



THE UNIVERSITY OF
SYDNEY

Centre for AI, Trust & Governance

Creating Surveys to Measure Institutional Trust: A Best Practice Guide

Author: Dr George Argyrous



This work was supported by the Australian Research Council, Laureate Fellowship [Grant number: FL230100075], *Mediated Trust: Ideas, Interests, Institutions, Futures*.

Author: Dr George Argyrous

Date: 25 February 2026

Recommended Citation:

Argyrous, G., (2026), *Creating Surveys to Measure Institutional Trust: A Best Practice Guide*. <https://doi.org/10.25910/d85p-gx62>

Centre for AI, Trust and Governance, The University of Sydney, Faculty of Arts & Social Sciences, NSW 2006, Australia.

Table of Contents

1. Who this guide is for.....	3
2. Guide.....	3
Start with a clear definition of what you want to measure	3
Decide the ‘trust object’ and level of specificity	3
Choose the item type: global trust, trust dimensions, or both	4
Write trust questions that are easy to answer	4
Wording principles.....	4
Prefer item-specific scales over Likert scales, especially agree–disagree options.....	5
Only ask for variation in trust that is meaningful.....	5
Midpoints and ‘Don’t know’	6
Rating and ranking	6
Layout and device mode	7
Build an institutional trust module that is interpretable	7
Test, refine, and document your questions	7
Analyse and report trust measures responsibly	8
Appendix A. Example institutional trust question set (adaptable template).....	9
Global trust	9
Trustworthiness dimensions.....	9
Optional open-ended item (to clarify referents and improvement ideas).....	9
Appendix B. Quick checklists	10
Pre-field checklist	10
Analysis checklist	10
Key references for follow-up.....	11
Contact	11

1. Who this guide is for

This guide is written for practitioners, researchers, and policy teams who want to design a short survey module (or a full questionnaire) to measure trust in institutions such as government agencies, regulators, the media, banks, courts, health services, or science. It focuses on the practical decisions that most affect data quality: defining the trust object, writing unambiguous items, choosing response options, and testing questions before fielding.

It draws on common issues identified in the trust-survey literature and on methodological evidence from survey design research, including experimental comparisons of alternative wordings and formats.

2. Guide

Start with a clear definition of what you want to measure

'Institutional trust' is not a single, universally agreed construct. In practice, survey questions can capture different things, including:

- A global evaluation (e.g., 'How much do you trust the health system?').
- Beliefs about trustworthiness (e.g., competence, integrity, fairness, benevolence).
- Confidence in performance (e.g., 'How confident are you that...?') which can overlap with, but is not identical to, trust.
- Distrust, mistrust, and scepticism, which may not sit neatly on a single 'trust' continuum from low to high (Bunting, Gaskell, and Stoker, 2021; Jennings et al., 2021).

Before writing any items, write a one-sentence measurement intent statement. For example: 'We want to measure public perceptions of the regulator's integrity and fairness when making decisions.' This will keep wording choices aligned with your purpose and avoid mixing distinct constructs in the same battery.

Decide the 'trust object' and level of specificity

Many trust questions use broad labels such as 'the government' or 'the media'. Evidence shows that respondents may answer these items by thinking of very different concrete referents (e.g., a particular ministry, politician, outlet, or platform). This can make results unstable over time and hard to interpret (Daniller et al., 2017; Tsfaty et al., 2023).

Use one of these practical strategies:

1. Use a specific, named institution (best for program evaluation and organisational reporting)
2. Use a broad institution label, but add a follow-up item that identifies the referent respondents should have in mind (useful for research and benchmarking)
3. If you must use a broad label, add brief clarifying text (e.g., 'by government, we mean 'Australian Government departments and agencies').

If you are comparing trust across multiple institutions, keep the level of specificity consistent (e.g., all at the national level, or all at the local-service level). Avoid mixing

sector (e.g., ‘banks’) with a single institution (e.g., ‘the Reserve Bank’) unless that is intentional.

Choose the item type: global trust, trust dimensions, or both

A single ‘direct’ trust item is quick and often useful for tracking trends, but it can be ambiguous about what respondents are assessing. Studies of trust in scientists, for example, show that global measures can blend multiple and differing underlying beliefs (Besley and Tiffany, 2023).

