

Using surveys and polling data in your journalism

November 2019



A wide variety of news stories incorporate or are even based on market and social results. Many private organisations, public sector organisations plus media and third sector organisations commission research where all or some of the results are published.

Some are commissioned to identify the opinions, attitudes, behaviour or knowledge about a) topical issues such as political events b) areas of ongoing interest among companies or organisations, or c) as media 'hooks' to provide points of interest.

If research results are collected and used in the right way, they can add real value to news stories providing context about the opinions, attitudes, behaviour or knowledge of a variety of subjects from a representative sample of people in a clearly defined audience. In some cases, they can provide additional information to help substantiate or illustrate a story.

Who publishes surveys and polls

Research agencies own research.



Government departments.



Regulators and other non-Government agencies.



Academics frequently publish research based on research data.



Private companies and charities who commission or undertake their own research.

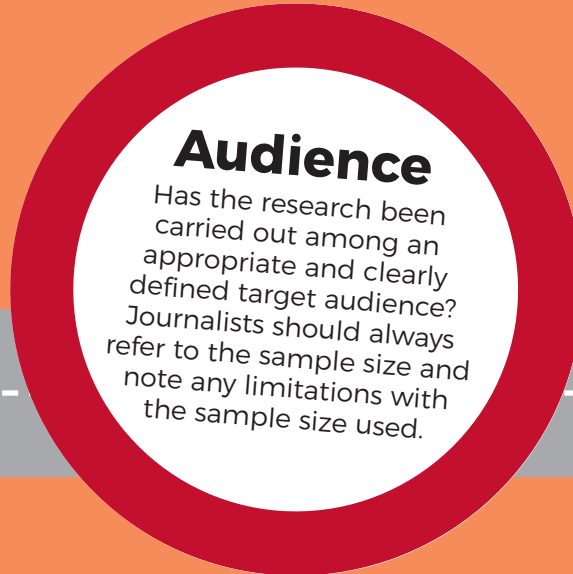


The ROADMAP to responsible reporting on surveys and polling data



Who

Who has undertaken the research? Check their credentials. Are they members of a recognised and reputable organisation? Are they members of a professional association, for example the Market Research Society? Journalists should exercise healthy scepticism about research and should make it clear in their reporting when research has been commissioned or paid for by an interested party.



Audience

Has the research been carried out among an appropriate and clearly defined target audience? Journalists should always refer to the sample size and note any limitations with the sample size used.

Think critically about the limitations of research involving children, young persons or vulnerable groups; it is important that these groups are represented, but this needs to be balanced against reliability and the informed nature of their responses (how consent was obtained).



What

Were the questions asked accurate, balanced and unambiguous or were participants led to a particular answer? Evaluate each question to see if each is accurate, balanced and unambiguous and fit for the purpose intended.

How

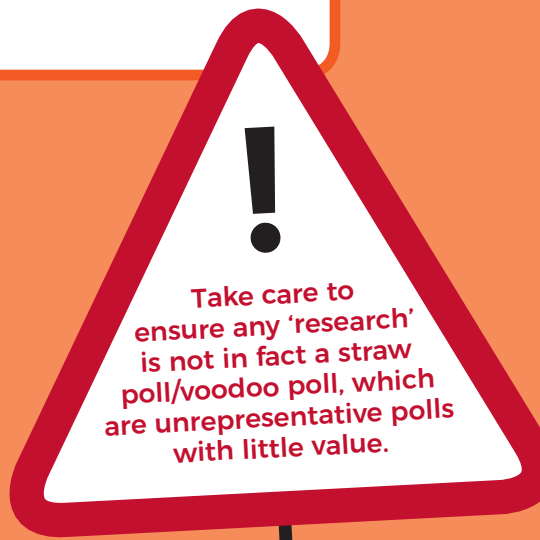
How is the methodological approach appropriate? For research results to have value they must be "representative" of a clearly defined target audience or population. For example, opinion polls should provide contextual information setting out how responses were gathered and analysed, and the steps taken to ensure that they are representative. If this information is not provided, journalists should ask for it, so that they can make a judgement on the quality of the opinion poll.

