

How can one person completely change the results of a survey?

October 21 2016, by



Surveys are challenging to design and execute, and there's no shortage of details on the nuances of how to do them well. Credit: Flickr

In public health, we rely heavily on samples, as measuring everyone you are interested in is often impractical. However, this requires a lot of thought and development in order to avoid unintentionally biasing your sample, as was the case for the USC Dornsife/LA Times Daybreak poll.

Last week, a story came out about how one [19-year old black man in Illinois](#) was single-handedly changing the standings of the US Presidential election. This was based on the results from the USC Dornsife/Los Angeles Times Daybreak, a survey of voter attitudes on "a

[wide range of political, policy, social and cultural issues.](#)" In this survey, Donald Trump has generally held the lead, until last week, when Hillary Clinton came out in front. Interestingly, this is markedly different than most other national polls, that have shown Clinton is [generally ahead](#), or a much closer contest than that poll would have you believe. So what happened?

Surveys give people different "weights"

When designing a survey, one question researchers have to ask is how many people they want to capture. If they can capture everyone of interest, then they've conducted a census – and everything is good. However, if they can't, due to either cost, practicality, or a host of other reasons, then a sample is chosen.

A popular way to pick a sample is at random. If you have a list of 50 names, you might pick 10 at random. Another way is to take a systematic random sample, such as every fifth person who registers at a medical clinic. When you do this though, you have to apply what are called "sample weights" to your survey, which adjusts your survey sample to be representative of the larger population. So if your sample had 25% women, you might weight women as 2, while weighting men as 0.67 in order to make the final ratio 50-50 ($25\% \times 2 = 75\% \times 0.67$). You could do the same for age as well, to make sure you have a representative number of women aged 30-39 (for example).

Is this a problem?

Not at all! This is fine, especially when the person you're weighting heavily is representative of the larger group you're interested in. It's standard practice, with national surveys such as NHANES in the US using sample weights to ensure their results are generalizable to the

national population. However, in the case of the USC Dornsife/LA Times poll, they had one person who was not representative of the larger group, and may have been responsible for (some of) the differences observed between this poll and others. This is where this survey ran into problems as it weighted for small groups. What that means is that if you don't have enough people in that group, they start being weighted more and more heavily. If you only have one person in that group (as an extreme example), they could be weighted many times more than someone in the next group, skewing the results. To quote the NY Times:

Our Trump-supporting friend in Illinois is a surprisingly big part of the reason. In some polls, he's weighted as much as 30 times more than the average respondent, and as much as 300 times more than the least-weighted respondent (emphasis mine).

Another way of looking at this is to consider if people have to pick between their favourite superhero between Batman and Superman. If this guy picks Batman, then there have to be between 30 to 300 people who pick Superman for the results to seem "even." Now that's fine, if the group that guy belongs to generally picks Batman. But if he's an outlier, and most people in his group pick Superman, then suddenly your results are very different from what you're trying to measure. This was further exacerbated by the LA Times poll is not a random sample every time, but surveying the same people repeatedly. This means that there's no opportunity for your outlier to fall out of the sample – they will contribute every time to the results, skewing the results every time you poll them.

How did the LA Times respond?

The [LA Times responded](#) to the NY Times piece, and had some very interesting, and sound, arguments for weighting groups as heavily as they did, and for trying to be representative. One issue is that you need to

make sure all ethnic groups are represented in your polls, especially when that is very highly tied to what you're trying to measure (e.g., who you will vote for). The LA Times acknowledged that this was what they were trying to accomplish, but unfortunately the one outlier they had pushed the support of Trump up among African Americans. Their argument was that this would increase their margin of error, and that this was ignored "in order to make a political point, but there's not much we can do about that."

Transparency and different choices

Now one thing I want to point out is that, while they made some interesting decisions, the survey was very transparent. In fact, anyone is able to look at their methodology here. So while there are decisions they made that I disagree with, they are willing to allow people the chance to review them, which deserves credit. Nate Silver of FiveThirtyEight has an interesting take on this:

You can [almost always find something "wrong" with a poll you don't like](#), even if you might have approved of its methodology before you saw its result.

It's probably also harmful for the profession as a whole when poll-watchers are constantly trying to browbeat "outlier" polls into submission. That can encourage [herding](#)—pollsters rallying around a narrow consensus to avoid sticking out—which is bad news, since herding reduces the benefit of averaging polls and [makes them less accurate overall](#).

Polls will all be different. The decisions made by the pollsters, the people who respond to the poll, as well as random fluctuations, will ensure that polls will all be slightly different. However, these decisions can be useful – as Silver points out above, sometimes the outliers are the

ones that have taken something into account that everyone else ignored. However, we still need to think carefully when designing surveys to make sure that we're actually capturing what we intend to, with both accuracy and precision.

This story is republished courtesy of PLOS Blogs: blogs.plos.org.

Provided by PLOS Blogs

Citation: How can one person completely change the results of a survey? (2016, October 21) retrieved 17 July 2024 from <https://phys.org/news/2016-10-person-results-survey.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.