

Part 1

KEYNOTE ADDRESS

WHAT WE, THE PUBLIC, NEED TO KNOW

## THE NUMBERS YOU COLLECT -- WHAT WE NEED TODAY AND TOMORROW

An address to the Seminar on Statistical Methodology in the Public Service, sponsored by the Council of Professional Associations on Federal Statistics.

By Victor Cohn, former science editor of the Washington Post, 1994-1966 scientific and public affairs fellow of the American Statistical Association.

You are important. You are important to us, the public, to us, the media, to business, to public policy makers, everyone.

We hear that numbers are dull. A turnoff. But we are ruled by numbers, by our incomes, our expenditures, our taxes, by the prices we pay and on and on. The numbers you collect, the numbers you convey to the public and how you convey them affect my life and the lives of millions. They affect public decisions, business decisions, my decisions.

Statistics uses numbers. But statistics, I have learned, is really about accurate and useful information. It is about straight thinking. The better and clearer we all can be, the clearer and easier become many lives.

What do we, the public, and we of the media want of you? And I don't say it's easy.

First of all and most of all, we want honest numbers. Accurate, of course -- and I know that's not always easy -- or at least your best possible estimates. But also politically honest numbers, data uninfluenced by the political or the personal considerations of the people you work for, or those who influence the people you work for. My hero is Janet Norwood, who spent 25 years at the Bureau of Labor Statistics, much of it as commissioner, and steadfastly resisted all attempts to paint rosy or unrosy pictures for the benefit of whatever Democratic or Republican administration.

Next, how can you explain what you do? How can you help us -- us, the public, and us, the media -- discern what to possibly believe and what to question when assertions that include or should include statistics, are thrown at us? How can we separate the truth from the trash? Or the probable truth from the probable trash?

I am here because I happened to write a little book called *NEWS & NUMBERS: A Guide to Reporting Statistical Claims and Controversies in Health and Other Fields*.\* It has received a

pretty good reception in the journalistic community, and, to my amazement, has gone into a sixth printing and even been recommended -- in the *Journal of the American Statistical Association*, *Chance* and medical, scientific and other journals -- to statisticians, physicians, scientists and even public policy makers.

Well, I have a confession to make. I was asked to write an article about the book for the quarterly publication of the National Association of Science Writers, and I not only wrote the article but I also wrote the headline, which was: MATH ILLITERATE WRITES GUIDE TO STATISTICS. Yes, the only thing I remember about high school math is that the instructor, a kind Mr. Peterson back in Minneapolis, said he'd give me a "D" for the last semester if I promised not to take any more math. I promised and avoided math, also physics and chemistry, all through high school and college. With this wonderful preparation, I naturally went into science writing.

I tell you this because my problem, like the problem of all reporters and of the public, became just this. When it comes to any controversial issue -- and what issue isn't, whether government, welfare, education, crime, health, risk, the environment or whatever? -- one side says, The sky is falling, the other, It isn't! One side says, We have a wonderful fix, the other, You'll ruin things!

Does this sound familiar about Washington? Does it sound familiar about our recent presidential race? Who to believe? How can we assess the credibility of what we hear?

I suspect strongly that this is not a problem of the media and the public alone. I know that I am addressing many trained statisticians and some who may not have statistics degrees, but are experienced in dealing with statistics. Surely many of you sometimes have to try to decide whether to believe or use some statistics and statements that are thrown at you. And you may sometimes wish for some simple, in part non-statistical questions and concepts that you can use -- either verbally or as you read -- to help separate the probable truth, something you can probably believe, from the probable trash.

When I was writing *NEWS & NUMBERS*, many fine statisticians and epidemiologists helped me find some simple, yet revealing questions.

Such as:

**How do you know?** Are you just telling us something you "know" or have "observed" or "found to be true"? Or have you done or found any studies or experiments?

**What are your data?** Your numbers? How or where did you get them?

**How sure** can you be about them? Are there any possible flaws or problems in your conclusions? Salespeople are sure. Honest folks admit uncertainties.

**How valid** -- and we have to explain to the media and the public that in science, valid means accurate -- are your numbers and conclusions?

**How reliable** -- and we have to explain that this means how reproducible -- are they?  
Have results been fairly consistent from study to study?

**What is your degree** of certainty or uncertainty by accepted tests?

**Who disagrees** with you? And why?

The interesting thing is that the very way someone answers, or dodges, such questions can reveal much.

I believe that we can then go a long way toward helping non-statisticians discern the probable facts from the probable trash, a long way toward judging claims and statistics that are thrown at us, by explaining that there are six basic concepts that apply to all science, all studies and virtually all knowledge of society and the universe. And that remembering these can help anyone ask, "How do you know?" with a considerable degree of sophistication.

I can't go into detail here on these, but they are:

**1. UNCERTAINTY** -- the certainty of uncertainty, the fact that all science is almost always uncertain, or uncertain to a degree. Nature is complex, people are more so, research is difficult. Almost all anyone can say about atoms or cells or human beings or the biosphere is: there is a strong possibility that such-and-such is true, and we may know more tomorrow.

