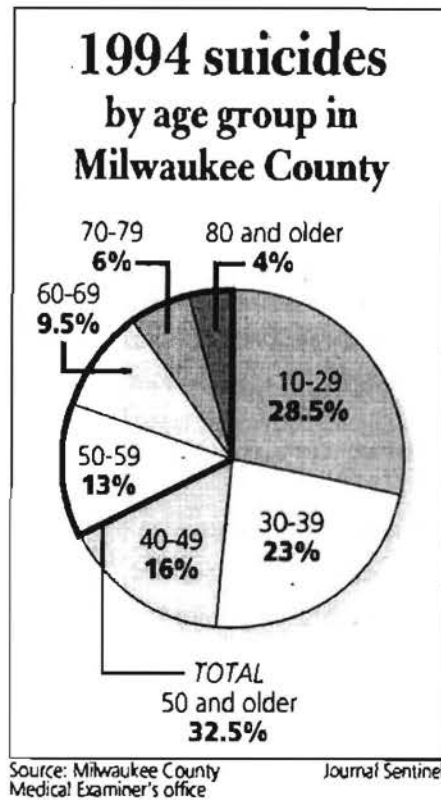


FIG. 13.1. 1994 Suicides by age group in Milwaukee County.

obscured when reporters employ only simple descriptions in the form of anecdotes, raw statistics, or even raw rates and percentages (e.g., reporting only the overall suicide rate for the population in general). A phenomenon such as illness or suicide now becomes a variable to be compared with another variable, such as age or sex differences, giving clues to sometimes subtle forces at work in society and in human lives. Of course, as Cohn (1989) noted, these comparisons can take many forms and must be rigorous and fair. However, looking for these dynamic comparisons is the second step in systematic thinking. By habitually asking "as compared to what?" and searching carefully for solid evidence of the answer, journalists can get new insights into the news and avoid some flimsy or misleading conclusions.

For example, *Discover* magazine once reported that 90% of the people who survived airplane crashes had formed in their minds a plan of escape before the accident happened (Nolan, 1986). Their recommendation? Look for the emergency exits and plan how to get off before you take off. This advice seems sensible enough, but it is not as factually based as it would seem. Note that *Discover's* advice is based on a comparison that is implied but for which there is no evidence.

As Gilovich (1997) observed, it is impossible to find out what percentage of the nonsurvivors had also formulated escape plans. In short, there is no way of knowing whether those who planned their way out actually fared any better.

In another case, an Associated Press (1995) article about the success of cardiopulmonary resuscitation (CPR) appeared under the headline: "Bystanders' CPR Efforts Often Backfire, Study Says." The lead paragraph read: "Chicago—Bystanders who attempted CPR on cardiac arrest victims got it wrong more than half the time, reducing patients' already slim chances of survival, a study found" (p. 6A).

Notice that the clear implication of the newspaper headline and of the first paragraph is that people who try to do CPR on a heart attack victim, but who do it improperly, are doing more harm to the victim than if they had done nothing at all. That is pretty important advice, advice with ethical and legal implications, as well as with implications for the life of the poor victim.

The news item was based on an article by Gallagher, Lombardi, and Gennis (1995) that appeared in the *Journal of the American Medical Association (JAMA)*. The brief Associated Press news story quotes one of the article's authors, John Gallagher of the Albert Einstein College of Medicine in New York, as stating that improperly administered CPR "does not seem to be any better than no CPR." He did not say that improperly administered CPR is worse than no CPR, which should have tipped off the reporter that something was amiss in the lead paragraph.

An abstract of the *JAMA* article, then readily available to journalists with Internet access, briefly explained the research method and basic findings of the study. Emergency hospital personnel, who arrived at the scene of a cardiac arrest, recorded whether any bystanders had attempted CPR on the victim and, if so, whether the technique they used was *effective*, that is, whether it was performed according to medical guidelines. The patient *survived* if he or she was able to return home from the hospital. The researchers also controlled for some other factors that might affect the outcome of the study. The abstract explained that:

the survival statistic for those receiving CPR was 19 out of 662 compared to 11 out of 1405 who did not receive CPR. . . . Of those patients who received effective bystander CPR, 14 out of 305 survived (4.6%) compared with 5 out of 357 (1.4%) who received CPR judged to be ineffective.