Evidence

Have the data tables been published and is there a Technical Note? A Technical Note should accompany any published research results outlining the way in which participants were selected, the profile and composition of the sample, how it was selected and what, if any, weights were applied to the research data. Scrutinise these technical details, the profile of the sample and evaluate whether or not it reflects the demographic profile of the target audience/population being represented. If the tables have not been published, ask for a copy.

Published opinion polls should include sufficient background and contextual information to enable users of the opinion poll to interpret the information. This should include information such as:

- ✓ Dates of data collection
- ✓ Method/s of obtaining participant responses (e.g. face-to-face, online, phone)
- ✓ The universe/population of interest (e.g. all adults, voters, parents, etc.)
- ✓ The percentages upon which any reported conclusions are based
- ✓ Size of the sample and geographic coverage
- ✓ Wording of questions
- ✓ A web address and/or other source where data tables may be viewed
- ✓ Name of organisation commissioning the opinion poll



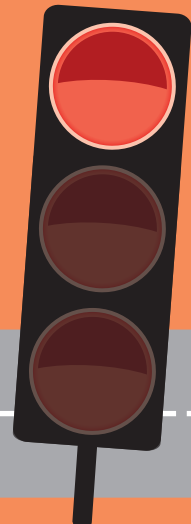
Take care to ensure any 'research' is not in fact a straw poll/voodoo poll, which are unrepresentative polls with little value.

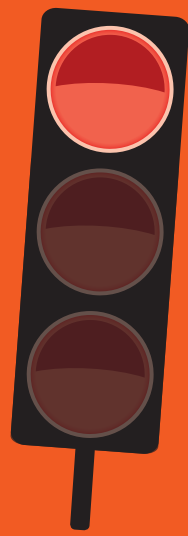
Analysis

Is the commentary on the results accurate? Have the results to the research been written up in an accurate and balanced way, to ensure they are a true reflection of the research results?

Influence

Do the people who paid for it have a particular interest in the outcome? Look at the way in which research has been conducted and reported upon. This will enable you to make a judgement about the balance of the entire exercise.





WARNING SIGNS

Results based on too small a sample size.

- Research should be conducted using a **robust sample** and be sufficiently large to allow for meaningful analysis. Typically, opinion polls are carried out among 1,000 participants.

Results based on samples that are not representative of a given audience.

- When research has been carried out among a sample that is **not representative** of a clearly defined target audience, the results produced will be skewed. In such circumstances it is not possible to draw reliable conclusions upon the target audience.

Results based on self-selecting samples where participants opted in.

Text used in a press release that subtly or significantly alters the question wording, or summarises the question, so that the meaning changes.

- The exact words of the question should be made available in full. Research results should not be cherry picked and should provide a balanced view of overall research findings.

Presenting shifts in opinion as relevant when they are not significantly different.

Where important details of a research project, such as question wording, sample size, date of interviewing, etc, are not available to the journalist to check, or where the journalist is not allowed/encouraged to report them (unless there are clear and obviously legitimate reasons for that).

- The format should make it easy and simple for journalists to be able to interpret the research findings and be able to read computer tables (or other research deliverables).

Results based on questions that are biased or leading.

More on bias

It is important to remember that “bias” is a technical term used by statisticians, and it applies to any situation where research results are systematically distorted, whether or not this is deliberate. Research findings (and therefore any reporting based on them) may be biased without anybody wanting to bias them or being aware that they are biased. Sometimes bias arises for reasons outside the control of everybody concerned with conducting research, and without their being aware that research is biased.

Bias can arise from any of a number of causes (and it is perfectly possible that more than one of these is operating at once). Bias might arise from research results that are not representative of the public as a whole, from a question(s) or other features of a research process causing participants to give inaccurate or insincere answers, or from inappropriate processing of the data once the interview is completed. Bias in the reporting is also possible, of course, even if research data itself is unbiased, since the reporting will include an element of subjective interpretation.

It would be very rare in the UK, and a serious breach of professional standards, for any reputable research market research organisation to consciously allow the sponsor or commissioner of a research project to influence research results, whether by manipulating the question wording or in any other way. It might happen when a project is conducted by a non-research professional.

