



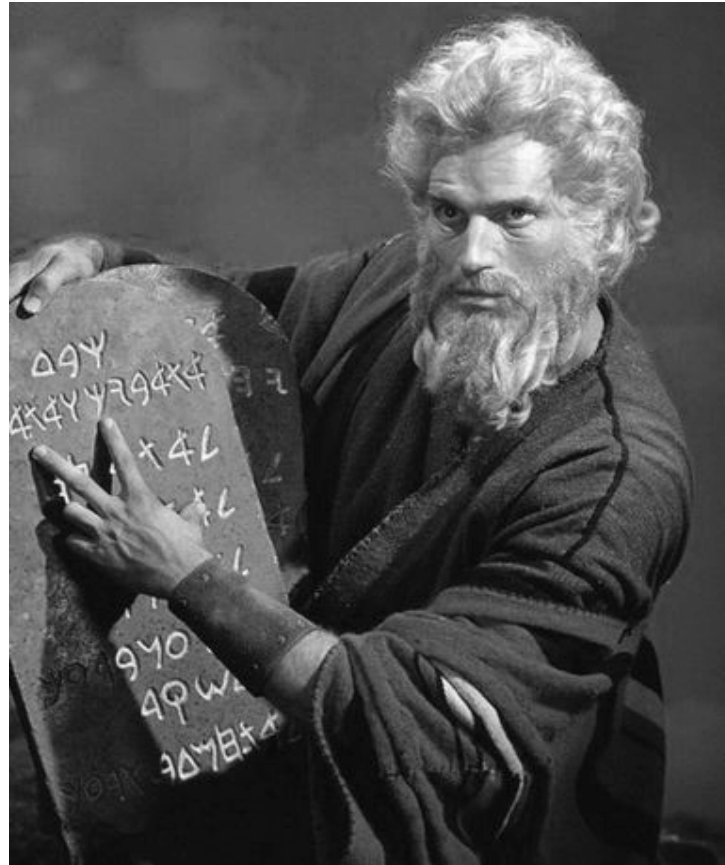
In Defense of Probability
(Has it come to this?)

Gary Langer
Langer Research Associates
glanger@langerresearch.com

American Association for Public Opinion Research
Orlando, Florida
May 18, 2012

From whence I come

- As well as a survey researcher and a small businessman, I'm a newsman by training and temperament.
- I'd like to give my clients the option of obtaining inexpensive data quickly.
- But I also only want to give them data and analysis in which they and I can be highly confident.
- I'm chiefly interested in estimating population values, in reliable time trends and in relationships among variables.
- I subscribe to the AAPOR Code, including the part that says we won't make unwarranted assertions about our data.
- I want to know not only whether a methodology apparently works (empirically), but how and why it works (theoretically).
- I have no stock options in probability sampling, or any other kind.





Good Data

- Are powerful and compelling
- Rise above anecdote
- Sustain precision
- Expand our knowledge, enrich our understanding, inform our judgments



Other Data

- Leave the discipline of inferential statistics and the generalizability to population values conferred by probability sampling
- Are increasingly prevalent; cheaply produced via the Internet, e-mail, social media
- Can misinform our judgment and misdirect our actions



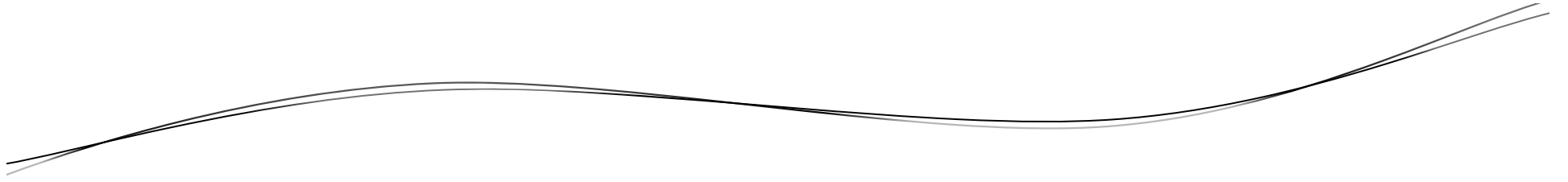
The Challenge

- Subject established methods to rigorous and ongoing evaluation
- Subject new methods to the same standards
- Value theoreticism as much as empiricism
- Consider fitness for purpose
- Go with the facts, not with the fashion



