

The Numbers Are Far Too Often Wrong

By *Everett Carll Ladd*

Social statistics have become a national necessity. Expectant, demanding, optimistic that problems can be solved, Americans want to know the answer to "How Are We Doing?" in virtually every area of social and political life.

Accurate information on the many dimensions of this question is essential to democracy itself. In our postindustrial era, citizens are expected to respond quickly to many issues that are at once complex and remote from personal experience. To do so intelligently we need reliable summary data. No one, no matter how diligent, can be expert in every area. If we're given wrong readings on social performance, our opportunity to set the direction of public policy is effectively curbed.

Researchers have devised quantitative measures of just about everything. In the economic sphere we track every twist and turn of the Gross Domestic Product (GDP), family income, poverty, inflation, productivity, and much more. On matters environmental, we seek to gauge the risks attendant on the use of an enormous array of chemicals in everything from pest control to aerosol cans. Human health and how dietary regimens affect it are explored with unmitigated persistence. The potential of assorted "megatrends and tendencies"—involving America's competitive position in the world economy, global warming, world population growth, energy supply and consumption, etc.—are insistently measured. And, I should not fail to note, how Americans see their society and politics, from health care issues to Bill Clinton's presidential performance to the prospects for young Americans' futures, are dissected in hundreds of opinion surveys each year. If effort, amounts spent, and volume of results are proper indicators in this passion for measurement, the end product is impressive.

Unfortunately, such testimonials are fundamentally irrelevant to the

main issue. What we need to know, of course, is how good our measures are. How accurately do they in fact describe the social conditions they claim to describe?

Examined individually, the measures we increasingly live by range from excellent to simply awful. When one is found occupying the latter status, we should want to see it scrapped and the effort either abandoned entirely or a far better substitute approach employed. But, in general terms, the fact that some

measures of social performance are flawed is relatively unimportant for the interested, non-specialist citizen. He or she would suppose, of course, that the sheer complexity of social life would inevitably lead to failure in at least some measurement efforts. If the *overall* record is sound, the citizen should be prepared to live with the reality of occasional failure. There comes a point, however, where the degree of error either in the measures themselves or the way they are used becomes so large as to constitute an unacceptable barrier to modern-day citizenship.

How, for example, is the American economy in fact doing vis-a-vis that of the Japanese? Is it relatively stronger or weaker? How is it trending? The thoughtful citizen might well be far more inclined to support change in economic policy if our relative position is declining seriously than if it is holding its own or even improving. How adequate are available summary data on this matter?

Everyone knows that in one sense it's true that "living is dangerous to your health." Social stress, occupational demands, dietary intakes, substances in the air we breathe—all impact on our health—sometimes negatively. But choices have to be made. Where are the threats to health greatest? What are the instances where if the impact is negative, it is modestly so and energy—both psychic and that involving national effort through government spending—should be constrained? Where are the measures sufficiently flawed that we simply don't know what the health impact is?

Since *Public Perspective* began publication five and a half years ago, we have believed that one of its principal missions should be to examine critically the adequacy of social data. Given our position in the field of public opinion and survey research, we concentrate on poll data. This mandate is in fact broad, since surveys are used to measure many things besides opinions.

MEASURING AMERICAN SOCIETY

pp. 4-6: R. C. Lewontin
Excerpted from *THE NEW YORK REVIEW OF BOOKS*,
April 20, 1995.

≈

pp. 7-9: Barbara Everitt Bryant & William Dunn
Excerpted from *MOVING POWER AND MONEY: THE
POLITICS OF CENSUS TAKING*

≈

pp. 10-11: Gregg Easterbrook
Excerpted from *A MOMENT ON EARTH*

≈

pp. 12-14: Bruce N. Ames and Lois Swirsky Gold
Excerpted from "The Causes and Prevention of
Cancer: The Role of Environment"

≈

pp. 15-17: Joseph W. Duncan and Andrew C. Gross
Excerpted from *STATISTICS FOR THE 21ST CENTURY*

≈

pp. 13, 17: John Allen Paulos
Excerpted from *A MATHEMATICIAN READS THE NEWS-
PAPER*

††

Measuring American Society — Everett Carll Ladd

What's more, there seems to be considerable commonality in the nature of measurement problems across social statistics. For one thing, as R. C. Lewontin, a population geneticist at Harvard University, observed in a recent essay, many efforts at measuring social behavior and performance confront the inevitable fact that "human societies are made by self-conscious organisms." [Excerpts from Lewontin's essay follow.] We care about our social performance and try to influence it. Sometimes in the process, we let our wishes shape our understanding of what is—even to the point of collecting data in a way which otherwise would be seen as flawed. Problems like these are not limited to opinion studies. They apply with equal force across many measures of economic performance, environmental impacts, human health correlates, etc.

