University of Waterloo W. H. Cherry

SAMPLING and SURVEY SAMPLING: Polling – *The polls dance to anybody's tune*

EM9113: The Globe and Mail, December 7, 1991, page D5

MASSAGING STATISTICS / *While public-opinion polls promise a quest for truth, many are little more than vehicles for pitching a product or opinion. The business of research has become pervaded by bias and distortion*

The polls dance to anybody's tune

By CYNTHIA CROSSEN
THE WALL STREET JOURNAL

CONSUMERS overwhelmingly preferred a Chrysler to a Toyota after test-driving both, contends a study sponsored by Chrysler. The vast majority of U.S. college students picked Levi's 501 jeans as the most "in" clothing, says a study sponsored by Levi's. And in separate studies funded by the cloth-diaper and disposable-diaper industries, guess what: Cloth diapers were shown to be better for the environment than paper – and vice versa.

In recent years, research studies like these have become one of America's most powerful and popular tools of persuasion. Once confined to a small circle of polling and research companies and a few universities, the business of studying public opinion and consumer habits has exploded in the past two decades. Today, studies have become vehicles for polishing corporate images, influencing juries, shaping debate on public policy, selling shoe polish and satisfying the media's – and the public's – voracious appetite for information.

Yet while studies promise a quest for truth, many today are little more than vehicles for pitching a product or opinion. An examination of hundreds of recent polls indicates that the business of research has become pervaded by bias and distortion. The result is a corruption of the information used every day by voters, consumers and leaders.

While described as "independent", a growing number of studies are actually sponsored by companies or groups with a real – usually financial – interest in the outcome. And often the study question is posed in such a way that the response is predictable:

When Levi Strauss & Co. asked students which clothes would be most popular this year, 90 per cent said Levi's 501 jeans. They were the only jeans on the list.

A survey for Black Flag said: A roach disc ..... poisons a roach slowly. The dying roach returns to the nest and after it dies is eaten by other roaches. In turn these roaches become poisoned and die. How effective do you think this type of product would be in killing roaches? Not surprisingly, 79 per cent said effective.

A Gallup poll sponsored by the disposable-diaper industry asked: It is estimated that disposable diapers account for less than 2 per cent of the trash in today's landfills. In contrast, beverage containers, third-class mail and yard waste are estimated to account for about 21 per cent of the trash in landfills. Given this, in your opinion, would it be fair to ban disposable diapers? Eighty-four per cent said no.

"There's been a slow sliding in ethics", says Eric Miller, who, as editor of the U.S. newsletter *Research Alert*, reviews some 2,000 studies a year. The scary part is, people make decisions based on this stuff. It may be an invisible crime, but it's not a victimless one.

'The research process, if done right, is much more difficult than people realize. It takes time and money. When deadlines and budgets are short, a lot of the niceties go down the tubes.'

The news media also play a role in disseminating sloppy or biased research to consumers. Journalists often publicize reports about a study without examining the study's methodology, or technical index, to see if it was conducted properly. Statistics are thrown around with abandon, even when sample sizes are so small they're meaningless.

The U.S. media themselves use call-in polls, including one last year conducted by *USA Today* showing that Americans love Donald Trump. A month later, *USA Today* reported that 5,640 of the 7,800 calls came from offices owned by one man, Cincinnati financier Carl Lindner. Mr. Lindner won't comment, but a spokesman told *USA Today* that Mr. Lindner's employees admire Mr. Trump. *USA Today* says its call-in polls aren't meant to be scientific and are strictly for fun.

"A distortion or adulteration has set in in the business", says Solomon Dutka, chief executive of Audits & Surveys, a New York research firm. "Every survey done should have to face an attorney across the room". Recognizing that surveys can be abused, the Council of American Survey Research Organizations has issued a code of standards, covering such topics as responsibility to the public.

There is still much good research being done, of course. In medicine and other physical sciences, research must be quantifiable and replicable to be taken seriously. Moreover, much consumer research is conducted strictly for internal consumption, not public distribution; it is therefore in a company's interest to get it right.

"We will eventually get to a dual standard of information", says Mr. Miller. "There will be a distinction made between research that's done with no hidden agenda, but to create useful information, and research that contains useful information that was generated for a very specific purpose".

In recent years, lean budgets have made everyone who does research, including formerly neutral colleges and universities, a little hungrier for work. "A funder will never come to an academic and say, 'I want you to produce finding X and here's a million dollars to do it'," says Paul Light, associate dean at the Hubert Humphrey Institute at the University of Minnesota. The subtext, he continues, is that if researchers produce the right finding, more work – and funding – will come their way. "Once you're on that treadmill, it's hard to get off. Many universities, which often get a cut of the fee, don't monitor the outside work done under their imprimatur."