Good Data

- Are powerful and compelling
- Rise above anecdote
- Sustain precision
- Expand our knowledge, enrich our understanding, inform our judgments





Probability samples





This is not a new concept

A “criterion that a sample design should meet (at least if one is to make important decisions on the basis of the sample results) is that the reliability of the sample results should be susceptible of measurement. An essential feature of such sampling methods is that each element of the population being sampled ... has a chance of being included in the sample and, moreover, that that chance or probability is known.”

Hansen and Hauser, *POQ*, 1945

A shared view

- “...the stratified random sample is recognized by mathematical statisticians as the only practical device now available for securing a representative sample from human populations...” Snedecor, *Journal of Farm Economics*, 1939
- “It is obvious that the sample can be representative of the population only if all parts of the population have a chance of being sampled.” Tippett, *The Methods of Statistics*, 1952
- “If the sample is to be representative of the population from which it is drawn, the elements of the sample must have been chosen at random.” Johnson and Jackson, *Modern Statistical Methods*, 1959
- "Probability sampling is important for three reasons: (1) Its measurability leads to objective statistical inference, in contrast to the subjective inference from judgment sampling. (2) Like any scientific method, it permits cumulative improvement through the separation and objective appraisal of its sources of errors. (3) When simple methods fail, researchers turn to probability sampling... Kish, *Survey Sampling*, 1965

The new school



Internet Opt-Ins

THREATS AND RESPONSES: Blair and Bush, Still in Step

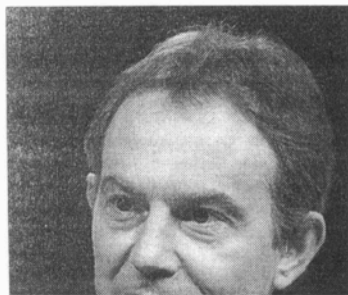
BRITAIN

Blair Pays a Price at Home For Supporting Bush on Iraq

By WARREN HOGE

LONDON, Jan. 26 — Prime Minister Tony Blair's moralizing diplomacy, his tough talk on Iraq and his steadfast loyalty to President Bush have gained him acceptance in the United States and an invitation to Washington this crucial week, but they have cost him popularity at home and hard-won influence in Europe.

In opinion surveys of the British public, a majority have been questioning his aggressive posture against Iraq and faulting him for being too subservient to the United States. Britons also increasingly accuse him of becoming distracted from the problems that affect their daily lives, like poor transportation, inattentive health services and rising



street crime.

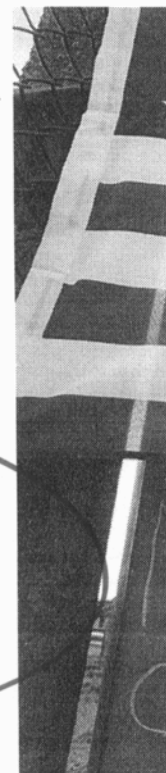
Mr. Blair, who has sent 26,000 troops to the Persian Gulf, reiterated his position in an interview with David Frost on the BBC this morning. He made arguments for the need to disarm Saddam Hussein and to limit additional inspection time before action might be taken to weeks rather than the months being sought by other European leaders.

In the hours that followed, the network conducted a call-in survey on whether viewers had been persuaded of the need to go to war, and 66 percent said no. A poll in The Sunday Times of London recorded 68 percent giving the same response to the question, "Has Tony Blair convinced you Saddam Hussein is sufficiently dangerous to justify war?"