A practical approach is to combine:

- One global trust item (for a headline indicator), and
- A small set of trustworthiness dimension items (to explain why trust is high/low and what to improve).

Table 1: Suggested trustworthiness dimensions

Dimension	What it means in plain language	Example
Competence / effectiveness	Does it do its job well?	How confident are you that [institution] can do its job effectively?
Integrity / honesty	Does it act truthfully and keep promises?	How often do you think [institution] acts honestly?
Fairness / impartiality	Does it treat people fairly and apply rules consistently?	How fair do you think [institution] is when making decisions?
Benevolence / acting in the public interest	Does it care about people like you?	To what extent does [institution] act in the public interest?
Transparency / accountability	Does it explain decisions and take responsibility?	How well does [institution] explain the reasons for its decisions?

Where relevant, consider separate items for distrust/mistrust rather than assuming they are just the opposite of trust (Bunting, Gaskell, and Stoker, 2021).

Write trust questions that are easy to answer

Good institutional trust questions reduce ambiguity in four places: the object, the behaviour or judgement being asked for, the time frame, and the response scale.

Wording principles

- Use plain language – avoid jargon (e.g., ‘governance’, ‘integrity framework’) unless your audience is specialised.
- Ask about one thing at a time – avoid double-barreled items (e.g., ‘honest and competent’).
- Avoid leading or emotionally loaded words – keep tone neutral.
- Specify a time frame that is realistic for respondents to reference – memory is notoriously unreliable so keep the reference period short (e.g., ‘in the past 1

month' rather than 'In the past 12 months'). If you want general trust, avoid time frames.

- If respondents may not have a view or enough information, provide a legitimate 'Don't know / not enough knowledge' option (OECD, 2017).

Prefer item-specific scales over Likert scales, especially agree–disagree options

Likert scales, especially widely used Agree–Disagree options, invite acquiescence effects and increase cognitive burden. Item-specific options (sometimes called construct-specific), on the other hand, can improve clarity. Item-specific questions build into each response option clear wording about the quality of trust you are interested in. This removes the need to interpret what the respondent was thinking if they 'disagree' with a statement, and avoids biases (such as social desirability) that arise from the ambiguity of Likert scales.

An example of the difference between the two is in the following table.

Likert scale example	Item-specific scale example
I trust the government to do the right things – Agree – Disagree	Which of the following statements best represents your opinion: – I generally trust the government to do the right things – I have mixed views about trusting the government -- sometimes the government will do the right things but other times it won't – I don't trust the government to do the right things

Only ask for variation in trust that is meaningful

Response scales often ask respondents to make very fine-grained distinctions on some dimension of trust. This is often done by:

- using verbal labels that are ambiguous and interpreted differently by different people (e.g. 'somewhat' and 'sometimes' have very wide interpretations and should be avoided)
- using unlabelled numerical scales (e.g. 0–10). Responses to these are also ambiguous and can introduce biases due to different perceptions of what 'good' constitutes on the scale (although they can be useful in international surveys where verbal labels may not have consistent interpretation across countries).

Ask yourself:

- do I need all the response options? Will I always combine some response options anyway when I do the analysis? If so, combine these options into one category in the survey.
- are the distinctions clear and will be understood consistently by all potential respondents?

Sometimes a simple dichotomous scale (only 2 options) might be sufficient to get the distinction you are interested in.

Midpoints and ‘Don’t know’

A midpoint can mean different things (neutral, mixed feelings, or uncertainty). If neutrality is substantively meaningful, include and clearly label the midpoint. Midpoints such as ‘Neutral’, ‘Neither Agree nor Disagree’, on the other hand, should be avoided because it is not clear why respondents select them. If you want respondents to take a directional stance, consider omitting a midpoint and offering a ‘Don’t know’ or ‘Can’t say’ option for genuine uncertainty.

Rating and ranking

Most trust surveys include questions that ask respondents to rate ‘trust’ against a Likert or Item-specific scale. An alternative to these rating scales is to ask respondents to rank items from ‘most’ to ‘least’ (or equivalent). There is some evidence to suggest that ranking questions are less susceptible to systematic bias and measurement error, potentially quicker, more engaging for respondents, and impose a lower cognitive load compared to traditional rating scales. Items that receive the same rating on a Likert scale are sometimes ranked differently when respondents are asked to rank-order items, suggesting that rating scales do not capture important variation in strength or quality of opinion on trust.