Getting social data to measure the social world adequately and accurately is, in any full form, unattainable. Social reality is too complex and uncertain, and we care too much about the outcomes that numbers purport to describe to ever really leave them alone. Nonetheless, most of us have a strong interest in accurate social description most of the time, and our resources for accurate information gathering are substantial. Greater awareness of the many deficiencies in existing measures of social performance can, given these things, spur improvement. A host of recently published books are in this context most helpful. We are pleased to be able to publish excerpts from five of them.

Barbara Everitt Bryant, who served as director of the US Census Bureau from 1989 to 1993, ably surveys the intense political pressures which operate upon the decennial Census in *Moving Money and Power: The Politics of Census Taking*, (which she wrote together with William Dunn). The distribution of many billions of dollars in federal funds hinges on the Census counts, adding immeasurably to the inherent difficulties in getting accurate and agreed-upon data on national population. In another book excerpt, Gregg Easterbrook of *Newsweek Magazine* examines the misreadings of trends in world population. The selection published here, from *A Moment on the Earth*, is part of a broad inquiry into our contemporary environmental record.

Occasionally, one or two analysts establish themselves as "policemen" of data in sectors of social measurement. They contribute, that is, a continuing stream of thoughtful assessment, which comes to be recognized as both informed and fair-minded. Bruce N. Ames and Lois Swirsky Gold have done this for the health effects of chemical substances. We excerpt here from their chapter on "The Causes and Prevention of Cancer: The Role of Environment," which appears in Ronald Bailey (ed.), *The True State of the Planet*. Mathematician John Allen Paulos has long been casting a wry and practiced eye on the mathematics of

“*There comes a point, however, where the degree of error either in the measures themselves or the way they are used becomes so large as to constitute an unacceptable barrier to modern-day citizenship. We need reliable summary information. No citizen, no matter how diligent, can be an expert in every area.*”

measuring social outcomes and performance. The two brief excerpts here are from *A Mathematician Reads the Newspaper*. Many readers will recall Paulos' acclaimed earlier work, *Innumeracy* (Hill and Wang, 1988).

Economics is surely not immune to measurement problems. They abound in this field, made fertile by powerful self-interest and by the subject's sheer complexity. Arguments are ongoing, in particular, about measurements of income, the distribution of income and wealth, rates of inflation, and the incidence of poverty. A new effort to recalculate poverty statistics, undertaken by the National Academy of Science, already promises to become a center of controversy.

An especially interesting and instructive, if unfortunate, chapter on economic mismeasurement is provided by work on "the energy problem." Back in the 1970s, in the days of Arab oil embargoes and wait-90-minutes-to-fill-your-tank gas lines at US service stations, the specter of long-term economic disaster

was raised by many supposed experts. The world was running out of oil. The vast imbalance between demand (high) and future supplies (low) would inevitably "empower" the (relatively) oil-rich nations of the world, especially those of the Middle-East. The United States and other advanced industrial nations would increasingly be held hostage, as the "oil weapon" was turned on them. Sweeping and expensive policy responses were needed, given the magnitude of the impending crisis. As President Carter put it in an address to the nation on April 18, 1977, "the energy crisis has not yet overwhelmed us, but it will if we do not act quickly....The oil and natural gas we rely on for 75% of our energy are simply running out....We now believe that early in the 1980s[!] the world will be demanding more oil than it can produce."

Were anything like this shown by sound data to be likely, actions that in less strenuous circumstances would seem outrageously wrong-headed, might be justified. One such policy, in fact acted on in this climate of near-panic, involved "alternative energy." With the price of oil held likely to rise to \$100 a barrel, the Public Utility Regulatory Policies Act of 1978 (PURPA) was passed, to encourage alternative sources of electricity. It provided for payments at way over current market prices to independent producers of electrical power, and it required the established utilities to buy the power.

Those of us in this regard unfortunate enough to own property (and thus use electricity) in Maine have experienced some of the consequences with notable force. Central Maine Power Company now has astronomically high rates, in no small part because it is saddled with contracts requiring it to pay per-unit costs for electricity from various independent (often wood-burning) producers at rates many times those prevailing in open markets.