Abroad, Mr. Blair's projection of Britain's power in the world through its alliance with America is undercutting his parallel desire to leverage the country's power in its own region.

"The Europeans do not see Britain as a trustworthy partner anymore," said Charles Grant, director of the London-based Center for European Reform. "Nobody wants to be seen holding hands with us in public — we're not kosher, we're not nice people to deal with."

Despite this damage to his popu-



Demonstrators

Blair has said that he has not acted on Iraq because he has not acted on Iraq. Blair has said that he has not acted on Iraq because he has not acted on Iraq.

An American official said Blair frequently administration put efforts to sway Blair. Blair has said that he has not acted on Iraq because he has not acted on Iraq. "For us, he is the



**Copyright 2003 Times Newspapers Limited
Sunday Times (London)**

January 26, 2003, Sunday

SECTION: Features; News; 12

LENGTH: 312 words

**HEADLINE: Blair fails to make a case for action that
convinces public**

BODY:

Tony Blair has failed to convince people in Britain of the need for war with Iraq, a poll for The Sunday Times shows. Even among Labour supporters, the prime minister has not yet made the case.

The poll of nearly 2,000 people by YouGov shows that only 26% say Blair has convinced them that Saddam Hussein is sufficiently dangerous to justify military action, against 68% who say he has not done so.

STATUS

NOT LOGGED IN

Login ▶

Voice your opinion
and get paid for it
REGISTER NOW

New registration ▶

How does it work? ▶

IRAQ
Latest
YouGov Data

UNDERSTANDING PEOPLE

INTERNET POLLING GETS TO THE HEART OF THE MATTER



TAKE P

About Us



Poll Reports



Open access polling

Services

Recent Reports

Music

Search



Survey Club

Established 2004

14,155,107 Survey Club Members and Counting!

Member Login

Email Password

Log

[Become a Member](#) | [Forget your ID?](#) | [Forget your Password?](#)

Join Today and Be Entered in Our \$1,000 Monthly Giveaway.

Grab Your FREE Membership!

Step 1 of 3

Email Password

Join

How it works...

Surveys
4 Cash

Product
Evaluation
Surveys

Focus
Groups





Greenfield ONLINE®



As a Greenfield Online member,
Your opinions count more!

More surveys - *Where you can express yourself*
More choice - *Of important topics that interest you*
More influence - *On the decisions of major companies*

The industry leader

Being part of the number one online survey community gives you access to surveys from the most influential players in the marketplace.

Quality matters

Because of our global focus on quality practices, companies trust the results of our surveys and pay more attention to your opinions.

Free membership

We will never charge you for membership. And as an added bonus, our members often receive cash and sweepstakes entries to thank them for making a difference.

Already a Member?

Existing members
login here

Email:

Password:

Remember me

[Forgot your Password?](#)

Not A Member?



Click below to
see if you qualify

[Click here if you live outside of the US](#)

-----Original Message-----

From: Students SM Team [<mailto:alumteam@teams.com>]

Sent: Wednesday, October 04, 2006 11:27 AM

Subject: New Job opening

Hi,

Going to school requires a serious commitment, but most students still need extra money for rent, food, gas, books, tuition, clothes, pleasure and a whole list of other things.

So what do you do? "Find some sort of work", but the problem is that many jobs are boring, have low pay and rigid/inflexible schedules. So you are in the middle of mid-terms and you need to study but you have to be at work, so your grades and education suffer at the expense of your "College Job".

Now you can do flexible work that fits your schedule! Our company and several nationwide companies want your help. We are looking to expand, by using independent workers we can do so without buying additional buildings and equipment. You can **START IMMEDIATELY!**

This type of work is Great for College and University Students who are seriously looking for extra income!

We have compiled and researched hundreds of research companies that are willing to pay you between \$5 and \$75 per hour simply to answer an online survey in the peacefulness of your own home. That's all there is to it, and there's no catch or gimmicks! We've put together the most reliable and reputable companies in the industry. Our list of research companies will allow you to earn \$5 to \$75 filling out surveys on the internet from home. One hour focus groups will earn you \$50 to \$150. It's as simple as that.