If biases are detected whether conscious or unconscious, journalists have a duty of care to highlight the nature of the bias and how this may have impacted upon research results. Sometimes, depending on the circumstances, a research project being biased may be the story in itself. On other occasions, particularly if there is reason to suppose that the bias is reasonably small, the findings may still be of interest once that bias has been taken into account.

Examples of leading or biased questions

Where the question leads the participant to a particular answer.

Double-barrelled questions that are not possible to answer and the participant may wish to answer in different ways to different parts of the question asked. E.g. Do you favour candidate X and higher taxes or candidate Y and lower taxes? **Look out for the word “and” to spot a double-barrelled question.**

Skewed scales, e.g. with different numbers of positive and negative options or with the wording making options at one end of the scale more attractive than the other.

Where not all possible response options are offered as an answer.

Omissions of a “none” response where participants do not wish to pick any response from the given list of options.

Where “don’t know” responses are missing.

Questions that are factually inaccurate.

GOOD PRACTICE

Recognising a Voodoo or Straw poll has been conducted.

- These are polls where:
 - results are not representative of a clearly defined target audience
 - the questions asked are inappropriate; and
 - the sample is too small.

Understanding and being clear about the target audience and who was questioned.

- Research needs to be carried out among relevant and clearly defined target audiences.

Identifying the difference between different levels of quality in research.

Identifying when research has been undertaken using an inappropriate method of data collection.

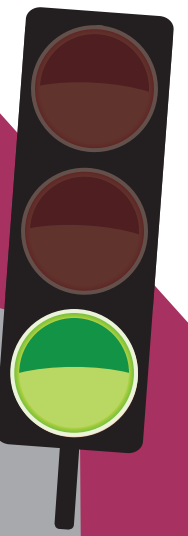
- For example, a research project evaluating social care to elderly people conducted entirely online. Internet penetration among the elderly, particularly those from the C2DE classes is lower. In this instance, a face-to-face method of data collection would be preferable.

Being realistic about the limitations of the research process.

- A poll taken weeks before an election cannot infallibly predict election results: the voters themselves do not all know how they will vote.
- Not every research project can realistically expect to receive fully honest answers from all its participants, any more than a doctor expects all his/her patients to accurately report how much they are drinking or smoking.
- No research can normally measure to fractions of a percentage point. Random probability research has a “margin of error” because they are based on samples, and this inescapable lack of precision must be acknowledged and form part of any reporting.

Noting the timing of the fieldwork.

- Research results are a snapshot of opinion at a particular point in time. When reporting upon research results it is crucial to consider the timing of fieldwork and consider whether any factors may have influenced participants responses to a particular question at a particular point in time.



COMMON MISTAKES



1 A lack of understanding of statistics.

For instance, not understanding the difference between the UK and Great Britain for statistical purposes.

2 Inaccurate reporting of results.

For example, reporting differences in results which may not be statistically significantly different. Or reporting statistically insignificant differences between groups of findings or between two research projects as evidence of a change in opinion over time (e.g. "support for the government has risen 2% since last week").

Be accurate when describing changes in % findings over time. For instance, a shift from 40% to 60% is not 'an increase of 20%'. It is an increase of 50% (the difference, 20, divided by 40), and can be described as 'an increase of 20 percentage points'. Again, reporting a movement from 5% to 10% as a 5% increase is inaccurate: it is a five-percentage point increase or a 100% increase.

3 Inferring causal explanations of research differences without adequate evidence.

4 Poor or unclear reporting of research results, altering the meaning of what the research is telling us.

Being wary that the text used in a news release does not subtly or significantly alter the question wording so that the meaning changes. The exact words of the question asked should be reported in full. Similarly, journalists should differentiate between what are research findings and what are not, which reporting is based on research findings, and what represents the personal views of the commissioner or the individual reporting the data.

References to research data should be clarified using the actual research results. For instance, simply stating 'most people feel they are not getting value for money' does not clearly reflect the research findings. Instead, the text should be 'two-thirds of those surveyed (67%) feel they are not getting value for money, for example. Do not extrapolate to the whole general population where the sample size does not support this.