This can help us tell people why things so often seem settled one way today and another tomorrow, and why so much is debated, whether the effects of global warming, a pesticide, a low-fat diet or a medical treatment. Yet uncertainty need not impede crucial action if society understands and uses the rest of these principles.

**2. PROBABILITY** -- the use of probability and how scientists live with uncertainty by measuring probability. And how the laws of probability and chance tell us to expect some unusual, even impossible sounding events. And that there is something else to remember when someone says, "How do they know this stuff isn't causing harm?" Science cannot prove a negative. No one can prove that Little Green Men from Mars have not visited earth. The burden of proof should be on those who say something is true.

**3. POWER** -- the power of large numbers in finding something if it's there, say an increase in cancer in workers exposed to some substance, and how we must be wary of studies with only a small number of cases. How we should ask too about any association's statistical strength, its odds.

**4. BIAS** -- the danger of bias. How bias in science means introducing spurious associations and reaching unreliable conclusions by failing to consider other influential factors -- confounding variables or co-factors, in plain language, other explanations. We can learn to watch for bias by asking, "Are there any other possible explanations?"

5. **VARIABILITY** -- its ubiquity. How a common pitfall of science is that everything measured or studied varies from measurement to measurement. Every human experiment, repeated, has at least slightly and sometimes markedly different results.

6. Finally, the fact that there is a **HIERARCHY** of **STUDIES**, from the least to the generally most believable, starting with simple anecdotes, going on to more systematic observation or "eyeballing," then proceeding to true experiments, comparing one population or sample with another, under controlled or known conditions.

All this tells us that one study rarely proves anything, that when someone tells us, "I've done a study," we should ask, "What kind? How confident can you be in the results? Were there any possible flaws?" Et cetera. It tells us that we must seek out the most credible evidence, the most likely probabilities. It tells us that we must look for a consensus among the best studies and the best, most neutral observers, those who remain scientists rather than salespeople.

Now, how can you get our message out? How can you get it to the public and the media? How can you make your information understandable and useful?

Reporters, the most conscientious, have a lot of constraints and demands. What we need are numbers and conclusions that are put to us clearly and in language we can understand. We're not the only ones. John Tucker at the Board of Mathematical Sciences at the National Research Council recently told me that part of the reason statisticians or people of any description responsible for collecting and disseminating numbers get let off the table, get left out of decision making, is that the very people they work with or work for don't understand their numbers.

Authors or investigators sometimes fail to include some numbers that any of us would want for intelligent decision-making. Researchers wrote in the *New England Journal of Medicine* that "one to three drinks" a day may help protect against heart attacks. They defined a drink as 13.2 grams of alcohol. But with whiskey, wine and beer all at various proofs or alcohol content, neither the journal's article or editorial nor any news reports I read or saw told in common language how much booze, wine or beer one should consume to drink no more than the prescribed grams. A free-pouring drinker could down far more.

The media and the public have a lot of trouble understanding risks. Part of the problem is a failure to understand the difference between a plain number and a rate, and the fact that a rate has to mean so many per so many per unit of time. A *Washington Post* headline once read, "Airline Accident Rate Is Highest in 13 Years." The story, like many others misusing the word "rate," reported no rate at all, merely death and crash totals. A correction had to be printed pointing out that the number of accidents per 100,000 departures -- the actual rate, the "so many per so many" - had been declining year after year.

Another problem: the word "safe." I believe we should pretty much quit using it. When the FDA or EPA or some other agency describes something as safe, it is almost never completely so.

We commonly have to accept some risk to achieve some gain. Current example: air bags in our cars. Instead of calling a prescription drug, for example, safe, maybe we should start to say "relatively safe" or "as safe as possible" or some such phrase and try to indicate in the best numbers and rates we can the degree of safety or risk, the rate of adverse events. When we fail to do this, when we announce something as "safe" and that turns out to be wrong -- the frequent case -- we create the public confusion and lack of confidence that exists today.

Some practical rules for dealing with the media?

Let me quote and add to some stated by Dr. John Bailar, former statistical adviser to the *New England Journal of Medicine*, now head of biostatistics and epidemiology at the University of Chicago.

\* "Give it all the time it takes." Indeed do this, I would add, when you are phoned or approached by a reporter who seems interested in learning something and doing a good job, and obviously devoted some time and effort to learning at least a little something before speaking to you. But if you are suddenly quizzed by someone who seems to know nothing and is interested only in a quick quote or sound bite -- and that's all the exposure you're going to get -- think carefully about what you want to say, say just it, and don't say much else if you think that's the honest and important message that should be conveyed.

This may sound cynical but the superficial reporter deserves no better. The public may deserve a great deal better, so it's important to say the right thing.

\* "Explain in short, simple sentences." And use plain language. Statistics and mathematics have their own useful languages, but these aren't the languages that the rest of us use. Words like "normal," "mean," "significance," "validity," "reliability," "power" have different meanings in statistical and common language.