A careful reading of the numbers shows that only 0.8% (11/1,405) of the victims who received no CPR had survived. In situations like this, a simple but systematic jotting down of the numbers on a notepad, such as a reporter might have done in Fig. 13.2, can help journalists—and their audiences—understand

<i>No CPR</i>	<i>Ineffective CPR</i>	<i>Effective CPR</i>
<i>0.8% survived</i>	<i>1.4% survived</i>	<i>4.6% survived</i>

FIG. 13.2. *JAMA* number comparisons.

what comparisons are being made. Although giving ineffective CPR might not really improve the victim's odds of survival as compared to administering no CPR at all, it certainly does not "backfire" by "reducing patients' already slim chances of survival," as the story headline and lead had erroneously reported.

Perhaps the statistics most in need of dynamic comparisons and reportorial finesse are vital statistics data representing health and disease, life and death. "They are much applied, misused, and misunderstood," stated Cohn (1989, p. 74). "Yet these statistics can yield fascinating stories if we learn something of their power and limits and the rather special vocabulary of human lives." Cautioning that disease data are often applied too broadly to the population, Crossen (1996) observed:

One in five American men will get prostate cancer. One in eight American women will get breast cancer. At least two million Americans are manic-depressive, and more than two million are schizophrenic. At least 60 million Americans have high blood pressure, 12 million have asthma and four million have Alzheimer's disease. One in three Americans is obese.

With numbers like these, it is amazing there is anyone still here—let alone people living happy, healthy lives. Projections of the incidence of disease are rampant these days, as a growing number of health advocacy groups compete for people's limited attention and money. Most of the numbers are extrapolations or estimates—at best. Yet as the media report them, often uncritically and without context, these conjectures assume the mantle of quantifiable fact. (p. B1)

Crossen suggested that a better way to present information to people about risks of a disease is to project, for example, how risks vary by gender and age. Figure 13.3 illustrates data her story provided on the probability of developing cancer in 10, 20, and 30 years, and eventually for men and women of various ages who are currently free of cancer. The chart shows some very interesting patterns. For women, for example, the overall risk of getting cancer sometime in life actually decreases as they get older. The same is not true for men. Of course, heart and circulatory problems claim more human lives than does

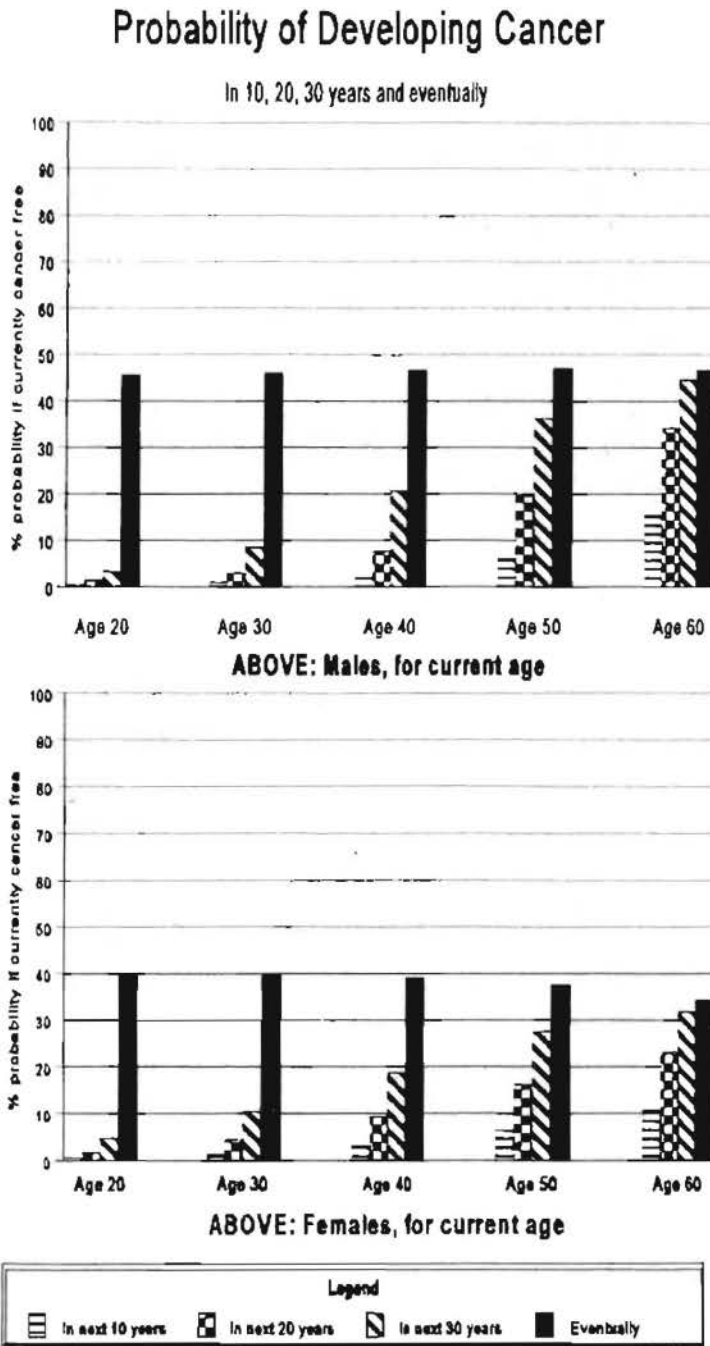


FIG. 13.3. Based on data from *The Wall Street Journal*.

cancer, Crossen said, and people should remember that, despite increasing longevity, no one lives forever.