Shortages of money and time also contribute to diminishing sample sizes in polls. Researchers say it's best to have at least 1,000 respondents if you hope to project results on to a large population. Yet most of the dozen or more U.S. national polls taken about Clarence Thomas's confirmation interviewed only 500 to 700 people. When broken into subgroups, such as women or blacks, the margin of error goes off the charts – as high as 12 per cent. So when an ABC-Washington Post poll interviewed about 500 adults, roughly half women, and found that more women believed Clarence Thomas (38 per cent) than believed Anita Hill (28 per

(*continued overleaf*)

cent), the opposite could also have been true. (Other surveys, however, did bear out the poll's results.)

Besides interviewing too few people, there are other ways a survey can be flawed: Those surveyed may not be representative of the population, the analysis of the data may be faulty or the conclusions may be screened so only the best are reported. What's more, many studies tackle issues that are so complex they are virtually unresolvable.

And then there are those studies that, though conducted with correct scientific protocol, may have predictable conclusions because the researchers hired to do the studies are known to have come to similar conclusions in the past.

"You can't have an industry study done by that industry be 100 per cent objective", says Carl Lehrburger, who has studied the environmental impact of cloth versus disposable diapers for the cloth-diaper industry.

There are at least four widely publicized studies on diapers that explore the issue of whether disposables are disproportionately responsible for burdening U.S. landfills and fouling the environment. Two studies were sponsored by the cloth-diaper industry and conclude that cloth diapers are friendlier to the world; two others, sponsored by the paper-diaper industry, conclude the opposite.

In studies like these, assumptions and statistics are crunched by computers but entered by humans: Put in one slightly different assumption – that babies use 65 instead of 85 cloth diapers a week – and the picture changes.

Last year, Procter & Gamble Co. released a study showing that holistically, cloth diapers take a greater toll on health, the economy and the environment than disposables. The report was commissioned in response to Mr. Lehrburger's report finding cloth diapers more environmentally friendly. "We found it not in our best interest to be directly critical of the Lehrburger report but to say, let's get some people with a track record in this kind of analysis to take a look", says Scott Stewart, a spokesman for P&G.

P&G's study was used by lobbyists in the California state senate to argue against a bill intended to encourage the use of cloth diapers, written by Senator Lucy Killea. "The studies aren't so much given weight in the actual decision-making as they're used to provide political cover for decisions made for other reasons", says Craig Reynolds, Sen. Killea's chief of staff. "No one wants to say, 'I opposed this thing because I don't want to offend the disposable-diaper industry.' They say, 'Yes, the environment is important, and I have this study showing that there are some concerns that cloth diapers could actually increase pollution'."

But duelling studies can also paralyze decision-making.

"It's gotten to the point where someone will produce a study that statistically demonstrates X or Y, and then the other side will rush out and get an expert to do a study for them", says Ray Sentes, a professor at the University of Regina, and a critic of many asbestos studies that minimize the effects the material has on health. "For 10 years we flash studies at each other. If the practical outcome of a scientific study ends up being a delay of any activity, shouldn't a scientist say, 'You don't need this study'?"

One of the fastest-growing areas of research today is so-called advocacy studies. These are commissioned by companies or industries for public-relations purposes. Simplesse, maker of Simple Pleasures frozen dessert, did a study last summer showing, among other things, that 44 per cent of people who eat a lot of ice cream are likely to take a tub bath. "It was interesting to a lot of people", says Russ Klettke, a spokesman for Simplesse, part of Monsanto Co. "We timed the study for when the media want to write about ice cream, and we have gotten a number of clips back".

Kiwi Brands, a shoe-polish company, commissioned a study earlier this year on the correlation between ambition and shiny shoes. The study found that 97 per cent of self-described "ambitious" young men believe polished shoes are important.

Public-opinion polling began in the 1930s and 1940s; market research became popular with the postwar explosion of consumer demand. But survey techniques were primitive then. Because early U.S. election polls were done face-to-face with as many as 4,000 people, and the results sent by mail and tabulated by hand, the last polls might be done two to three weeks before an election – with plenty of time left for opinion to shift. Partly because of this, the 1948 presidential election was a polling debacle, with all the major pollsters predicting Thomas Dewey, not Harry Truman, would win.

That wasn't as bad as *Literary Digest* magazine's 1936 presidential poll, however. It is widely considered the most inaccurate polling result in history. The poll, which declared Republican Alf Landon a big winner, sampled only people who had telephones or cars or were subscribers – all likely to be affluent.

Today, surveys can be done with astonishing swiftness. With Computer Assisted Telephone Interviewing, called CATI, interviewers sitting in booths can see questions flash on to a computer screen. As each respondent's answer is recorded, the next question automatically flashes on the screen. Jack Honomichl, who publishes *Inside Research*, an industry newsletter, says: "You could, right now, develop a question, call a research company, do 1,000 interviews tonight and have the data on your desk tomorrow.

