SURVEYS · NONRESPONSE, WEIGHTING AND DATA QUALITY

Attention Checks and Instructed Response Items: Do They Work?

Attention checks are questions designed to show whether online survey respondents are reading. This article describes the main types, reviews the evidence on how well they identify inattentive respondents and explains how to use them without harming data quality.

ModQu research article · 1,019 words · 13 references (APA 7th edition)

Online surveys have no interviewer to notice when a respondent stops paying attention. To fill that gap, researchers include **attention checks**: questions with an obvious correct answer that only an inattentive respondent would get wrong. They are now common in academic, commercial and public opinion surveys. Their popularity raises a practical question: do they identify the right people, and does using them improve the quality of survey knowledge? (Berinsky et al., 2014; Ward & Meade, 2023).

Types of attention check

Instructed response items appear within a set of ordinary questions and tell the respondent which answer to give, for example a row in a grid reading "For this item, please select 'Somewhat disagree'". Respondents who are reading will comply; those clicking without reading will often choose another option (Meade & Craig, 2012; Gummer et al., 2021).

Instructional manipulation checks hide an instruction within a longer block of text. A question may appear to ask about the respondent's hobbies, but the final sentence of its introduction asks them to ignore the question and click a particular button. Only those who read the full text will follow the instruction (Oppenheimer et al., 2009).

Bogus or infrequency items ask about something that is almost certainly true or false for everyone, such as "I have been to every country in the world". Agreeing suggests inattention (Huang et al., 2012; Meade & Craig, 2012).

Figure 1 illustrates how pass rates tend to differ between types.

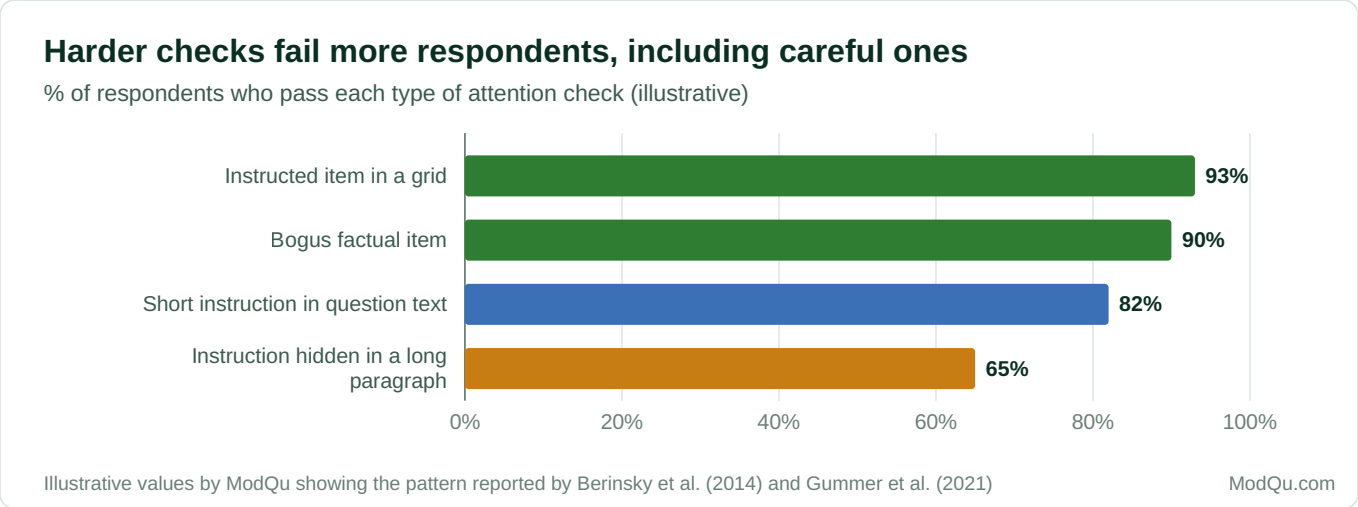


Figure 1. Illustrative pass rates for different types of attention check. Illustrative values by ModQu showing the general pattern in the literature.

What the evidence shows

They identify less attentive respondents

Respondents who fail attention checks do behave differently. Oppenheimer et al. (2009) found that participants who failed an instructional manipulation check showed weaker responses to experimental manipulations, so that excluding them increased statistical power. Berinsky et al. (2014) found that respondents who failed attention screeners in political surveys answered experimental treatments less consistently and gave lower-quality data on several measures. Gummer et al. (2021) found that failing instructed response items was associated with other signs of low-quality responding, such as straightlining and speeding.

But they also catch careful respondents

Attention checks are imperfect. Some respondents fail through a brief lapse rather than general inattention, some deliberately refuse to follow an instruction they find odd, and some misunderstand it. Long hidden instructions are failed by many respondents who are otherwise answering carefully (Anduiza & Galais, 2017). Respondents who fail tend to be younger and less educated, so excluding them changes the composition of the sample (Berinsky et al., 2014; Alvarez et al., 2019).

They can change how people answer

Attention checks may alter respondents' behaviour. Hauser and Schwarz (2015) found that completing an instructional manipulation check made respondents more careful and analytical on later tasks, which could change the results of experiments that depend on intuitive responses. Clifford and Jerit (2015) tested whether measures intended to increase attention also increased socially desirable answers and found limited effects. Kung et al. (2018) found that including attention check items did not harm the measurement properties of the scales around them.