Our companies just want you to give them your opinion so that they can empower their own market research. Since your time is valuable, they are willing to pay you for it.

If you want to apply for the job position, please email at:
job2@alum.com Students SM Team



-----Original Message-----

From: Ipsos News Alerts [mailto:newsalerts@ipsos-na.com]

Sent: Friday, March 27, 2009 5:12 PM

To: Langer, Gary

Subject: **McLeansville Mother Wins a Car By Taking Surveys**

McLeansville Mother Wins a Car By Taking Surveys

Toronto, ON- McLeansville, NC native, Jennifer Gattis beats the odds and wins a car by answering online surveys. Gattis was one of over 105 300 North Americans eligible to win. Representatives from Ipsos i-Say, a leading online market research panel will be in Greensboro on Tuesday, March 31, 2009 to present Gattis with a 2009 Toyota Prius.

Access the full press release at:

<http://www.ipsos-na.com/news/pressrelease.cfm?id=4331>



- Opt-in online panelist
- 32-year-old Spanish-speaking female African-American physician residing in Billings, MT



HarrisPollOnlineSM Rewards

You are not currently logged in to your account.

[Login](#) | [Terms & Conditions](#) | [Privacy](#)

My HIpoints Account

What I Can Earn

Redeem My HIpoints

How HIpoints Work

HIpoints Support

How HIpoints Work

Because you are a valued Harris Poll Online panel member, we want to thank you for your continued support and participation in our surveys. This program is a fast and easy way to earn exciting rewards just by completing Harris Poll Online surveys. We will automatically create an account for you as soon as you complete an eligible survey. All you need to do is come to this site and login to activate your account.

Since we value your privacy, your personal information is not provided to third parties except for our agents who are working to mail your rewards. We have contracted with our agents to make sure your personal information remains private. Neither Harris Interactive, its agents nor any other third parties will use your personal information to send any special product offers from outside companies.

Choose What You Want From a Variety of Great Rewards

When you reach a redemption level, you will be able to choose a reward from a HIpoints Folio that includes a variety of merchandise and gift certificates. The reward that you choose will be sent directly to you or to the person designated on the delivery form **at absolutely no cost**.





Professional Respondents?

Among 10 largest opt-in panels:

- 10% of panel participants account for 81% of survey responses;
- 1% of participants account for 34% of responses.

Gian Fulgoni, chairman, comScore, Council of American Survey Research Organizations annual conference, Los Angeles, October 2006.



Questions...

- Who joins the club, how and why?
- What verification and validation of respondent identities are undertaken?
- What logical and QC checks (duration, patterning, data quality) are applied?
- What weights are applied, and how? On what theoretical basis and with what effect?
- What level of disclosure is provided?
- What claims are made about the data, and how are they justified?

One claim: Convenience Sample MOE

- Zogby Interactive: "The margin of error is +/- 0.6 percentage points."
- Ipsos/Reuters: "The margin of error is plus or minus 3.1 percentage points."
- Kelton Research: "The survey results indicate a margin of error of +/- 3.1 percent at a 95 percent confidence level."
- Economist/YouGov/Polimetrix: "Margin of error: +/- 4%."
- PNC/HNW/Harris Interactive: "Findings are significant at the 95 percent confidence level with a margin of error of +/- 2.5 percent."
- Radio One/Yankelovich: "Margin of error: +/-2 percentage points."
- Citi Credit-ED/Synovate: "The margin of error is +/- 3.0 percentage points."
- Spectrem: "The data have a margin of error of plus or minus 6.2 percentage points."
- Luntz: "+3.5% margin of error"

YouGov **Polimetrix** What the world thinks

HOME

COMPANY

SERVICES

NEWS

CAREERS

CONTACT



Online Survey

Traditional phone-based survey techniques suffer from deteriorating response rates and escalating costs. YouGovPolimetrix combines technology infrastructure for data collection, integration with large scale databases, and novel instrumentation to deliver new capabilities for polling and survey research.

