

Research

AUSTRALIAN MARKET & SOCIAL RESEARCH SOCIETY | VOLUME 36 | NUMBER 3 | AUG - OCT 2019

NEWS

IMPACTFUL INSIGHTS

PAGE 8
CELEBRATING HUMANITY
THROUGH IMMERSION

PAGE 10
CELEBRATING ALL KINDS
OF BEING ACTIVE

PAGE 14
MARGIN FOR ERROR:
2019 ELECTION POLLING

PAGE 19
THE ART TO
RESEARCH SPARK

PAGE 22
FIVE TECHNIQUES FOR
VISUAL STORYTELLING

The official magazine for members of the
Australian Market & Social Research Society.

AMRS 

Margin for Error: 2019 Election Polling

The failure of the published polls to predict the 2019 election outcome could have been avoided, and is a wake-up call to the wider industry to think more deeply about the way it measures opinion and behaviour, writes Jim Reed.

They say that the first step in solving a problem is to admit to having one. Well, we have a problem.

Australia's published political polls got it wrong at the recent Federal election.

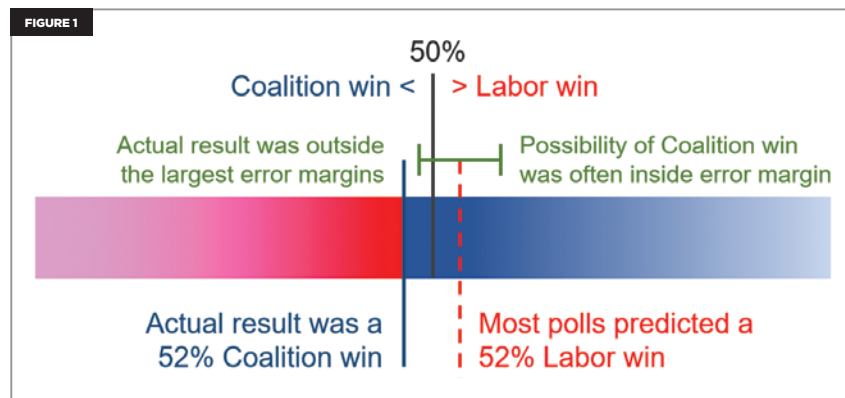
Full stop. All major polls predicted a swing to Labor to deliver them government at a two-party preferred result of around 52%, whereas the reality was a swing to the Coalition to retain power.

I say 'we' have a problem because this should be important to us all.

It does not matter if you are in the majority who conduct or commission research on other topics. These published polls use the same methods, the same sample sources, the same question

styles, and are arguably at the forefront of making sure error does not creep in. If they are getting it wrong, and systemically so, the chances are that others are too.

And both your clients and respondents are watching. Most of our research is not published or immediately tested against reality in the same way, but if the public-facing



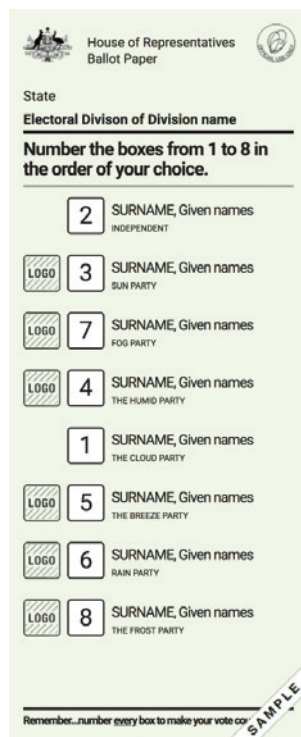
part of our industry is seen to be getting it wrong it affects us all. Why commission, listen to or take part in research if it is perceived as nonsense?

We need to fix this. Now.

Arguably, we should have seen this coming. Pollsters in other jurisdictions with trickier voting systems, like the UK and US, have already experienced this, we have seen closer than expected results in recent state elections, and we have seen it when asked to leave the comfort zone of our compulsory voting system to predict the same-sex marriage vote.

We have been lulled into a false sense of security by our long track record of accuracy, so we need to take a look at what has changed in our methods, the electorate and in the voting system. What might have gone wrong and what lessons can we learn?