Using attention checks well

The evidence supports using attention checks, but with care (Berinsky et al., 2021; Ward & Meade, 2023):

- **Use more than one check** and base decisions on overall performance rather than a single failure. Berinsky et al. (2021) found that using several items gave a more reliable measure of attention.
- **Prefer simple instructed items** to long hidden instructions, which penalise careful respondents who skim introductions.
- **Decide rules in advance**, before seeing the results, to avoid choosing exclusions that favour particular findings.
- **Report results with and without exclusions**, so readers can see whether conclusions depend on them.
- **Consider weighting** after exclusions, because those who fail differ demographically from those who pass (Alvarez et al., 2019).
- **Combine with other indicators**, such as completion time and straightlining, rather than relying on checks alone (Meade & Craig, 2012).

The wording and placement of a check also matter. A good check is short, unambiguous and clearly part of the survey, so that respondents who are reading understand it immediately. It should not look like a trick, which can annoy respondents and lead some to refuse it on principle. Checks placed inside grids catch respondents who click down a column without reading the rows, while a check placed late in the questionnaire shows whether attention has faded. Respondents should never be penalised in ways they cannot foresee, such as losing an incentive for one missed item, because this raises ethical concerns and may encourage complaints. In panels where respondents take many surveys, the same checks should not be repeated too often, since experienced respondents learn to spot them, which reduces their value as a measure of attention (Hauser & Schwarz, 2016).

Some researchers use attention checks as **covariates** rather than exclusion criteria, reporting how results differ between more and less attentive respondents. Alvarez et al. (2019) argued that this approach keeps the sample representative while still showing the influence of inattention.

An example

A web survey of 1,500 adults includes three instructed response items placed in different grids. Ninety per cent pass all three, 7 per cent fail one and 3 per cent fail two or more. The research team, following a rule set before fieldwork, excludes the 3 per cent who failed two or more checks and keeps those who failed one, flagging them. Comparing results, the team finds that excluding the flagged group as well would change the main estimates by less than one percentage point but would reduce the share of respondents under 30 by 2 points. It therefore keeps the flagged respondents and notes the comparison in its report (illustrative example).

Conclusion

Attention checks do work, in that they identify respondents whose answers are less reliable. They are not perfect: they also catch some careful respondents, they are failed more often by some groups than others, and difficult versions can do more harm than good. Using several simple checks, setting rules in advance, combining checks with other indicators and reporting the effect of exclusions allows researchers to gain the benefits of attention checks while limiting their costs.

References

Alvarez, R. M., Atkeson, L. R., Levin, I., & Li, Y. (2019). Paying attention to inattentive survey respondents. *Political Analysis, 27*(2), 145–162. https://doi.org/10.1017/pan.2018.57

Anduiza, E., & Galais, C. (2017). Answering without reading: IMCs and strong satisficing in online surveys. *International Journal of Public Opinion Research, 29*(3), 497–519. https://doi.org/10.1093/ijpor/edw007

Berinsky, A. J., Margolis, M. F., & Sances, M. W. (2014). Separating the shirkers from the workers? Making sure respondents pay attention on self-administered surveys. *American Journal of Political Science, 58*(3), 739–753. https://doi.org/10.1111/ajps.12081

Berinsky, A. J., Margolis, M. F., Sances, M. W., & Warshaw, C. (2021). Using screeners to measure respondent attention on self-administered surveys: Which items and how many? *Political Science Research and Methods, 9*(2), 430–437. https://doi.org/10.1017/psrm.2019.53

Clifford, S., & Jerit, J. (2015). Do attempts to improve respondent attention increase social desirability bias? *Public Opinion Quarterly, 79*(3), 790–802. https://doi.org/10.1093/poq/nfv027

Gummer, T., Roßmann, J., & Silber, H. (2021). Using instructed response items as attention checks in web surveys: Properties and implementation. *Sociological Methods & Research, 50*(1), 238–264. https://doi.org/10.1177/0049124118769083

Hauser, D. J., & Schwarz, N. (2015). It's a trap! Instructional manipulation checks prompt systematic thinking on "tricky" tasks. *SAGE Open, 5*(2). https://doi.org/10.1177/2158244015584617

Hauser, D. J., & Schwarz, N. (2016). Attentive Turkers: MTurk participants perform better on online attention checks than do subject pool participants. *Behavior Research Methods, 48*(1), 400–407. https://doi.org/10.3758/s13428-015-0578-z

Huang, J. L., Curran, P. G., Keeney, J., Poposki, E. M., & DeShon, R. P. (2012). Detecting and deterring insufficient effort responding to surveys. *Journal of Business and Psychology, 27*(1), 99–114. https://doi.org/10.1007/s10869-011-9231-8

Kung, F. Y. H., Kwok, N., & Brown, D. J. (2018). Are attention check questions a threat to scale validity? *Applied Psychology, 67*(2), 264–283. https://doi.org/10.1111/apps.12108

Meade, A. W., & Craig, S. B. (2012). Identifying careless responses in survey data. *Psychological Methods, 17*(3), 437–455. https://doi.org/10.1037/a0028085

Oppenheimer, D. M., Meyvis, T., & Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology, 45*(4), 867–872. https://doi.org/10.1016/j.jesp.2009.03.009

Ward, M. K., & Meade, A. W. (2023). Dealing with careless responding in survey data: Prevention, identification, and recommended best practices. *Annual Review of Psychology, 74*, 577–596. https://doi.org/10.1146/annurev-psych-040422-045007