"But the research process, if done right, is much more difficult than people realize", adds Mr. Honomichl. "It takes time and money. When deadlines and budgets are short, a lot of those niceties go down the tubes".

The "niceties" include things such as careful wording of questions. But even meticulously crafted surveys can get skewed, particularly when they bump up against human shortcomings, like pride and guilt.

"Eighty per cent of Americans are environmentalists", says Mr. Miller of *Research Alert*, citing survey results. "Hog-wash. Eighty per cent of Americans are embarrassed to tell a stranger that they're not for the environment".

With the help of sophisticated statistical techniques, finding 1,000 Americans who can speak for 240 million others has become more reliable. Yet even here there are pitfalls. Poor people and minorities are notoriously underrepresented in telephone sur veys; in surveys taken in shopping malls they are rarely interviewed. What's more, research companies say it's getting more difficult to find people willing to spend 15 minutes answering questions. How representative is a sample of people who will agree to that kind of invasion of privacy, they wonder.

So, in many cases, research simply relies on unrepresentative samples: "There's good news for the 65 million Americans currently on a diet", trumpeted a news release for a diet-products company. Its study showed that people who lose weight can keep it off. The sample: 20 graduates of the company's program who endorse it in commercials.

The Chrysler study showing its cars were preferred to Toyota's included just 100 people in each of two tests. But more important, none of the people surveyed owned a foreign car, so they may well have been predisposed to U.S.-made vehicles. Chrysler says its intent was to survey people who might buy a foreign car.

A study sponsored by American Express Co. and the French government tourist office found that old stereotypes about French unfriendliness weren't true. The respondents: "More than 1,000 Americans who have visited France more than once for pleasure over the past two years" – that is, a select group of people who are already inclined to travel in France. American Express says that it didn't try to hide the makeup of the sample group and that most travellers to France are repeat travellers anyway.

The text of questions such as these, along with the methodology used in the studies, should be readily available to anyone who wants it. But in practice, technical indexes frequently are not offered, often on the ground that the material is proprietary. A

(*continued*)

SAMPLING and SURVEY SAMPLING: Polling – *The polls dance to anybody's tune* (continued 1)

survey done for a coupon-redemption company, Carolina Manufacturer's Service, found that a "broad cross-section of Ameri- cans find coupons to be true incentives for purchasing products". The technical index was available only for a price: $2,000.

Still, even with a convincing methodology, studies conducted by companies or industry groups often have a very predictable outcome. In September, an asbestos study paid for partly by the U.S. government and partly by three asbestos-related industries – real estate, insurance and asbestos manufacturers – found that breathing in buildings with undisturbed asbestos was the same as breathing outside, a conclusion disputed by asbestos activists.

Did that surprise Prof. Sentes of the University of Regina? "I would have had to be sitting down if they'd come up with any other results", he says. "Those guns have been hired for a long time."

The asbestos study was reported in several major publications, including *The Wall Street Journal* and *The Globe and Mail*. Most articles described the panel of scientists as "independent", while noting that the study was partly financed by asbestos-related industries. (*The Globe* mentioned only the Environmental Protection Agency.)

In the end, it's the news media that disseminate the findings of studies – both good ones and bad. "Only if journalists aren't doing their jobs does the public have a problem", says Karen Anderson, public information manager for Battelle Human Affairs Research Centers. "It's the journalist's problem to look at the report or interview the researcher."

But if the journalist doesn't, the consumer of news is often left in a confusing stew of statistics. Many newspapers include explanations of methodology with their polls, but they can be difficult to understand. "The average consumer doesn't know what two standard deviations are", says Mark Clements, head of a New York research firm. Yet, they nonetheless seize on surveys and studies of all types.

As originally presented, the article EM9113 reprinted in this Statistical Highlight #82 emphasized the following eleven of its points by repeating them as separate statements:

a. *When Levi Strauss & Co. asked students which clothes would be most popular this year, 90 per cent said Levi's 501 jeans. They were the only jeans on the list.*

b. *A survey for Black Flag described how its roach disk works: The dying roach returns to the nest and after it dies is eaten by other roaches. In turn these roaches become poisoned and die. How effective do you think this type of product would be in killing roaches? Not surprisingly, 79 per cent said effective.*

c. *A U.S. call-in poll conducted by* USA Today *showed that Americans love Donald Trump. A month later,* USA Today *reported that 5,640 of the 7,800 calls came from offices owned by one man, Cincinnati financier Carl Lindner.* USA Today *was told Mr. Lindner's employees admire Mr. Trump.*