Figure 13.3 is an example of what statisticians call a *multivariate analysis*. Tables such as these are very useful because they illustrate the ways different

factors (here, age and sex) might combine in different ways to produce different outcomes (here, the risk of cancer). Notice that this is a much more dynamic and realistic picture of the way life really is than one gets from the more crude descriptions provided by simple anecdotes and undigested statistics. However, because data such as these can become quite complicated for people to understand, journalists will need to become more adept at reasoning from data and interpreting data for audiences. Innovative and clear graphic displays of statistical information (e.g., Tufte, 1983, 1997; Utts, 1996; Wainer, 1997) are essential to that task. In particular, computers and other new technologies offer exciting opportunities to use hypertext (Fredin, 1997), animation, and interactive environments to help people understand data dynamics.

## Error and Uncertainty

Journalists also should remember that even the most carefully gathered statistical information contains some uncertainty in the form of error. Some of the error rests in the techniques and measures used to gather the information, such as relying on the completeness of health department records, the precision of a medical test, or a person's ability or willingness to report illnesses, socially undesirable activities, or highly personal information to a survey interviewer. Some of the error rests in extrapolating or generalizing from, for example, animal tests to humans or from a sample to the population. So, in making dynamic comparisons, it is important to take into account what the range of error might be. As a general rule, we should be cautious about small differences between groups, especially if the differences are only a few percentage points. They might simply be the result of error and, therefore, have no real meaning. If the data are from a well-designed probability sample survey (of people in a city, for example) or laboratory experiment in which subjects were randomly assigned to conditions (a placebo versus a new drug, for example), then a reporter can more readily get a handle on how much room to allow for error by relying on the reported statistical significance of the results or by applying the standard formula to determine the margin of error in surveys. Of course, even these statistical tests will not compensate for errors in the measures and techniques themselves, so it is still wise to exercise some caution.

So, a good revision of the question posed at the beginning of this section is this: "As compared to what, given error?"

**SIDEBAR:**  
EXAMPLES AND SAMPLES, OR  
HOW A LITTLE HOMEWORK CAN GO A LONG WAY

Parents from the Milwaukee suburb of Shorewood had complained that their seventh-grade offspring were being given an unusual amount of homework, so much so that their backpacks were becoming laden with books and too heavy for a kid to carry.

So, to check out these parental allegations, reporters decided to compare the heaviness of the backpacks of Shorewood Intermediate School students to the weight of backpacks carried by their peers at two schools in neighboring communities. In a large, color graphic that accompanied the front-page article ("Full Load," 1997), the paper listed the weight of each of the 14 backpacks sampled at each of the three schools along with the average backpack weight at each school: 20.43 pounds at Shorewood, 16.93 pounds at University School, and 11.93 pounds at Morse Middle School. The data are reproduced in Figure 13.4. The reporters concluded that "for the most part, the [Shorewood] parents were right." To further illustrate the point, the newspaper used an example as part of the graphic: a photo of a student from the Shorewood school whose backpack was heavier (at 31 pounds) than any of the other backpacks from any of the schools. "What's it like carrying around 31 pounds?" asked a tag line above the student's photo. The answer, also illustrated, is that 31 pounds is equal to the weight of nearly two bowling balls or 191 toy "beanie babies" and is greater than the weight of a Trek 750 bicycle. The story itself is peppered with quotes from students at the three schools about what it is like to lug their burdens.

All in all, the reporters endeavored to conduct a dynamic comparison across three schools and tried to use comparative data on the average weight of backpacks at each institution to provide a moderating context for the quotations from students. Without the weigh-ins, the entire story might have had to rely on the use of anecdotes as evidence instead of as examples. The interesting graphic tried to translate the weight of backpacks into terms many readers might understand.

