A

SPOTTING QUESTIONABLE NUMBERS

The billion is the new million. A million used to be a lot. Nineteenth-century Americans borrowed the French term millionaire to denote those whose wealth had reached the astonishing total of a million dollars. In 1850, there were 23 million Americans; in the 1880 census, New York (in those days that meant Manhattan; Brooklyn was a separate entity) became the first U.S. city with more than one million residents.

At the beginning of the twenty-first century, a million no longer seems all that big. There are now millions of millionaires (according to one recent estimate, about nine million U.S. households have a net worth of more than $1 million, not counting the value of their principal residences).[1] Many houses are priced at more than a million dollars. The richest of the rich are billionaires, and even they are no longer all that rare. In fact, being worth a billion dollars is no longer enough to place someone on *Forbes* magazine's list of the 400 richest Americans; some individuals have annual *incomes* exceeding a billion dollars.[2] Discussions of the U.S. economy, the federal budget, or the national debt speak of trillions of dollars (a trillion, remember, is a million millions).

The mind boggles. We may be able to wrap our heads around a million, but billions and trillions are almost unimaginably big numbers. Faced with such daunting measurements, we tend to give up, to start thinking that all big numbers (say, everything above 100,000) are more or less equal. That is, they're all *a lot*.

Envisioning all big numbers as equal makes it both easier and harder to follow the news. Easier, because we have an easy way to make sense of big numbers. Thus, we mentally translate statements like "Authorities estimate that HIV/AIDS kills nearly three million people worldwide each year" and "Estimates are that one billion birds die each year from flying into windows" to mean that there are *a lot* of HIV deaths and *a lot* of birds killed in window collisions.

But translating all big numbers into *a lot* makes it much harder to think seriously about them. And that's just one of the ways people can be confused by statistics—a confusion we can't afford. We live in a big, complicated world, and we need numbers to help us make sense of it. Are our schools failing? What should we do about climate change? Thinking about such issues demands that we move beyond our personal experiences or impressions. We need quantitative data—statistics—to guide us. But not all statistics are equally sound. Some of the numbers we encounter are pretty accurate, but others aren't much more than wild guesses. It would be nice to be able to tell the difference.

This book may help. My earlier books—*Damned Lies and Statistics* and *More Damned Lies and Statistics*—offered an approach to thinking critically about the statistics we encounter.[3] Those books argued that we need to ask how numbers are socially constructed. That is, who are the people whose calculations produced the figures? What did they count? How did they go about counting? Why did they go to the trouble? In a sense, those books were more theoretical; they sought to understand the social processes by which statistics are created and brought to our attention. In contrast, this volume is designed to be more practical—it is a field guide for spotting dubious data. Just as traditional field guides offer advice on identifying birds or plants, this book presents guidelines for recognizing questionable statistics, what I'll call "stat-spotting." It lists common problems found in the sorts of numbers that appear in news stories and illustrates each problem with an example. Many of these errors are mentioned in the earlier books, but this guide tries to organize them around a set of practical questions that you might ask when encountering a new statistic and considering whether it might be flawed. In addition, all of the examples used to illustrate the various problems are new; none appear in my other books.

This book is guided by the assumption that we are exposed to many statistics that have serious flaws. This is important, because most of us have a tendency to equate numbers with facts, to presume that statistical information is probably pretty accurate information. If that's wrong—if lots of the figures that we encounter are in fact flawed—then we need ways of assessing the data we're given. We need to understand the reasons why unreliable statistics find their way into the media, what specific sorts of problems are likely to bedevil those numbers, and how to decide whether a particular figure is accurate. This book is not a general discussion of thinking critically about numbers; rather, it focuses on common flaws in the sorts of figures we find in news stories.

I am a sociologist, so most of the examples I have chosen concern claims about social problems, just as a field guide written by

an economist might highlight dubious economic figures. But the problems and principles discussed in this book are applicable to all types of statistics.

This book is divided into major sections, each focusing on a broad question, such as: Who did the counting? or What did they count? Within each section, I identify several problems—statistical flaws related to that specific issue. The discussion of each problem lists some things you can "look for" (that is, warning signs that particular numbers may have the flaw being discussed), as well as an example of a questionable statistic that illustrates the flaw. (Some of the examples could be used to illustrate more than one flaw, and in some cases I note an example's relevance to points discussed elsewhere in the book.) I hope that reading the various sections will give you some tools for thinking more critically about the statistics you hear from the media, activists, politicians, and other advocates. However, before we start to examine the various reasons to suspect that data may be dubious, it will help to identify some statistical benchmarks that can be used to place other figures in context.

B

BACKGROUND

Having a small store of factual knowledge prepares us to think critically about statistics. Just a little bit of knowledge—a few basic numbers and one important rule of thumb—offers a framework, enough basic information to let us begin to spot questionable figures.

B1 Statistical Benchmarks

When interpreting social statistics, it helps to have a rough sense of scale. Just a few benchmark numbers can give us a mental context for other figures we encounter. For example, when thinking about American society, it helps to know that:

- The U.S. population is something over 300 million (you may recall the hoopla when the magic figure was reached in late 2006).
- Each year, about 4 million babies are born in the United States (the 2004 total was 4,112,052).[1] This is a surprisingly useful bit of information, particularly for thinking about young people. How