Martin O'Shannessy, OmniPoll:

Were polls accurate?

It's not yet possible to tell whether polls were accurate in 2020 as all the votes are yet to be counted.

What has been the issue with polls that are wrong?

Pending anything more concrete...

The main issue in 2016 and ongoing, has been reliance on Online polls which are unrepresentative because they exclude 98% of voters.

This is because no online panel has more than about 2% of the adult population in it – 98% do not have a chance of participating in the poll because they are not on the panel.

The most reliable polls were the national polls that were conducted on the phone and they have compared favourable to national results in recent years.

These findings and others I will discuss below were made by AAPOR - the US national polling industry body when they looked into the failure of the 2016 polls.

How to do it right and why it often is not

As *states* elect the electoral colleges, state based polls are the only way to predict the presidential outcome.

To do a good job, one would need to conduct a reliable, live operator telephone poll using mobiles and land lines of (say) 1000 interviews in each of the 50 states shortly before the election.

This is very expensive and has limited news value as it needs to be done just a few days out.

So we observe that state polls tend to be done cheaply and erratically because of the sheer cost of doing so many while national polls can be done at a good quality level because only one at a time is required.

The result of this problem in 2016 was that the national polls were done well more often and more often pointed to a Democrat win of the national popular vote – which is what happened.

However, they were irrelevant to the outcome because they did not predict state results.

At the same time, the state level polls were numerous but rarely of good quality (8/10 were done online and therefore were not representative) and generally failed to point to a Trump win.

They were the polls that pollsters needed to get right but they were often wrong.

The amalgamators – Real Clear Politics, Nate Silver etc. sucked up all the polls good and bad and the bad being more numerous, skewed their results towards the wrong answer.

Is failure to answer mobile phones and issue?

People who posit this apparently do not know the difference between *failing to have a representative sample* and having a *high non-response rate.*

These two issue are quite different.

What is representative?

To be representative, a poll must (should) give *every potential respondent a chance to participate in the survey.*

Everybody (practically) has a mobile phone so using a good list of mobile phones is a good way to try to make your survey representative by including everybody.

Comparing live operator mobile phone to online, we see that just about 95% of all respondents can be theoretically reached in a given time frame by phone, so this method is more likely to be representative as *virtually nobody is excluded* by the method.

Compare this to Online polls where only 2% of the adult population is on any typical panel. This method *excludes 98% of the population* and this therefore very unlikely to be representative.

What is response rate?

This is the proportion who answer your survey – high response rates are desirable but poor response rate to a representative survey is more likely to result in an accurate result than a good response rate to an un-representative one.

We have found in our Australian work that we need to call about 10 - 20 people to get one interview on the mobile. About half of these people we *don't* interview are excluded by us as we already filled their age/sex quota.

But these people did have a chance to be selected and would have been interviewed if we had called their number the day before. We randomise numbers to allow this dynamic to happen.

The other half who do not participate do so for their own reasons but statistically, the issue is that they had a *chance* to be interviewed making the survey more likely to be representative.

So, right now, using real people as interviewers and a combination of mobile and landline, we are seeing response rates of about 30%.

Our inclusion rate – the proportion of the population that was eligible sits above 90%.

As non-response goes up, we are concerned about why and what impact it has but this is a much smaller issue than starting with a non-representative method such as panel – or pushing non-response beyond 90% by using robo-dialling.

To back this up

In 2013, I conducted the final national Newspoll using a mix of landline and mobiles. Our response rate overall was about 20% after exclusions and refusals.

The absolute error between my poll and the national two party result was less than 0.5%

Here is a comparison of the last National Federal Newspoll I conducted and the actual result.

REF DATE	POLL DATE	TYPE			Primary					
					COALITION	ALP		GREENS		OTHERS
05-Sep-13	Newspoll 3-5 September 2013	Newspoll			46	33		9		12
07-Sep-13	**Election 7 September 2013**	**Election**			**45.6**	**33.3**		**8.7**		**12.4**
				Two Party						
REF DATE	POLL DATE	TYPE	COALITION	LABOR						
05-Sep-13	Newspoll 3-5 September 2013#	Newspoll	54	46						
07-Sep-13	**Election 7 September 2013**	**Election**	**53.5**	**46.5**						

So – it's possible but expensive to do it right.

And most people I hear commenting mix up basic concepts or have no concept of what is actually happening in a survey.