Unfortunately, two enduring journalistic problems—using unrepresentative examples and unrepresentative samples—seemed to mar this otherwise laudable effort:

1. The illustration used the decidedly atypical backpack—the one that weighed more than any of the others and half again as much as the average for Shorewood Intermediate School—rather than a typical backpack (the mean,

*(continued)*

## Sidebar (continued)

## It's tough being a backer packer

It all began when some Shorewood parents complained how heavy their kids' backpacks have become because of the amount of homework brought home. Do their concerns carry weight? We found out, by weighing backpacks from 14 students from each of the schools below. The results? Well, for the most part, the parents were right. And how do the kids feel about carrying all that weight? Some shrugged it off, but others said it feels like the weight of the world.

Shorewood Intermediate School		University School		Morse Middle School	
Sampled backpack weights	TOTAL WEIGHT: 286 LBS.	Sampled backpack weights	TOTAL WEIGHT: 237 LBS.	Sampled backpack weights	TOTAL WEIGHT: 167 LBS.
11 20 25 15 20 25 15 20 30 15 20 31 18 21	AVERAGE WEIGHT: <b>20.43</b> pounds	6 17 19 11 18 20 13 18 21 16 19 24 16 19	AVERAGE WEIGHT: <b>16.93</b> pounds	8 11 13 9 12 14 11 12 14 11 13 15 11 13	AVERAGE WEIGHT: <b>11.93</b> pounds
What's it like carrying around 31 lbs.?					

FIG. 13.4.

LOUIS K. SALDIVAR/Journal Sentinel

median and mode are 20 pounds) to represent that school. Journalists too often use the atypical as an example, and in this case the reader can too easily be confused about the average weight of backpacks at that school.

2. The reporters used what they described in the text of the story as a "very unscientific survey" (p. 1A) to choose which backpacks to weigh. Although they did not describe how they selected backpacks, "unscientific" usually means "nonrandom sample" and opens the possibility that expectations about finding heavier backpacks at the Shorewood school might have unconsciously affected reporters' choices. A random sample of some type would probably not have been too hard to conduct, especially given the effort reporters had already devoted to the story. And it would have given the reporters a big advantage: they would have had a means to control for sampling error. As it is, there is no way to reduce the large amount of uncertainty in their results. In short, it is not clear what they actually found. In fact, had the same data been the result of a random sample, the results would show that the average backpack weight at Shorewood Intermediate School is not different, beyond sampling error, from the average backpack weights at University School, according to a statistical

(continued)

*Sidebar (continued)*

test called the “analysis of variance” commonly available on desktop computer statistical packages. The real finding would be that Morse Middle School backpacks are the lightest and that, for the most part, the Shorewood parents were wrong.

## WHY AND SO WHAT?

Interpretation is an essential component of good reporting and usually requires answering the questions “why?” and “so what?” The answer to the question “why?”—for example, “Why does urban sprawl affect water quality?” or “Why would eating fatty foods increase a person’s risk of heart attack?”—calls on the reporter to address cause-and-effect relations. Similarly, the answer to the question “so, what?”—for example, “So, what policy solutions can mitigate the effects of sprawl on water quality?” or “So, what personal actions can one take to lower the risk of a heart attack?”—also requires analysis of causality.

Because people tend to base their preferences for solutions to a problem on their perceptions of what caused the problem (Didier, 1987), it is important for reporters to present causal information carefully to audiences. Doing so means being especially alert to some of the common mistakes we all make in everyday causal reasoning. Many of these miscues stem from the need to quickly overcome uncertainties in our world in order to go about everyday life pressures that certainly affect reporters at least as much as everyone else. For a reporter, however, taking a more careful and responsible approach to interpreting causality requires taking a more systematic look at what causes what, adopting in less formal ways the standards of proof scientists require.

### Third Step: Use Causal Caution

Dynamic comparisons reveal what are commonly known as *correlations* between variables. In the sciences, establishing a correlation between two variables is an essential step in determining whether one may be a cause of the other. Suppose there were no correlation between being a cigarette smoker and having a higher risk of lung cancer. Under those conditions, smoking could not be a cause of lung cancer. However, the opposite is not true: Simply showing an association between smoking and cancer did not prove that smoking contributed to the risk of lung cancer. Additional rigorous evidence was needed,



as is true with any area of scientific inquiry into causality. Along with showing correlation, evidence of causality requires evidence 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.

“One of the first things they teach in introductory statistics is that correlation is not causation,” quipped Sowell (1996, p. 11) in *The Country Chronicle*. “It is also one of the first things forgotten.” Remembering these rigorous standards of causal proof and adopting the caution they impart constitute the third step in systematic thinking.

**Correlation and Causal Conclusions.** In everyday life we often jump to conclusions about causality. Sometimes those judgments are based on what is, at best, incomplete or flimsy evidence of correlation, much of it anecdotal and based on what we happen to have noticed or experienced. Unlike the scientist, who relies on systematic sampling and statistical techniques, the layperson “must rely upon intuitions and subjective impressions based on limited access to relevant data,” explained Ross and Anderson (1982, p. 140). The result is often a biased base on which to build causal conclusions.