d. *Researchers say it's best to have at least 1,000 respondents if you hope to project results on to a large population. Yet most of the polls taken about Clarence Thomas's confirmation interviewed only 500 to 700 people. When broken into subgroups, such as women or blacks, the margin of error goes off the charts – as high as 12 per cent.*

e. *Two studies, on whether disposable diapers are disproportionately responsible for fouling the environment, were sponsored by the cloth-diaper industry and conclude that cloth diapers are friendlier to the world; two others, sponsored by the paper-diaper industry, conclude the opposite.*

f. *Simplesse, maker of Simple Pleasures frozen dessert, did a study showing among other things that 44 per cent of people who eat a lot of ice cream are likely to take a tub bath.*

g. *The* Literary Digest *magazine's 1936 presidential poll is widely considered the most inaccurate polling result in history. The poll, which declared Republican Alf Landon a big winner over Franklin Roosevelt, sampled only people who had telephones or cars or were subscribers – all likely to be affluent.*

h. *There's good news for the 65 million Americans currently on a diet, trumpeted a news release for a diet-products company. Its study showed that people who lose weight can keep it off. The sample: 20 graduates of the company's program who endorse it in commercials.*

i. *A Chrysler study showing its cars were preferred to Toyota's included just 100 people in each of two tests. More important, none of the people surveyed owned a foreign car, so they may well have been predisposed to U.S.-made vehicles.*

j. *A study sponsored by American Express Co. and the French government tourist office found that old stereotypes about French unfriendliness weren't true. The respondents: More than 1,000 Americans who have visited France more than once for pleasure over the past two years; that is, a select group of people who are already inclined to travel in France.*

k. *In September, an asbestos study paid for partly by the U.S. government and partly by three asbestos-related industries found that breathing in buildings with undisturbed asbestos was the same as breathing outside, a conclusion disputed by asbestos activists.*

(*continued overleaf*)

1 The article EM9113 reprinted in this Statistical Highlight #82 uses the word *research* a number of times, starting in the summary given above the headline on the first side (page HL82.1), and continuing in the second and third paragraphs, in the highlighted statement in the second columm, the paragraphs above and below it and in the last line of this column, and in the three middle paragraphs of the third column, for instance. List each use of the word in context and then categorize each statement by its meaning of *research*, explaining carefully for each category the sense in which the word is used.

2 The bombing of the U.S. federal building in Oklahoma City occurred on the morning of Wednesday, April 19, 1995. On Monday, May 1, 1995, the following item apeared in *The Globe and Mail* on page A16 in the Social Studies section: **Post-bombing mood**: In a *USA Today*/CNN poll, 39 per cent of Americans agree that "the federal government is so large and power- ful that it poses an immediate threat to the rights and freedoms of ordinary citizens." Discuss critically the implications of this news item in light of the information provided about *USA Today* polls in the second-last paragraph of the second column of the article [on the first side (page HL81.1) of this Highlight #82 and emphasized by statement *c* on the third side (page HL82.3)].

3 In the third paragraph of the third column of the article EM9113 [(on the first side (page HL82.1) of this Highlight #82], which begins *We will eventually get to a dual standard of information* ....., two categories of investigation are described. What is a *third* category that should be included?

- Suggest the relative *sizes* of the three categories in terms of the numbers of investigations they contain; briefly justify your answer from the information presented in the article EM9113.

4 Starting in the last paragraph on the first side (page HL82.1) of this Highlight #82, and also in statement *d* on the third side (page HL82.3), the article EM9113 discusses how polling was used to assess public reaction to Clarence Thomas's confirmation hearings. By means on an appropriate calculation, justify the statement ..... *as high as 12 per cent*; indicate clearly the *assumptions* which underlie your calculation.

- Comment critically on the *units* ('per cent') used in this statement.
  - Explain the basis of the statement (later in the same paragraph, at the top of page HL82.2) ..... *the opposite could also have been true* in terms of other information given in last paragraph on the first side (page HL82.2) of this Highlight #82.

5 In its second-last paragraph [on the third side (page HL82.3) of this Highlight #82], the article EM9113 describes journalists' responsibility for making proper assessments of the reliability of the Answer(s) of any investigation they report in the media. What is *not* mentioned in the article EM9113 but is clearly needed by journalists in making such assessments.

- Assuming that the articles reprinted in these Statistical Highlights are representative of those in Canadian newspapers which report the findings of 'research' investigations, indicate briefly how well you consider that journalists are meeting their statistical responsibilities.
  - Based on the experience of your studies at the University of Waterloo (or elsewhere), outline the step(s) that should be taken to bring about an *improvement* in the statistical quality of such media reports.

The article EM9113 in this Statistical Highlight #82 is also used in Figure 8.14 in the STAT 220 Course Materials and in Figure 2.7 in the Materials for STAT 332.

1995-04-20