5 Failing to check the figures in a press release carefully – and how they are described.

For instance, if the figures relate to a sub-sample of those who 'ever use a mobile phone', this must be clear. It would be wrong to state "x% send at least 15 text messages a day" as this implies it is x% of British adults – which is incorrect. It should say 'x% of those surveyed who ever use a mobile phone say they send at least 15 text messages a day'.

6 Omitting to report any important contextual factors.

For example, that a research project was conducted before some major recent news event that might have changed people's opinions.

7 Reporting on samples that are too small to be accurate.

Whilst the number of individuals questioned as part of an opinion poll may differ, they are typically carried out among a robust sample of between 1,000 and 2,000 individuals. Whilst it is perfectly feasible to have a smaller sample, the smaller the sample, the larger the margins of error. Research of the general public or of other large groups a sample should be several hundred. Smaller sample universes such as MPs, for example, would necessarily use a much smaller sample. Taking this example, research with MPs will have a wider margin of error than research of voters, but in such a case this is unavoidable.

Regional press releases are popular, but often the sample sizes involved (such as in Wales) are too small to justify separate analysis and reporting (i.e. they are not statistically reliable). As a guideline, to report the percentage results based on any sub-sample, the (un-weighted) sample size should be at least 100; between bases of 50 and 99 indications of proportional results could be given – along with a clear health warning about the small sample size. This, of course, does not apply where the universe questioned is a census i.e. such as an employee research project.

8 Not reporting that participants are picking from lists.

When participants are asked to pick one or more items from a list of items journalists should make it clear that this is the case – otherwise the write up may be misleading.

9 Not distinguishing between spontaneous versus prompted responses.

Some questions ask participants to pick from lists ('prompted') and others ask participants to answer in their own words ('spontaneous'). It is important to know when each technique has been used.

10 Charts or tables not providing sufficient technical details for the reader to know on what basis percentage results are based.

If charts or graphs are being used, these should make clear the full question wording, fieldwork dates, source the research agency and indicate the base size must for that given audience (either full or for sub-sample groupings being reported upon).

11 Reporting research results to decimal places

Research results should be reported as whole numbers, not to decimal points which may suggest a higher degree of accuracy.

12 Excluding "don't know" responses and re-basing findings after taking out those who responded with "don't know".

When these figures are then reported as represented the whole population, e.g. "60% of those who expressed an opinion either way". This can lead to a minority of views being inaccurately reported as a majority.

13 Not being clear upon target audience.

Not being clear about the sample audience or not making clear the proportion of the audience sampled who provided that answer. For example, reporting upon the proportions of "people" and not the true research audience, which may be "British adults 16+", for example.

Weighting is used by researchers to adjust the results of a study to bring them more in line with what is known about a population. For instance, a study may have too many women respondents and be adjusted for that accordingly. When reporting on the sample audience, journalists should refer to the unweighted, not the weighted base.

PRESS RELEASES - WHAT TO LOOK FOR



British public continue to be more satisfied with tea than coffee.

PRESS RELEASE

1 New research from WeFindOut (Regulated) shows that a majority of **GB adults aged 18-75** found tea to be more satisfying.

- 4 in 10 are satisfied with tea, compared to just 15% for coffee - 7 in 10 feel dissatisfied with drinking coffee.
- Majority consume coffee 'not very often' or 'never'. **9**
- 44% of GB adults aged 18-75 support an immediate end to the sale of coffee while 27% oppose. These numbers are similar to May where 43% supported an immediate end to coffee sales and 32% opposed.

Commenting on the findings, WeFindOut Research Director Ken Smith said:

"The British public continue to be divided on the contentious issue of tea and coffee. Looking ahead, tea producers will likely be encouraged by findings that public satisfaction with their produce remains high, meanwhile coffee brands will want to improve perceptions if they are to hold on to market share."

TECHNICAL NOTE

3 This poll was conducted by WeFindOut on behalf of BeansRus Ltd. **7**
4 WeFindOut interviewed a **representative sample of 1,084 adults aged 18-75** across Great Britain. Interviews were **conducted online: 5**
11th-14th June 2019. Data are **weighted** to match the profile of the population. All polls are subject to a wide range of potential sources of error. **6**

For more information contact Ken Smith on 0206667778 **8**

A note on headlines

Don't take headlines at face value. They may give the impression that all those within the given audience researched believe or think in one way (as no qualification of the proportions are provided in the headline).