Some of those premature causal conclusions find their way, unexamined, into the media, as illustrated by the story about the effects on survival if passengers mentally map their emergency routes out of an aircraft. Some are even embarrassing for the media. Take, for example, the media’s response to information proffered by various advocates for battered women including, ironically, the group Fairness and Accuracy in Reporting (FAIR) that reports of domestic violence rise 40% on Super Bowl Sunday. In short, the causal implication is that watching the Super Bowl makes men more likely to batter their female partners. As Hohler (1993) of *The Boston Globe* related:

The image was alarming. Men across America, incited by booze, gambling losses and the body-slammings exploits of their football heroes, could make Super Bowl Sunday the worst day of the year for domestic violence.

Activists trumpeted the warning, saying national studies supported their claim. Much of the nation’s media echoed the alarm.

And NBC, heeding the prediction, aired as its only public service announcement in the countdown to the Super Bowl a 20-second television spot that dramatized for 40 million viewers the horror of domestic violence.

But in an embarrassing setback for the campaign against domestic violence—and for the news media—some of the groups that pressured NBC to air the free spot,

including Fairness and Accuracy in Reporting, acknowledged yesterday that they had based their predictions in part on incomplete, inaccurate or anecdotal information. (p. 1)

According to Hohler, the media watchdog group FAIR had extrapolated the 40% figure, which FAIR described as anecdotal, from a book of photo essays on domestic violence. There were other errors in advocates' claims as well, but there did not seem to be any significant, systematic evidence of an increase in domestic violence after the Super Bowl game.

"People are extraordinarily good at ad hoc explanation," Gilovich (1991, pp. 21-22) observed. "To live, it seems, is to explain, to justify, and to find coherence among diverse outcomes, characteristics, and causes." Gilovich noted that people can find patterns even in random phenomena—randomness being the ultimate in uncertainty and lack of correlation—and quickly explain the patterns in terms of their own preexisting theories and beliefs about causality.

People can have a number of intuitive, preexisting theories about causality (see, e.g., Hilton, 1988; Kahneman, Slovic, & Tversky, 1982). Many concern human behavior—the stuff that much reporting is made of. These intuitive beliefs can bias our perceptions of causality, especially when causes and effects are otherwise uncertain. For example, people tend to overestimate the role of forces inside the individual, such as personality, ability, disposition, and motivation, as causes of human behavior and to underestimate the role of environmental or situational factors, such as the varied opportunities and obstacles that exist for people in different social classes. Heider (1958) called this bias the *fundamental attribution error*, and it affects us more when we interpret the behavior of others, as reporters tend to do, rather than our own. When applied to whole classes of people, the attribution of behavior to shared internal states can form the basis for social stereotypes (Hamilton, 1979), such as we might have of others of a different race or sex when we believe that whole groups of people are inherently lazy, ignorant, insensitive, and so forth. Media portrayals might be associated with the ways audiences attribute to internal or external causes the way members of certain groups in society behave (Griffin & Sen, 1995). Attributional biases, of course, could have a lot of ramifications if they influence media content, audience perceptions, and social policy.

Although it would be nearly impossible for reporters to effectively counteract all sources of error when making statements about "why" and "so what," adopting a more rigorous set of standards for causal proof will reduce the likelihood of inaccurate causal conclusions. To that end, here is a brief overview of the other steps scientists take in finding evidence of causality.

**The Right Time Slot.** People develop causal beliefs because they repeatedly witness the association between an event and something that follows it (Hilton, 1988). Scientists use the same approach, although more rigorously. In essence, a purported cause must be shown to precede its effect in real time, whether it be epochs or nanoseconds. Many scientific procedures, such as most controlled laboratory experiments or panel design surveys, directly observe before-and-after changes as they happen. In many other cases, such as in geology or deep space astronomy, the evidence of before-and-after is often gathered indirectly. Sometimes, attempts are made to discern causal sequences from only one point in time, such as in a cross-sectional epidemiological or public opinion survey. In all cases, however, it is wise to be cautious about the evidence of time sequence. It is often helpful for a reporter to determine whether the proposed order of the cause and effect variables could just as realistically be reversed.

**The Network.** Most of the things that happen are the result of a variety of factors, many interwoven with one another. One of the most difficult tasks in systematic thinking and investigation is to separate the influence of an apparent causal agent from other variables, often called *confounds*, which also might affect the outcome attributed to the causal agent.