[\(read more\)](#)



RRs

- Surveys based in probability sampling cannot achieve pure probability (e.g., 100% RR)
- Pew (5/15/12) reports decline in its RRs from 36% in 1997 to 9% now.
- Does this trend poison the well?



A Look at the Lit

Keeter et al., *POQ*, 2006:
Comp. RR 25 vs. 50

"...little to suggest that unit nonresponse within the range of response rates obtained seriously threatens the quality of survey estimates."

(See also Keeter et al., *POQ*, 2000, RR 36 vs. 61)



And the latest

Pew 5/15/12

Comp. RR 9 vs. 22

“Samples Still Representative

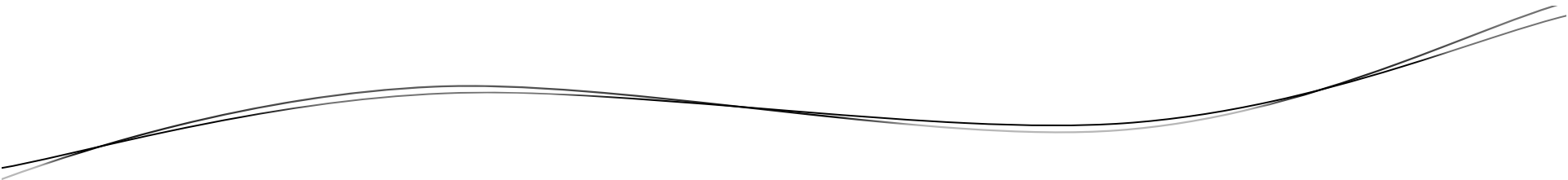
“...despite the growing difficulties in obtaining a high level of participation in most surveys, well-designed telephone polls that include landlines and cell phones reach a cross-section of adults that mirrors the American public, both demographically and in many social behaviors.

“Overall, there are only modest differences in responses between the standard and high-effort surveys. Similar to 1997 and 2003, the additional time and effort to encourage cooperation in the high-effort survey does not lead to significantly different estimates on most questions.”



Holbrook et al., 2008
(in “Advances in Telephone Survey Methodology”)

“Nearly all research focused on substantive variables has concluded that response rates are unrelated to or only weakly related to the distributions of substantive responses (e.g. O’Neil, 1979; Smith, 1984; Merkle et al., 1993; Curtin et al., 2000; Keeter et al., 2000; Groves et al., 2004; Curtin et al., 2005).”



Holbrook et al.: Demographic comparison, 81 RDD surveys, 1996-2005; AAPOR₃ RRs from 5% to 54%

“In general population RDD telephone surveys, lower response rates do not notably reduce the quality of survey demographic estimates. ... This evidence challenges the assumption that response rates are a key indicator of survey data quality...”

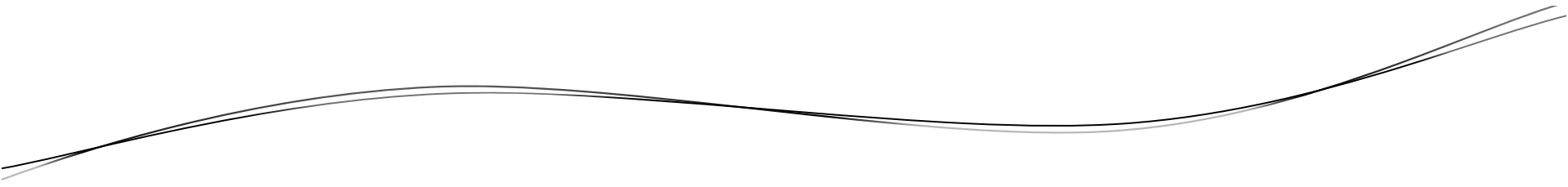


Groves, *POQ*, 2006

“Hence, there is little empirical support for the notion that low response rate surveys de facto produce estimates with high nonresponse bias.”