To minimize cognitive burden on respondents, especially when asking them to rank a large number of items (i.e. 5 or more items), a useful option is not to ask them to order the whole list, but instead to:

- Choose the ‘most/highest/...’ favored item
- Choose the next 1 or 2
- Choose the ‘lowest/least/...’ favored item

The following box provides an example of this.

Which of the following issues do you think the government can be MOST TRUSTED to act on fairly and in the public interest? (*choose one*)

1. **Immigration and border management** (e.g. visas, migration, border processes, and humanitarian/refugee programs)
2. **Anti-corruption and accountability** (e.g., honesty in public office, preventing abuse of power)
3. **Elections and democracy** (e.g., voting systems, electoral administration, political donations rules).
4. **National security and cyber security** (e.g., intelligence, counter-terrorism, critical infrastructure protection).
5. **First Nations affairs** (e.g., services, equity, Closing the Gap priorities)
6. **Digital media, digital services and the internet** (e.g. access and inclusion, digital safety, privacy, access to platforms, age-appropriate content, preventing abuse)

On which of these issues do you trust government THE LEAST? (repeat list)

If a complete ordering of the list is required, use a series of pairwise comparisons, whereby two items are chosen at random, and respondents are asked to order only these two items. Then a third item from the complete list is chosen and respondents are then asked to insert it into the list, and so on with the remaining items.

Layout and device mode

If your survey is online, test the layout on mobile and desktop. Rank orders, grids, and long batteries of questions are particularly sensitive to small layout differences (Revilla and Couper, 2018).

Build an institutional trust module that is interpretable

For most research needs, a short, interpretable module may be sufficient that includes:

1. A global trust item for each institution (or a small subset of institutions) of interest.
2. 2-4 trustworthiness dimension items for each institution (choose dimensions that match your use case).
3. One open-ended question with a clear prompt to capture aspects of trust that you did not anticipate and may surprise you. Think about having this at the start of the survey before respondents are fatigued and primed by the closed-ended responses that you provide to them.
4. Basic demographics and 'experience' questions (e.g., recent contact with the institution, service use).

Test, refine, and document your questions

The fastest way to improve a trust survey is to test questions early. Even a small amount of pretesting can uncover major interpretation problems.

A practical testing pathway could be:

1. Desk review: check wording for ambiguity, double-barrels, and inconsistent specificity across institutions
2. Cognitive interviews: ask 8-15 people to 'think aloud' while answering and probe what they understood by the institution label and response options (Fowler, 2014; Tourangeau, Rips, and Rasinski, 2000)
3. Pilot survey: run a small pilot in your intended mode to check completion time, missing data, 'Don't know' rates, and response distributions
4. Split-ballot experiments (optional but powerful): randomly assign alternative wordings (e.g., broad vs specific referents; open vs closed items) and compare results (Daniller et al., 2017; Bernhard-Harrer and Pfaff, 2025)
5. Document decisions: keep a short 'question log' noting why you chose each wording and scale so future waves remain comparable (OECD, 2017).

Analyse and report trust measures responsibly

- Avoid over-interpreting small differences (especially across institutions with different public audiences and visibility). This is especially important with large sample surveys where a difference might be statistically significant but unimportant from a practical or theoretical standpoint
- Report the exact wording, response options, and survey mode whenever you present results, at least in an appendix
- When tracking trends, keep items and context stable. If you must change wording, run an overlap period or split-ballot bridge
- Use dimension items to translate results into action (e.g., low fairness suggests different interventions than low competence)

Appendix A. Example institutional trust question set (adaptable template)

Below is an example module you can adapt. Replace [institution] with the institution you are interested in. Choose one response scale and keep it consistent.

Global trust

Thinking about [institution] how would you describe your feelings of trust towards it? (open-ended to get unanticipated responses)

Which of the following statements best summarizes your feelings of trust toward [institution]?

1. I trust [institution] to do what is right
2. My feelings of trust are mixed: sometimes [institution] does what is right but sometimes it doesn't
3. I mostly don't trust [institution]: they generally don't do what is right
4. Don't know/Can't say

Trustworthiness dimensions

How confident are you that [institution] can do its job effectively?