Cohn (1989), for example, related the story of a scientist who proposed the possibility that left-handed batters were overrepresented among the best hitters in baseball because of hemispheric lateralization of the brain. A more baseball-savvy critic had a simpler explanation: Left-handed batters happen to enjoy a natural physical advantage in the game, specifically, most pitchers are right-handed, which gives left-handed batters an edge, and left-handed batters are already moving toward first base after they swing, making it easier for them to reach the base.

Usually the simpler, more parsimonious explanation is preferred as long as it predicts the phenomenon—in this case, the better hitting performance of lefties in baseball—at least as well as the more complicated one. Of course, it is also quite possible that both the scientist and the baseball-wise observer were right, because a phenomenon can (and usually does) have multiple causes. Thus, it is wise for a reporter to ask what other causal agents might be on the scene and how they have been accounted for.

Here are some of the other common patterns of relationships among causal agents that a reporter might encounter. For convenience, the suspected causal agent will be referred to by its common nickname, *X*, the variable it apparently influences as *Y*, and the other dynamic variable in the mix as *Z*. *X* is also often

referred to as an *independent variable*, *Y* as a *dependent variable*, and *Z* as a *third variable* (no matter how many there are).

**Contingency.** Sometimes a variable (*Z*) can work like a switch or catalyst for the relationship between *X* and *Y*. Only if *Z* takes on certain characteristics, for example, does *X* affect *Y*. In Fig. 13.3, sex might be considered to be *Z*, age as *X*, and lifetime cancer risk as *Y*. If *Z* (sex) is female, then advancing age (*X*) decreases overall cancer risk (*Y*). If *Z* (sex) is male, then age and overall cancer risk are unrelated.

**Intervening Variable.** For a variable to effect change in another variable, the two must be functionally related. That is, the link between the variables should be clear and the processes by which *X* affects *Y* well defined. Sometimes *X* can affect *Y* only through an intervening variable *Z* that is the more proximate cause of changes in *Y*. Thus, intervention is like a chain of relationships. For example, researchers often look for, and find, differences in all sorts of social, psychological, and even health-related variables based on demographic differences in the population. Flynn, Slovic, and Mertz (1994) found that white men perceive risks differently than white women and minority men and women. Yet, to explain why this difference occurs, the authors suggested that white males may simply feel more in control of their environment than everyone else, and that sense of control affects their perceptions of being at risk. Thus, a sense of control is posed as an intervening variable (*Z*). In general, demographic variables often need such assistance. The alert reporter, trying to assess whether the services of an intervening variable are needed, might ask whether it is crystal clear *why* *X* might affect *Y*. Take another look at Fig. 13.3 in that light.

**Lurking Variable.** This one's a real con artist. In what is sometimes called a *spurious relationship*, the lurking variable *Z* deceives you into thinking that *X* and *Y* are related when they really are not. In reality, the lurking variable itself affects both *X* and *Y*, making them correlate with one another without any real connection between them. For example, suppose a study were to show that Internet users who browse on-line news services know more about international current events than other Internet users. It might indeed be tempting to say that their greater use of net news (*X*) is making these folks more savvy about what is happening in the world (*Y*). Might a *Z* be lurking about? Perhaps these folks are better educated than those who do not use the on-line news services. They might visit the on-line news sites as part of a pattern of greater attention to a lot of news channels, including television, newspapers, and news

magazines. Their superior knowledge of international current events might be a byproduct of their educational preparation coupled with their use of these more traditional news media. A good study will control for alternative possibilities such as this.

In general, the authors of most scientific studies are very cautious about claiming causal relationships, preferring instead to claim not much more than association. Nonetheless, systematic thinking on the part of a journalist can serve as a check on unwarranted causal claims.

### SUMMARY: PREPARING FOR THE 21ST CENTURY

*"So certain are you."*

—Yoda, *Star Wars*, Century uncertain

Computers, the Internet, and the other new communication technologies are changing the information landscape. For journalists and the public alike, landmarks that used to identify trustworthy sources and valid information are disappearing. For example, rumors can attain about the same status as news on the Internet, to the point that even veteran journalists have been befuddled. Whether new landmarks appear in this world of uncertain information is anyone's guess. If these trends continue, the new millennium will require people to have cognitive tools, or at least considerable guidance, to verify what is valid and what is not. More than ever, journalists will need to apply the tools of systematic thinking, tools that, by their grounding in the sciences, are effective in sorting out facts from fantasy, puffery, politics, and even the preening of scientists.