**GREENFIELD ONLINE'S OFFERING
\$4 PER COMPLETE ON SAMPLE
50% INCIDENCE OR HIGHER!**



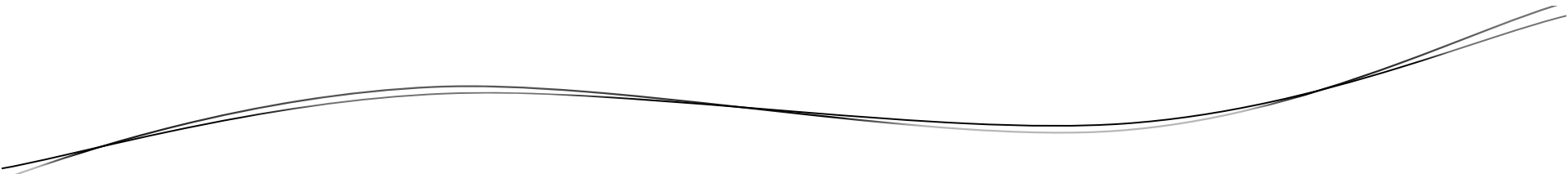


“Inside Research” estimates that spending on online market research will reach \$2.05 billion in the United States and \$4.45 billion globally this year, and that data gathered online will account for nearly half of all survey research spending.

I asked Laurence Gold, the magazine’s editor and publisher, what the industry is thinking about in its use of non-probability samples.

“The industry is thinking fast and cheap,” he said.

<http://blogs.abcnews.com/thenumbers/2009/09/study-finds-trouble-for-internet-surveys.html>



Not Everyone (the Dedeker rocket)

“I’ve been talking recently about the importance of high-quality research to support business decision-making. It’s vitally important for P&G. ... The area I feel is in greatest need of help is representative samples. I mention online research because I believe it is a primary driver behind the lack of representation in online testing. Two of the biggest issues are the samples do not accurately represent the market, and professional respondents.”

Kim Dedeker

Head of Consumer & Market Knowledge, Procter & Gamble

Research Business Report, October 2006

Enter Yeager, Krosnick, et al., 2009

(See also Malhotra and Krosnick, 2006; Pasek and Krosnick, 2010)

- Paper compares seven opt-in online convenience-sample surveys with two probability sample surveys
- Probability-sample surveys were “consistently highly accurate”
- Opt-in online surveys were “always less accurate... and less consistent in their level of accuracy.”
- Weighted probability samples sig. diff. from benchmarks 31 or 36 percent of the time for probability samples, 62 to 77 percent of the time for convenience samples.
- Opt-in data’s highest single error vs. benchmarks, 19 points; average error, 9.9.



ARF FOQ via Reg Baker, 2009

- Reported data on estimates of smoking prevalence: similar across three probability methods, but with as many as 14 points of variation across 17 opt-in online panels.
- “In the end, the results we get for any given study are highly dependent (and mostly unpredictable) on the panel we use. This is not good news.”

AAPOR's "Report on Online Panels," April 2010

- “Researchers should avoid nonprobability online panels when one of the research objectives is to accurately estimate population values.”
- “The nonprobability character of volunteer online panels ... violates the underlying principles of probability theory.”
- “Empirical evaluations of online panels abroad and in the U.S. leave no doubt that those who choose to join online panels differ in important and nonignorable ways from those who do not.”
- “In sum, the existing body of evidence shows that online surveys with nonprobability panels elicit systematically different results than probability sample surveys in a wide variety of attitudes and behaviors.”
- “The reporting of a margin of sampling error associated with an opt-in sample is misleading.”