FIGURE 2 - BALLOT PAPER GENERIC



Good research is essentially asking the right people the right things in the right way – the ‘who’, ‘what’ and ‘how’ of research – then analyzing and presenting the findings of those questions properly. Let’s look at how Australia’s published polling stacks up on these measures.

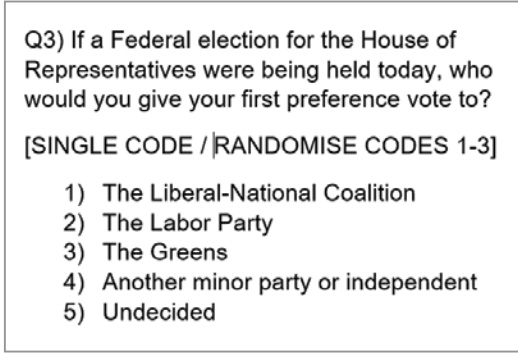
Asking the Right Things: Question Design

The questions asked by pollsters are probably the area to have changed least over the years simply because our voting system has remained fairly constant. Each polling company has its own particular ‘house style’ when asking and recording vote, but in truth they are all quite similar in their approach to asking people’s first preference vote on that day.

In doing so they are attempting to replicate as far as possible the real-world behaviour of casting a ballot because the way in which something is asked and recorded influences the result. Arguably, they do not get close enough.

For example, the actual ballot paper lists a series of candidates and their parties in a particular order specific to an electorate, voters are asked to

FIGURE 3 - SURVEY QUESTION



number every square in order of their preference, and they are obliged to do so lest their vote becomes ‘informal’.

Most polls fail to emulate this process. They name and record only some of the parties (including combining the Liberals and Nationals), they often present those parties in random order or have the major parties first, they name only the parties and not the candidates and so dilute any incumbency advantage, and they provide the option of ‘don’t know’ when this is never an option. Each may have an effect, but the latter is the most important.

Most of the final polls at the 2019 election stated that they had excluded 5-10% of their sample who remained undecided, and if the majority of them finally decided to keep with the status quo this could explain the election result on its own. The evidence suggests that there was no ‘late swing’ from one party to another, but this would be a failure to capture which way undecided voters were going.

Where the voting system has changed is in the growth in early voting, where electors may cast their ballot many days in advance of polling day in person or by post. When a third of people may have already cast a vote before the final polls are taken, what is the sense in asking them how they would vote on that day? I am yet to see a pollster display the results or questions of early votes, so we must assume that this ‘part exit poll’ was not factored in.

The other factor at play here is whether respondents have started lying to our questions or, rather, have begun ‘gaming the system’ to send a message of protest to an incumbent to do better, before then reverting to their actual

© OVER PAGE

vote. It is nigh impossible to detect a lie in survey research, especially when the protest is real.

However, there are means by which to guard against them. You can use interviewers to determine the possibility, you can use parallel qualitative research to understand the motivations, you can use supporting diagnostic questions to identify those likely to change, and you can mask the source of the survey so that it is not seen to be a public message respondents are sending. All of this applies to non-political research too.

But we are assuming here that vote is the right question to ask. This attempts to measure a behaviour, and one often months or years away. It tends not to move around greatly as even when there is a dislike for a party, leader or policy, a certain proportion of their 'base' will vote for them regardless. If we are really interested in measuring opinion, we might ask a different question, like how the government and opposition are performing.

Much of what is outlined here could have been addressed by deeper thought

and proper pilot testing of questions. There is a worrying tendency to think that a 'soft launch' and review of data is enough, but this does not tell you whether your questions are understood, whether the response categories are appropriate or whether they are measuring what you want them to.

There is no proxy for proper testing with your target respondents and industry peers.

Asking the Right People: Sampling Design

In sampling, we are looking to ensure that the right people are selected for the study, and that those people selected and responding are a decent approximation of the population at large.

Screening is where question design crosses over with achieved sample, as it makes sure that who we are asking to take part should take part. In my experience, too many studies default to 'general population' when they should be looking at the real decision-makers. This habit results in a watered-down or warped view of reality.

In the case of political polling, it may surprise many to know that only around three-quarters of Australia's adult population casts a vote after non-citizens, those not enrolled, those not turning out and those casting an informal vote are omitted. To put that in context, it's like the population of Victoria not voting.