Fortunately, there are many fine sources that journalists can use to prepare themselves. Philip Meyer's now classic works, *Precision Journalism* (1979) and *The New Precision Journalism* (1991), continue to offer reporters superb guidance in the use of surveys, sampling, statistics, and related tools of the social sciences, and—just as valuable—guidance in the thinking that goes along with using them. Victor Cohn (1989) similarly provided an excellent guide to reporting a wide range of scientific controversies, methods, and thinking in *News & Numbers*, a book that includes a fine chapter on scientific uncertainty. Even though audiences might not generally be familiar with the role of uncertainty in science, clear reporting of uncertainty may indeed help people understand it better (see Johnson & Slovic, 1995).

Unfortunately, the mass media have provided critical observers with a wealth of examples of how not to report and display scientific and statistical information. Journalists can learn a lot about how to do things better—and learn a lot about systematic thinking in the bargain—by reading books such as Paulos' *A Mathematician Reads the Newspaper* (1995), his best-seller *Innumeracy: Mathematical Illiteracy and its Consequences* (1988), and Gilovich's *How We Know What Isn't So* (1991). Journalists also should take advantage of the movement within the teaching of statistics and mathematics that urges universities to ensure that college graduates are quantitatively literate, that is, that they can apply mathematical thinking, beyond mere formulas, to everyday problems (Subcommittee on Quantitative Literacy Requirements, 1995). Various universities are responding to this need by offering courses incorporating or devoted to quantitative reasoning. The Chance Project at Dartmouth College in Hanover, New Hampshire, has a wealth of examples of correct and incorrect quantitative reasoning in the news and also has a storehouse of related teaching materials, all available on-line ([www.dartmouth.edu/~chance](http://www.dartmouth.edu/~chance)). One superb how-to book that employs the systematic thinking approach to statistics and that definitely belongs on the journalist's bookshelf is Jessica Utts' very readable *Seeing Through Statistics* (1996).

The growth of computer-assisted reporting will require journalists to think adeptly and systematically about the numbers they encounter. In fact, journalists without the requisite computer and quantitative skills will probably find themselves at a real disadvantage in the journalism job market (Feola, 1995). The growth of new media, which overwhelmingly use visual displays to present information, will demand that reporters and illustrators become masterful at presenting quantitative information graphically and accurately. Computerized animation and interactive formats could help audiences understand dynamic comparisons and even the influence of third variables. For example, one's chances of contracting heart disease at a given age might be illustrated by a curve on an attractive graph that changes as the viewer provides the computer program with different information about his or her own unique background and habits. The curve might be animated to show change based on lifestyle alterations, such as stopping smoking or adopting a low-fat diet, that the viewer might be considering—and a range of uncertainty could be illustrated around the curve. Of course, to date, media have sometimes distorted quantitative information when presenting it graphically. Some fine books that show how to prepare valid, attractive, still-life graphic displays of data include Utt's *Seeing Through Statistics* (1996), Tufte's *Visual Explanations* (1997), and especially Tankard's "Visual Crosstabs" (1994) and Wainer's *Visual Revelations* (1997).

Journalists in the 21st century must be able to reason from verbal and quantitative information, know how to assess what information is missing, be able to gather and validate the required information, be adept at understanding and explaining the uncertainty that scientific information inevitably contains, and be able to interpret information to nonexpert audiences verbally, quantitatively, graphically and, most of all, accurately. These cognitive and communication skills are absolutely essential if journalists are to meet their responsibilities to society and their audiences in the new millennium.

## ACKNOWLEDGMENTS

Special thanks to Dr. Laurie Snell, professor of mathematics and Director of the Chance Project at Dartmouth College in Hanover, New Hampshire, for helping to provide many of the examples used in this chapter, and to Ruth Boulet, master of arts graduate in journalism from Marquette University, for her superb help with the research.

## REFERENCES

- Associated Press (1995, December 27). Bystanders' CPR efforts often backfire, study says. *Milwaukee Journal Sentinel*, p. 6A.
- Barnett, A. (1994, October). How numbers can trick you. *Technology Review*, 39-45.
- Bauer, F. (1996, April 1). Senior suicides on rise. *Milwaukee Journal Sentinel*, pp. G1, G3.
- Becker, L. B., & Graf, J. D. (1994). *Myths & trends: What the real numbers say about journalism education*. Arlington, VA: The Freedom Forum.
- Biskupic, J. (1997, December 16). Trial judges have wide discretion on scientific testimony, Court says. *The Washington Post*, p. A2.
- Cohn, V. (1989). *News and numbers*. Ames, IA: Iowa State University Press.
- Crossen, C. (1996, April 11). Fright by the numbers: Alarming disease data are frequently flawed. *The Wall Street Journal*, p. B1, B4.
- Didier, M. (1987). Attributions in the evaluation of alternative solutions to social problems. *Journal of Social Psychology*, 127, 357-365.
- Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Fort Worth, TX: Harcourt Brace.
- Feola, C. J. (1995, March). Making the cut: Enhance your job opportunities by developing your computer skills. *The Quill*, 24-26.
- Flynn, J., Slovic, P., & Mertz, C. K. (1994). Gender, race, and perception of environmental health risks. *Risk Analysis*, 14, 1101-1108.
- Fredin, E. S. (1997, September). Rethinking the news story for the Internet: Hyperstory prototypes and a model of the user. *Journalism and Mass Communication Monographs*, 163.