1. I have no confidence that [institution] can do its job effectively
2. I have some confidence that [institution] can do its job effectively
3. I am very confident that [institution] can do its job effectively
4. Don't know/can't say

[adapt this response set for the following questions]

- How fair do you think [institution] is when making decisions?
- How honest do you think [institution] is in the information it provides to the public?
- To what extent does [institution] act in the public interest?
- How well does [institution] explain the reasons for its decisions?

Optional open-ended item (to clarify referents and improvement ideas)

When answering the questions above, what were you mainly thinking about?

Optional follow-up: What is one thing [institution] could do to increase your trust?

Appendix B. Quick checklists

Pre-field checklist

1. Have you defined the trust object and level (institution, sector, or system)?
2. Do items separate 'trust' from 'confidence in performance' where that distinction matters?
3. Are items free from double-barrels and leading language?
4. Are response options labelled clearly (endpoints at minimum)?
5. Is there a planned approach for 'Don't know' responses?
6. Have you tested items on the devices respondents will use (mobile/desktop)?

Analysis checklist

1. Are missing/'Don't know' rates acceptable and similar across institutions?
2. Do distributions show strong floor/ceiling effects that limit discrimination?
3. Do dimension items cohere (reliability) and relate to expected correlates (construct validity)?
4. Can you clearly explain what respondents were likely thinking of (referents) when answering broad labels?

Key references for follow-up

The following sources provide practical guidance and empirical evidence on measuring institutional trust and designing robust survey questions.

Bernhard-Harrer, J., and Pfaff, K. 2025. Question Form Matters: Examining Trust in Government through Open and Closed Survey Items. *Journal of Survey Statistics and Methodology*, 13(4), 370–392. <https://doi.org/10.1093/jssam/smaf014>

Bunting, G., Gaskell, J., and Stoker, G. 2021. Trust, Mistrust and Distrust: A Gendered Perspective on Meanings and Measurements. *Frontiers in Political Science*, 3, 1–15. <https://doi.org/10.3389/fpos.2021.642129>

Daniller, A., Allen, D., Tallevi, A., and Mutz, D. 2017. Measuring Trust in the Press in a Changing Media Environment. *Communication Methods and Measures*, 11(1), 76–95. <https://doi.org/10.1080/19312458.2016.1271113>

DeCastellarnau, A., and Revilla, M. 2017. Two approaches to evaluate measurement quality in online surveys: An application using the Norwegian Citizen Panel. *Survey Research Methods*, 11(4), 415–433. <https://doi.org/10.18148/srm/2017.v11i4.7226>

Dillman, D. A., Smyth, J. D., and Christian, L. M. 2014. *Internet, Phone, Mail, and Mixed-Mode Surveys: The Tailored Design Method*, 4th ed. Hoboken, NJ: Wiley.

Dykema, J., Garbarski, D., Wall, I. F., and Edwards, D. F. 2019. Measuring Trust in Medical Researchers: Adding Insights from Cognitive Interviews to Examine Agree-Disagree and Construct-Specific Survey Questions. *Journal of Official Statistics*, 35(2), 353–386. <https://doi.org/10.2478/jos-2019-0017>

Fowler, F. J. 2014. *Survey Research Methods*, 5th ed. Thousand Oaks, CA: Sage.

Höhne, J. K., Revilla, M., and Lenzner, T. 2018. Comparing the Performance of Agree/Disagree and Item-Specific Questions across PCs and Smartphones. *Methodology*, 14(3), 109–118. <https://doi.org/10.1027/1614-2241/a000151>

OECD. 2017. *OECD Guidelines on Measuring Trust*. Paris: OECD Publishing.

Tourangeau, R., Rips, L. J., and Rasinski, K. 2000. *The Psychology of Survey Response*. Cambridge: Cambridge University Press.

Tsfati, Y., Strömbäck, J., Lindgren, E., Boomgaarden, H. G., and Vliegenthart, R. 2023. What News Outlets Do People Have in Mind When They Answer Survey Questions About Trust in 'Media'? *International Journal of Public Opinion Research*, 35(2). <https://doi.org/10.1093/ijpor/edad008>

Contact

Centre for AI, Trust and Governance

Email: aitg.centre@sydney.edu.au
sydney.edu.au

CRICOS 00026A