In effect, this means that Australia's system is more akin to voluntary voting systems in other countries than we think, and holds the same uncertainties this brings into the system. Pollsters should filter these people from the equation if they are asking vote, but it appears from stated methodologies that this is not always the case.

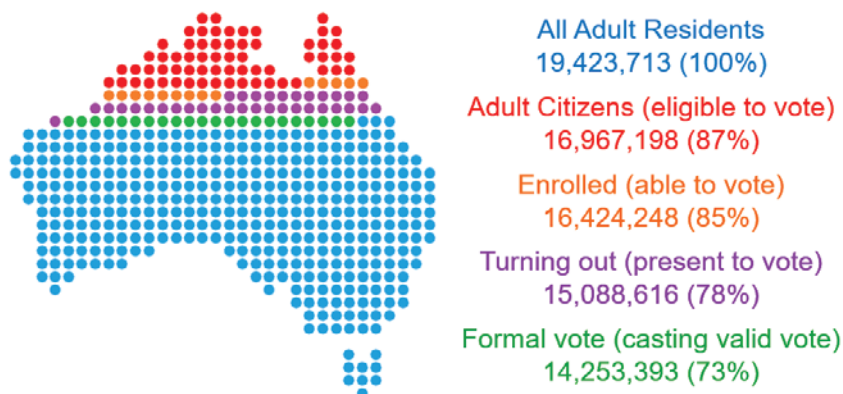
Including the theoretical voting intention of a quarter of the population that do not vote (including younger people who enrolled for the same-sex marriage vote, but then did not vote in the election) could easily account for a few percentage points of error.

This is before we get to the issues of who to sample before screening and how to source them.

Australia's voting system is not just about national opinion. Rather than a nationwide census, it is actually 151 surveys of electorates where the winner is predicted by a majority of leads in those smaller surveys. For example, in the 1998 election John Howard's Coalition won government quite comfortably with a two-party preferred vote of 49%. That is, overall public opinion was against him and he still won a majority of seats.

The result is already known in many of those electorates, so it is only the closer seats – the 'marginals' – that matter. In such a case it would make sense to include only marginal seats in polling, but our published polls do not even sample or breakdown results at this level. Some do publish individual

FIGURE 4



seat polls, but these have proven even more prone to inaccuracy.

Regardless of whether a sample is national or local, we are now finding greater difficulty in reaching and interviewing a representative group of people. This is a major cause for concern that affects us all.

Response rates are one thing that can lead to bias, with some pollsters suggesting that Greens voters, for example, are more likely to take part and give us their views. This requires its own solutions, such as industry-wide reputation building and incentives to encourage response.

But before we get to that stage, we must ensure that we have a representative group to approach in the first place. This was achievable in the past, where pollsters had the time, budget and physical access to randomly sample households to door-knock or make landline calls to, but those days have passed. In the absence of greater investment or access to the IPND – the Government’s complete list of all phone numbers – we are reduced to other avenues to reach people.

Asking the Right Way: Data Collection Method

Household and landline phone surveys are now not optimal, and where used may over-sample older people who tend to be right-leaning, so we have instead seen growth in on-line, social media, automated robo calls and mobile phone (call and SMS) surveys.

Whilst many of these newer methods are cost-effective and acceptable for many polling applications, each has its own pros and cons that must be thought through.

For example, cost-effective on-line polls are best at reaching people in a

way and at a time convenient to them. One might also argue that they better emulate the private voting decision of writing in a voting booth. Indeed, I used it in the Marriage Equality research to avoid social bias, but vote decisions are less sensitive (there was certainly no

Including the theoretical voting intention of a quarter of the population that do not vote (including younger people who enrolled for the same-sex marriage vote, but then did not vote in the election) could easily account for a few percentage points of error.

‘shy tory’ effect at work in 2019).

On-line research is a go-to method for the industry now, but it is too often forgotten that on-line surveys rely on panels of willingly recruited respondents completing surveys for an incentive. They are not and never will be a reliable means to reach the entire population afresh, they are self-selecting sub-samples with their own response bias and so margins of error simply do not apply.

They also tend to be skewed to younger, less affluent and left-leaning segments, can have panel members who are more engaged and vocal on current affairs, and some suggest that regular exposure repeat tracking studies can lead to morphed views. Social media surveys and qualitative recruitment lists suffer similar disadvantages, and must be carefully managed.