- Freimuth, V. S., Edgar, T., & Hammond, S. L. (1987). College students' awareness and interpretation of the AIDS risk. *Science, Technology, & Human Values*, 12, 37–40.
- Full Load. (1997, November 11). *Milwaukee Journal Sentinel*, p. 1A.
- Gallagher, E. J., Lombardi, G., & Gennis, P. (1995). Effectiveness of bystander cardiopulmonary resuscitation and survival following out-of-hospital cardiac arrest. *Journal of the American Medical Association*, 274, 1922–1925.
- Gilovich, T. (1991). *How we know what isn't so*. New York: The Free Press.
- Gilovich, T. (1997, March/April). Some systematic biases of everyday judgment. *Skeptical Inquirer*, 21(2), 31–35.
- Griffin, R. J., & Sen, S. (1995). Causal communication: Movie portrayals and audience attributions for Vietnam veterans' problems. *Journalism and Mass Communication Quarterly*, 72, 511–524.
- Hamilton, D. L. (1979). A cognitive-attributional analysis of stereotyping. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 12, pp. 53–84). New York: Academic Press.
- Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley.
- Hilton, D. F. (1988). Introduction: Images of science and commonsense explanation. In D. F. Hilton (Ed.), *Contemporary science and natural explanation: Commonsense conceptions of causality* (pp. 1–8). Brighton, Great Britain: Harvester.
- Hohler, B. (1993, February 2). Super Bowl gaffe: Groups back off on violence claims. *The Boston Globe*, p. 1.
- Johnson, B. B., & Slovic, P. (1995). Presenting uncertainty in health risk assessment: Initial studies of its effects on risk perception and trust. *Risk Analysis*, 15, 485–494.
- Kahneman, D., Slovic, P., & Tversky, A. (Eds.). (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge, England: Cambridge University Press.
- Kahneman, D., & Tversky, A. (1982). Variants of uncertainty. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 509–520). Cambridge, England: Cambridge University Press.
- Marin, R., & Gegax, T. T. (1996–1997, December 30–January 6). Conspiracy mania feeds our growing national paranoia. *Newsweek*, 64–71.
- Meyer, P. (1979). *Precision journalism* (2nd ed.). Bloomington: Indiana University Press.
- Meyer, P. (1991). *The new precision journalism*. Bloomington: Indiana University Press.
- Nolan, D. (1986, October). Airline safety: The shocking truth. *Discover*, 30–31, 52–53.
- Paulos, J. A. (1988). *Innumeracy: Mathematical illiteracy and its consequences*. New York: Vintage.
- Paulos, J. A. (1995). *A mathematician reads the newspaper*. New York: Basic Books.
- Ross, L., & Anderson, C. A. (1982). Shortcomings in the attribution process: On the origins and maintenance of erroneous social assessments. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 129–152). Cambridge, England: Cambridge University Press.
- Singer, E., & Endreny, P. (1987). Reporting hazards: Their benefits and costs. *Journal of Communication*, 37, 10–26.
- Sowell, T. (1996, June 26). By the numbers. *The (Grafton, NH) Country Chronicle*, p. 11.
- Subcommittee on Quantitative Literacy Requirements. (1995). Quantitative reasoning for college graduates: A complement to the standards [report]. Mathematical Association of America. [Online]. Available: <http://www.maa.org>.



- Tankard, J. (1994). Visual crosstabs: A technique for enriching information graphics. *Mass Comm Review, 21*, 49-66.\
- Tufte, E. (1983). *The visual display of quantitative information*. Cheshire, CT: Graphics Press.
- Tufte, E. (1997). *Visual explanations*. Cheshire, CT: Graphics Press.
- Tversky, A., & Kahneman, D. (1982). Judgment under uncertainty. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 3-20). Cambridge, England: Cambridge University Press.
- Utts, J. M. (1996). *Seeing through statistics*. Belmont, CA: Duxbury/Wadsworth.
- Wainer, H. (1997). *Visual revelations*. New York: Copernicus.