AAPOR's conclusion

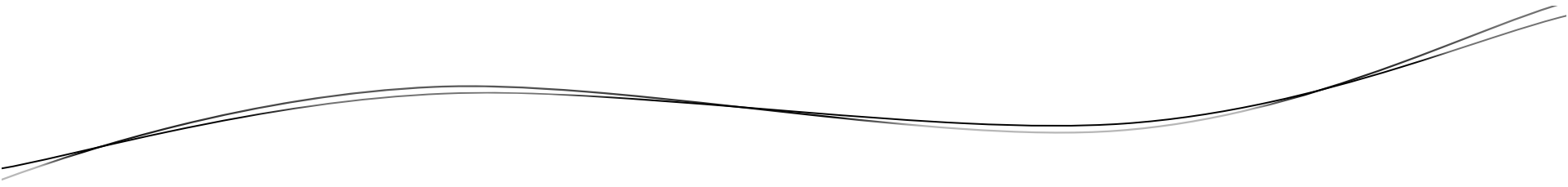
“There currently is no generally accepted theoretical basis from which to claim that survey results using samples from nonprobability online panels are projectable to the general population. Thus, claims of ‘representativeness’ should be avoided when using these sample sources.”

AAPOR Report on Online Panels, April 2010



Horse race polls

- Are a poor indicator of data quality
- They represent modeling to estimate a single specific outcome, not sampling for broader generalizability
- The modeling elements are ill-disclosed; e.g., are non-probability samples weighted to probability samples?
- Unmodeled comparisons are far more useful, e.g., comparison of unweighted as well as weighted data to benchmarks



OK, but apart from population values, we can still use convenience samples to evaluate relationships among variables.

Right?



Pasek and Krosnick, 2010

- Comparison of opt-in online and RDD surveys sponsored by the U.S. Census Bureau assessing intent to fill out the Census.
- “The telephone samples were more demographically representative of the nation’s population than were the Internet samples, even after post-stratification.”
- “The distributions of opinions and behaviors were often significantly and substantially different across the two data streams. Thus, research conclusions would often be different.”
- Instances “where the two data streams told very different stories about change over time ... over-time trends in one line did not meaningfully covary with over-time trends in the other line.”



Pasek/Krosnick conclusion

“This investigation revealed systematic and often sizable differences between probability sample telephone data and non-probability Internet data in terms of demographic representativeness of the samples, the proportion of respondents reporting various opinions and behaviors, the predictors of intent to complete the Census form and actual completion of the form, changes over time in responses, and relations between variables.”

Pasek and Krosnick, 2010



This is *really* not a new concept

“Diagoras, surnamed the Atheist, once paid a visit to Samothrace, and a friend of his addressed him thus: ‘You believe that the gods have no interest in human welfare. Please observe these countless painted tablets; they show how many persons have withstood the rage of the tempest and safely reached the haven because they made vows to the gods.’

“ ‘Quite so,’ Diagoras answered, ‘but where are the tablets of those who suffered shipwreck and perished in the deep?’ ”

“*On the Nature of the Gods*,” Marcus Tullius Cicero , 45 B.C.

(cited by Kruskal and Mosteller, *International Statistical Review*, 1979)



Can non-probability samples be ‘fixed?’

- Bayesian analysis?
 - What variables? How derived?
- Sample balancing?
 - “The microcosm idea will rarely work in a complicated social problem because we always have additional variables that may have important consequences for the outcome.” (See handout)

Gilbert, Light and Mosteller, Statistics and Public Policy, 1977

The parking lot...

“The Chairman” is said to have asked his researcher whether an assessment of a parking lot reflects “a truly random sample of modern society.”

Maybe not, the researcher replied, “ ‘but we did the best we could. We generated a selection list using a table of random numbers and a set of automobile ownership probabilities as a surrogate for socio-economic class. Then we introduced five racial categories, and an equal male-female split. We get a stochastic sample that way, with a kind of ‘Roman cube’ experimental protocol in a three-parameter space.’ “

“It sounds complicated,” said the Chairman.