COMPARING POLLS OR SURVEYS



Very few polls and surveys are like for like. If journalists intend on drawing comparisons between different polls and surveys they must first evaluate the following before attempting any comparison:

- What is the reputation and experience of a research organisation which conducted the research?
- Are the target audiences identical?
- How do the sample profiles compare?
- Was the sampling methodology the same?
- If a quota sample, were identical quota controls set and were data weighted and, if so, using the same weights?
- If they use a quota sample, were the quotas set appropriate?
- Are the fieldwork dates identical, or did they differ?
- Were they both carried out using the same method of data collection? Perhaps one was carried out online and one face-to-face, for example.
- Were the questions asked identical, and asked in the same order, or was this different?
- What was the sample size - were they both robust?
- Were the data analysed and reported upon in a similar way?

QUICK CHECK - WHAT DOES RELIABLE LOOK LIKE?

Reliable	Unreliable
✓ Representative sample	✗ Unrepresentative/skewed/self-selecting sample e.g. Voodoo or straw polls
✓ Robust/large sample size	✗ Small sample size
✓ Reporting on differences that are significantly different	✗ Reporting on differences that are not significantly different
✓ Providing a well-rounded balanced summary of research results	✗ Biased reporting or cherry-picking research results
✓ Reporting whole numbers	✗ Reporting to decimal places
✓ Describing the target audience of a research project in an appropriate way	✗ Describing the target audience of a research project as "people"
✓ Appropriate research methodology	✗ Inappropriate research methodology
✓ Highlighting key items that need to be highlighted about the way the question was asked - e.g. picking from a list/spontaneous response	✗ Not highlighting key items that need to be highlighted about the way the question was asked - e.g. picking from a list/spontaneous response
✓ Appropriate technical note on how research was carried out - accompanies the commentary	✗ No or inaccurate technical note
✓ Showing all the research results for a given question	✗ Not showing all the Research results for a given question
✓ Only reporting on the results to appropriate/balanced/accurate question	✗ Reporting the results from questions that are inappropriate/unbalanced/factually inaccurate
✓ Accurate description of the geographic area of coverage of the sampling area	✗ Inaccurate description of the geographic area of coverage of the sampling area e.g. using UK instead of GB
	✗ Re-basing the results by omitting those who responded "Don't know", or similar re-basing of research results

Look for the following technical pieces of information that should be present alongside the published results:

- 1** Audience.
- 2** Sample size.
- 3** Sampling methodology (quota, random probability sample, etc. - if appropriate).
- 4** Fieldwork dates (reported where relevant with reference to other events or the timing of other research conducted so that the viewer/reader can make allowance for anything that might change their understanding of the findings)
- 5** Method of data collection (face-to-face, online, self-completion, etc.).
- 6** Whether the data were weighted and, if so, to what.
- 7** Name of commissioning organisation.
- 8** Contact details.
- 9** Response options.

It is important that journalists mention the sample size, if at all possible, because it is an important bit of information, but they do also need to understand that it is not the only thing that matters. A small but representative sample is better than a large but poorly designed sample which will give an inaccurate measurement despite a big sample size.

Remember to watch out for potential warning signs (see p. 4). If any information is missing that would allow you to assess data correctly ensure you ask for more info.

SUPPORT AND GUIDANCE



If I find a poll or survey confusing, but would like to write a story about it, what should I do?

Ask for the following to help you understand how research was undertaken:

- A copy of the questionnaire.
- A copy of the computer tabulations.
- Research data (if available and you feel confident to analyse the data).
- The technical note which indicates how research was carried out e.g. when it was conducted, how it was conducted, etc.

How can I identify a reputable polling company?

Those organisations that are:

- MRS Company Partners
- MRS Members
- Organisations regulated by recognised Codes of Conduct or Practice (they should identify which Codes they use on their websites).

If I think there are dodgy polls or surveys going around, is there somewhere I can report them?