Writing about the magnitude of the overall problem resulting from incredibly bad 1970s estimates of production and consumption of oil, Jeff Bailey of the *Wall Street Journal* concludes: "In all, the hundreds of power plants qualified under PURPA now represent, in effect, a \$37 billion roadblock to lower utility prices. That is how much the plants will receive through the year 2000 over and above estimated market prices, as calculated by Resource Data International, Inc., a consulting firm in Boul-

der, CO. The US, meanwhile has a glut of generating capacity. This glut plus major advances in turbine technology would, in an unregulated market, be driving electricity prices sharply downward in many regions now.”¹

“
Why, indeed, would any researcher start with a presumption that sufficient numbers of people will ‘level’ with regard to their sex lives in interviews with persons whom they have never known and have no reason to trust, in interviews where they have no real incentive beyond a small cash payment and perhaps a desire to say something?
”

In *Statistics for the 21st Century*, Joseph W. Duncan and Andrew C. Gross further describe the magnitude of the 1970s energy miscalculations. The forecasts made in 1973 anticipated a growth in demand so great that by 1985 world production of crude oil would be straining to reach a total output of roughly 200 quadrillion BTUs of oil-derived energy, with all that such a figure would mean for oil prices. In fact, in 1985 the production of crude oil worldwide was just over half as great as the 1973 estimate had foreseen—and still there was a glut of oil. Bad numbers can have grievous consequences. Among other things, nations may as a result elect economic policies that have the practical consequence of robbing millions of consumers.

My own field of opinion research has its own fair share of mismeasurement. Survey research offers a valuable tool for gathering information efficiently, through small, carefully designed samples of general populations. But no matter how carefully a sample is designed and executed, it cannot escape problems resulting from the inadequacy

of questions designed to elicit the information sought after. There can be many different reasons for “question failure”—including some which occur even when the questions are framed as ably as current knowledge permits—and, perhaps, as ably as it is ever possible for them to be framed.

Reviewing one failure in survey measurement, R. C. Lewontin writes that “there are some things in the world that we will never know and many that we will never know exactly. Each domain of phenomena has its characteristic grain of knowability.”² This has been a hard lesson for some in survey research to learn. If we want to know something about people’s thinking, or their behavior in areas where asking them is the only way of finding out about it, it’s tempting to think that a properly designed battery of questions will get us there. Perhaps. But maybe not.

Lewontin wrote the above in a review of a book by Edward O. Laumann, John H. Gagnon, Robert T. Michael, and Stuart Michaels, *The Social Organization of Sexuality: Sexual Practices in the United States*. Much publicized when it appeared a few months ago, *The Social Organization of Sexuality* reports on a major national survey done by the National Opinion Research Center of the University of Chicago. The study was sponsored by the Robert Wood Johnson, Rockefeller, Kaiser, Mellon, MacArthur, and Ford foundations. Lewontin’s review is a devastating critique of the survey, and it has prompted an angry protest by the book’s authors.³

Laumann and his colleagues complain that Lewontin is writing about a field outside his own area of special competence. That’s true, though population geneticists are hardly unfamiliar with statistical measures. The important point, however, involves not Lewontin’s disciplinary background, but rather the substance of what he has to say.

The main issue, Lewontin argues, is whether those who responded to this survey—which asked sensitive questions about sexual practices—in fact “told the truth.” “...Why should anyone lie on a questionnaire [on personal sexual be-

havior] that was answered in a face-to-face interview with a total stranger?,” Lewontin asks, his frustration yielding to sarcasm. Why, indeed, would any researcher start with a presumption that sufficient numbers of people will “level” with regard to their sex lives in interviews with persons whom they have never known and have no reason to trust, in interviews where they have no real incentive beyond (for some) a small cash payment and perhaps a desire to say something?

The point isn’t, of course, that people inevitably lie, or more precisely misrepresent, their actual views or behaviors when asked in surveys. Every experienced survey researcher knows that there are great differences in the willingness and ability of respondents to “tell the truth” across different areas of inquiry. We are vastly less inclined to reveal our income to an interviewer than we are to tell him or her how we will vote for president, for example. In the long list of subjects examined through survey research, none poses a higher likelihood of misrepresentation than personal sexual practice. That this is so isn’t the “fault” of the investigators.

Everything we might wish to know can’t be known. Some things can’t be known now, some things never. The task of a more mature social science of measurement is finding out where one can get reliable information and what good can come from having it—and then designing studies, whatever the domain, that gather the needed information as free from all sorts of contamination as possible. The final step is presenting the data cleanly and clearly to interested audiences, stressing their unavoidable limitations. In this task, there is more than enough for all of us in empirical social science to do in all our career lifetimes.

Endnotes:

¹ Jeff Bailey, “Carter-Era Law Keeps Price of Electricity Up in Spite of a Surplus,” *The Wall Street Journal*, May 17, 1995, pp. 1, 8.

² Richard C. Lewontin, “Sex, Lies, and Social Science,” *The New York Review of Books*, April 20, 1995, p. 28.

³ Letter by Laumann, et al, “Sex, Lies, and Social Science: An Exchange,” *New York Review of Books*, May 25, 1995, pp. 43-44.