“Oh, no. The only real trouble we’ve had was when we had to find an Amerindian woman driving a Cadillac.”

Hyde, *The Chronicle of Higher Education*, 1976

cited by Kruskal and Mosteller, *International Statistical Review*, 1970



Some say you can

“The survey was administered by YouGovPolimetrix during July 16-July 26, 2008. YouGovPolimetrix employs sample matching techniques to build representative web samples through its pool of opt-in respondents (see Rivers 2008). Studies that use representative samples yielded this way find that their quality meets, and sometimes exceeds, the quality of samples yielded through more traditional survey techniques.”

Perez, *Political Behavior*, 2010



AAPOR's task force says you can't

“There currently is no generally accepted theoretical basis from which to claim that survey results using samples from nonprobability online panels are projectable to the general population. Thus, claims of ‘representativeness’ should be avoided when using these sample sources.”

AAPOR Report on Online Panels, 2010



It has company

- “Unfortunately, convenience samples are often used inappropriately as the basis for inference to some larger population.”
- “...unlike random samples, purposive samples contain no information on how close the sample estimate is to the true value of the population parameter.”
- “Quota sampling suffers from essentially the same limitations as convenience, judgment, and purposive sampling (i.e., it has no probabilistic basis for statistical inference).”
- “Some variations of quota sampling contain elements of random sampling but are still not statistically valid methods.”

Biemer and Lyberg, *Introduction to Survey Quality*, 2003



And the Future?

In probability sampling:

- Continued study of response-rate effects
- Concerted efforts to maintain response rates
- Renewed focus on other data-quality issues in probability samples, e.g. coverage
- Development of probability-based alternatives, e.g. mixed-mode, ABS



The Future, cont.

In convenience sampling:

- Continued study of appropriate uses (as well as inappropriate misuses) of convenience-sample data
- Continued evaluation of well-disclosed, emerging techniques in convenience sampling
- The quest for an online sampling frame



Thank you!

**Gary Langer
Langer Research Associates
glanger@langerresearch.com**

**American Association for Public Opinion Research
Orlando, Florida
May 18, 2012**

Addendum: Langer handout

“People often suggest that we not take random samples but that we build a small replica of the population, one that will behave like it and thus represent it. ... When we sample from a population, we would like ideally a sample that is a microcosm or replica or mirror of the target population – the population we want to represent. For example, for a study of types of families, we might note that there are adults who are single, married, widowed, and divorced. We want to stratify our population to take proper account of these four groups and include members from each in the sample. ... Let us push this example a bit further. Do we want also to take sex of individuals into account? Perhaps, and so we should also stratify on sex. How about size of immediate family (number of children: zero, one, two, three...) – should we not have each family size represented in the study? And region of the country, a size and type of city, and occupation of head of household, and education, and income, and... The number of these important variables is rising very rapidly, and worse yet, the number of categories rises even faster. Let us count them. We have four marital statuses, two sexes, say five categories for size of immediate family (by pooling four or over), say four regions of the country, and six sizes and types of city, say five occupation groups, four levels of education, and three levels of income. This gives us in all $4 \times 2 \times 5 \times 4 \times 6 \times 5 \times 4 \times 3 = 57,600$ possible types, if we are to mirror the population or have a small microcosm; and one observation per cell may be far from adequate. We thus may need hundreds of thousands of cases! ... We cannot have a microcosm in most problems. ... The reason is not that stratification doesn't work. Rather it is because we do not have generally a closed system with a few variables (known to affect the responses) having a few levels each, with every individual in a cell being identical. To illustrate, in a grocery store we can think of size versus contents, where size is 1 pound or 5 pounds and contents are a brand of salt or a brand of sugar. Then in a given store we would expect four kinds of packages, and the variation of the packages within a cell might be negligible compared to the differences in size or contents between cells. But in social programs there are always many more variables and so there is not a fixed small number of cells. The microcosm idea will rarely work in a complicated social problem because we always have additional variables that may have important consequences for the outcome.”