Report concerns to the Market Research Society CodeLine here:
<https://www.mrs.org.uk/standards/codeline>

I made a mistake when reporting on a poll or survey, what happens now?

If a mistake is noticed after the research results have been reported, Rule 64 of the MRS Code of Conduct which requires MRS members and MRS Company Partners to do the following:

Rule 64:

If members are aware, or ought reasonably to be aware, that findings from a project have been incorrectly or misleadingly reported by a client they must at the earliest opportunity:

- a) Refuse permission for the client to use the Member's name further in connection with the incorrect or misleading published findings; and
- b) publish in an appropriate forum (e.g. their website) the relevant technical details of the project to correct any incorrect or misleading reporting.

The IMPRESS Standards Code offers additional guidance. Code clause 1. Accuracy is particularly relevant, as well as the Guidance available on it. Visit www.impress.press/standards for more information.'

This guidance was co-written by IMPRESS and the Market Research Society (November 2019).

KEY STATISTICAL CONCEPTS



Research findings are statistical calculations of what is represented if the whole target audience or population under scrutiny is measured. As only a sample of the target audience or population is researched, particular **levels of significance** apply to research and these reflect the probability that the research findings are a true reflection of the target audience or population characteristics.

Statistical significance is the practical application of confidence intervals. **Confidence intervals** are an attempt to place precise mathematical limits to the error which may have occurred. Strictly speaking, confidence intervals only apply to research with random sampling, which most opinion polls and research projects in Britain do not use.

In reporting research results, we frequently want to know whether two different measurements reflect a **real difference**. If the two findings are sufficiently different that we can be confident that the difference is **not caused purely by chance** in the random sampling, then we say that the difference is **statistically significant**, and it can be reported as a **factual finding** of the poll. But if they are not very different, the difference may be caused only by chance, not reflecting the underlying reality, and would be statistically insignificant.

Statistically insignificant differences should not be reported, because the research does not provide adequate evidence that what they are suggesting is true.

An **opinion poll** normally tries to obtain a representative view of the population as a whole or a significant defined section of it (e.g. "over 60's", "parents", "diesel car owners") by reaching an appropriate sample.

A **survey** normally tries to obtain a representative view from a smaller more-specific group of people – individuals like MP's, university vice-chancellors, or organisations like health trusts, FTSE 100 companies, charities.

A **questionnaire/voodoo poll/ straw poll/online vote/text vote** all try to obtain a view by contacting as many people as possible to answer questions. The sample will always be self-selecting, and the numbers will have no statistical significance. Reporters may find this a valuable and useful editorial tool producing good anecdotal material or provides a different but the limitations of this type of information gathering must be clear and the results should not be included in news reporting.

When describing the methodology of the research, journalists should consider the **sampling** approach when assessing the quality of the research. Simplistically, there are three methods of selecting a sample of participants:

- 1. Probability sampling** is the 'gold standard' of sampling where each person in the population has a known chance of being selected to the sample – for example, some postal, much telephone research. Indeed, all statistical tests assume a sample has been selected using a random sampling technique
- 2. Quota sampling** is where "quotas" are set on key characteristics or variables of importance to a research project to ensure the sample profile matches that of the population. For example, research among British adults 16+, quotas may be set on variables such as age within gender, geographic area, work status, etc.so the proportions in the sample match those in the population of the clearly defined target audience as a whole.
- 3. Self-selected sampling** is where little or no control is placed upon the way in which the sample is selected, or how participants take part in a research project. Participants may opt into a research project, or no controls are placed upon the selection of participants. Care needs to be exercised when analysing results from self-selected sampling.

Journalists may be tempted to use and make judgments on the use of **averages** when reporting on research. Different types of averages have different meanings and mistaking one for another can make a report completely misleading.

- 1. Mean** = An arithmetic mean is calculated by adding several quantities together and dividing the sum by the number of quantities.
- 2. Mode** = is the most frequent value in a set of data.
- 3. Median** = represents the middle number in a given sequence of numbers when it's ordered by rank

Only a mean should be described as an "average" without some qualification, but ideally the reporting should always specify which of these is being used.