\* Help your audience understand the issues by familiar examples or phrases. Instead of saying "the U.S. homicide rate is 10.4 per 100,000 per year," say "one person in every 10,000 will be a homicide victim."

\* If you say "homicides were up by X percent," add the actual number, the "10.4 per 100,000" that tells people their own chances.

\* "If you don't know or aren't sure, say so. Your stock will rise."

\* Be prepared to answer the questions: What would you recommend that the public do about this? What would you do, or what are you doing? What would you want your family to do?

But be clear whether you are responding from an official position or stating your personal opinion. "Distinguish between what is the science," the facts, and "what do they mean?" -- how they should be interpreted in your opinion.

\* "Any constraints on use -- that is, 'off the record' -- must be on the table before the words are out. Afterwards is too late. . . . Be clear about what you mean by 'off the record,' and make sure there is agreement."

In Washington, at least, not necessarily elsewhere, such phrases have a pretty precise meaning.

"Off the record" means "you can't use this," though the reporter may be free to get the same information from another source.

"On background" means "you can use it, but don't use my name" -- and get an understanding on how it may be attributed, whether to a Department of Something official" or "Congressional aide" or "government official" or whatever.

"Deep background" means you can use it but only on your own authority, not with attribution to anyone.

I don't mean to encourage you to have such recourse. Have the courage to say what you mean and to stand behind what you say. Statements from anonymous sources lack the credence of "Max Smith said . . ."

\* If you would increase your chances of being reported accurately, provide a copy of your paper and charts and perhaps other background material. Tell your information people to add a copy of the full paper to their press release.

\* Offer to review what is written for accuracy of facts -- and facts alone, not writing style or interpretation. But do not demand control. Put it generously -- "I'd be glad to have you call me if you want to recheck any of the facts or the way you've put it." If you are called, don't try to wiggle out of what you have previously said, unless you think there's something plainly wrong. Again, have the courage of your convictions, even when they sound blunt when shown or read back to you.

In doing all this, use your public information people, but find the good ones, those who not only have shown they know how to get ink and tube time, but also have taken the trouble to learn your subject, and come sniffing around from time to time to find out what your are dong. Value and try to promote these people.

But also be available yourself to media calls and inquiries. Reporters value and use public information people who know their subject. But they also want to get their information from the horse's mouth. They want to talk to the people at the top who are responsible, and they want to talk to the people who put together and really know the data. When Cap Weinberger was secretary of then-HEW, he and his deputy, Frank Carlucci, put out directives telling HEW people at every level: talk to reporters, answer reporters' phone calls. And they did. And HEW had a remarkably good press. Silence, reluctance to talk breeds suspicion. Openness builds trust.

In 25 years at the *Washington Post*, I had a lot of occasions to call the National Center for Health Statistics, sometimes getting to someone through public affairs but often phoning directly, and never failed to get immediate and cheerful help. The same goes for the statistics people at the National Cancer Institute, whose numbers were in the news a lot, and not always with good news about progress against cancer. I'm sure the same would be true of some other agencies. I've heard good things about the Census Bureau and the BLS. And not such good things about some others.

Finally, let me repeat: candor, accessibility build trust. I said at the outset that you are important. This means that we depend on you to tell us candidly how your conclusions, your data meet the tests of validity and reliability and the other statistical rules that you know and use. Tell us that in a thousand difficult situations -- risks, apparent risks, dangerous diseases, population trends, educational progress or lack of it, the efficacy or lack of efficacy of welfare programs, the performance of many agencies and systems, yes, and in that controversial count, the census -- your conclusions, your data are not written in stone, but rather, "Here is the best we can say today, and we may know more tomorrow. The pursuit of knowledge is a continuing struggle."

About the census, now in contention. I can't give you a single formula for getting across the message that the best count in today's United States inevitably must include the best possible estimates, which includes sampling. Some of the opponents of sampling of course know that, which is why they oppose the best possible estimates.

I can only say: Educate, educate, educate, and, where appropriate to your job, politic too -- it's a respectable activity in Washington. Don't just wait to be called on to testify. Keep repeating your best opinions and the best possible facts in every possible way, and get them again and again to the political leaders and their staffs. Get them to the media. Do not give in. Obey, of course, what directives you must. But do not retreat on your beliefs. And eventually, truth may prevail.

The candor I have recommended on this and all issues -- including the admission at times of uncertainty -- flies in the face of the usual Washington argument for one course or another, which is "Here are the immutable facts, my friends, and if you follow my recommendations, everything is going to be wonderful." Honesty may make selling some ideas a bit harder, but it distinguishes an honest seeker after truth from a snake oil seller. It can build trust. It can prevent future disappointment and disillusionment and lack of trust.

A wise person once said, "If you would have public confidence, confide in the public." Any other course may work temporarily, then disastrously backfire and build the mistrust that is so common in America today.

\* *Iowa State University Press, Revised Edition, 1994.*