Mobile phone surveys are probably the best solution to reaching a majority, but they are more expensive given the use of interviewers. Cheaper robo and SMS telephone polling has grown significantly in recent years, but their interview length is limited, the quality of their samples can vary greatly, and some argue that they lead to lower response rates across the industry through sheer annoyance and a lack of incentive.

In an attempt to reach a greater variety of people cost-effectively, some polling companies have begun weaving these methods together. This definitely has a place in the absence of IPND access, but the optimal mix of these efforts is a point of debate and constant refinement.

Though difficult to pin down, the choice of method and its associated sample source could easily account for the errors we are seeing.

Data Treatment, Analysis & Presentation

As with any research project, the dilemma of how to best treat, analyse and present the findings also creates challenges in published polling.

The difficulties of sampling and interviewing a representative sample place greater emphasis on weighting: how each geo-demographic and lifestyle group is emphasised in the results. This is standard practice in general population research as applying weighting factors (and quotas) iron out any obvious skews in a sample, but should employ interlocking weights and be minimal.

Published political polls often refer to their samples being weighted, but there is no frame of reference for what

© OVER PAGE

the three-quarters voting look like, so what are pollsters weighting to? If it is general ABS data for adults they risk weighting accurate results into inaccurate ones of the kind we are seeing. Far better would be to include those screened out into the weighting decision.

Much has been made of the fact that numerous polls arrived at a very similar two-party preferred results in the final days of the election campaign, and that there may have been a ‘herding’ effect pushing them to agree through the use of weighting. I think this unlikely as the polls did report differing primary vote figures, with the application of similar preference flows from the last election pushing them closer together.

On the assumption that this and other aspects of basic data checking, treatment and analysis have been performed correctly and checked, we turn our attention to how the findings are presented.

The pollster does not always control the commentary of course. This is as true of research findings presented by our clients to their stakeholders as it is to the media’s use of published polling data, but there are obvious ways in which to safeguard against misunderstanding or misuse.

Results should be wrapped up with a clear and unequivocal commentary, accompanied by a verbal debrief and opportunity for questions from your client. This ensures that the leeway for misinterpretation is minimal.

And we should be clear about what research represents. Whilst most of the polls got the election result wrong outside their error margin, the chance of a win for the Coalition was within their error margin. (See Figure 1.) This, coupled with the 5-10% undecided, should have led to a clear caveat that the result was ‘too close to call’, but

The 2019 election showed that predictions are dangerous things, but if we can deliver quality, I confidently predict a bright future for research.

that call was absent from the reporting of the polls.

At the very least, the error margin and exclusion of undecided voters should have been more prominent. No matter how well-designed research is, all results have uncertainties. Omitting such details makes your reporting of them a ‘promised prediction’. This is never true of a survey’s snapshot in time, and it is certainly not true in qualitative research.

Quality, not (just) Quantity

I hope this article has gone some way to answer the main ‘who’, ‘what’ and ‘how’ methodological questions that need to be asked of the 2019 election polls, and to provide a recurring theme throughout for the rest of us.

That theme is to ask ourselves questions as we design, conduct and present research. At every stage we should be asking; if we are targeting the right people, reaching as many of them as we can, sampling and interviewing them in the best way; if we are asking the right questions in the right way with

the right response options; and, if we are treating, analyzing and presenting the data honestly, fairly and accurately. Where can something go wrong and error creep in?

In the end, this all comes down to one factor: quality.

I agree with the pollsters who have admitted that the time of ‘polling on the cheap is over’ because cheap solutions are giving the wrong results, and that is worse than having no results at all. As an industry we must strive for quality at all costs, even if that means fewer companies doing less research, because distrust amongst clients and respondents will inevitably lead to no companies doing any research.

This is the ‘why’ of polling – we should be doing it to deliver quality research that can be trusted – and it’s up to us to educate our audience about quality and then deliver on that promise.

The 2019 election showed that predictions are dangerous things, but if we can deliver quality, I confidently predict a bright future for research.



JIM REED

Jim Reed is the Founder of Resolve. He is a pollster with over two decades quantitative and qualitative research experience, including private and published political polling and campaigns.