Supporting Information

1 REMINDER knowledge scales

Both the general and the immigration scale derived from the REMINDER data set (Meltzer et al. 2020) were constructed using `mirt` (Chalmers 2012).

1.1 General knowledge scale

The three general items loaded well onto the factor: G1 at 0.51, G2 at 0.894, and G3 at 0.511. The discrimination value for each was acceptable (in each case greater than 1), and the test information plot suggested that the test provided most information right below $\theta = 0$, corresponding to mean estimated ability. Unidimensionality was investigated using parallel analysis by way of the `psych` (Revelle 2022) package, which suggested one factor/dimension. Local independence was investigated using Yen's Q3 (Yen 1993). The largest Q3 value was -0.411. Yen suggests a cut-off value of 0.2, but as pointed out by Ayala (2009), a Q3 test tends to give inflated negative values for short tests. Indeed, Yen's own suggestion was in the context of scales with at least 17 items. For that reason, a value of -0.411 would seem acceptable, given the short scale. Model fit was evaluated using empirical plots, which suggested acceptable fit. The measured θ values ranged from -1.358 to 0.729, with a mean of 0.

1.2 Immigration knowledge scale

The three immigration items loaded well onto the factor as well: I1 at 0.731, I2 at 0.643, and I4 at 0.464 The discrimination values where acceptable, with all greater than 1, except for I4 at 0.891. Parallel analysis suggested one factor/dimension, and the test information plot showed the scale to discriminate best right below $\theta = 0$. Local independence was investigated using Yen's Q3, with the largest value being -0.318, which in light of the above observation about short scales should be considered acceptable. The empirical plots suggested an acceptable fit. The measured θ values ranged from -1.166 to 0.909, with a mean of 0.

2 BES knowledge scale

The BES data was taken from Wave 17 of the 2014-2023 British Election Study Internet Panel (Fieldhouse 2020) (N = 34,366). 0.4% of the data was missing across the variables used (education, age, gender, and four knowledge items). These were imputed using multiple imputation in R's `Hmisc` package (Harrell (2022)). The BES knowledge scale, too, was constructed using `mirt` (Chalmers 2012). The four items loaded well onto the factors: BES1 at 0.7, BES2 at 0.852, BES3 at 0.725, and BES4 at 0.653 The discrimination value for each was acceptable, in

each case greater than 1, and the test information plot suggested that the test provided most information right below $\theta = 0$. Unidimensionality was investigated using parallel analysis in the `psych` (Revelle 2022) package, which suggested one factor/dimension. Local independence was investigated using Yen's Q3 (Yen 1993). The largest Q3 value was -0.365. For reasons discussed above, this would seem an acceptable value, given the short scale. Model fit was evaluated using empirical plots, which suggested acceptable fit. The measured θ values ranged from -1.332 to 1.007, with a mean of 0.

3 ANES knowledge scale

The ANES data was taken from the 2019 Pilot Study (ANES 2019) (N = 3,000). 0.2% of the data was missing across the variables used (education, age, gender, and three knowledge items). These were imputed using multiple imputation in R's `Hmisc` package (Harrell 2022). The knowledge scale was constructed using `mirt` (Chalmers 2012). The three items loaded well onto the factors: ANES1 at 0.954, ANES2 at 0.915, and ANES3 at 0.738 The discrimination value for each was acceptable, in each case greater than 1, and the test information plot suggested that the test provided most information at around $\theta = 0$. Unidimensionality was investigated using parallel analysis in the `psych` (Revelle 2022) package, which suggested one factor/dimension. Local independence was investigated using Yen's Q3 (Yen 1993). The largest Q3 value is -0.796. For reasons discussed above, this would seem an acceptable value, given the short scale. Model fit was evaluated using empirical plots, which suggested acceptable fit. The measured θ values ranged from -1.095 to 0.902, with a mean of 0.

4 CES knowledge scale

The CES data was taken from the 2020 Cooperative Election Study (Ansolabehere and Luks 2021) (N = 61,000). 765 of the observations had missing data in place of the correct answer (e.g., the correct political affiliation of the relevant state governor), and could as such not be coded for accuracy. These observations where therefore removed. An additional 97 observations were missing at least one response on one of the knowledge questions. Given the small number of missing responses here, a decision was taken to simply remove these rather than impute them. The knowledge scale was constructed using `mirt` (Chalmers 2012). The three items loaded well onto the factors: CES1 at 0.911, CES2 at 0.925, CES3 at 0.913, and CES4 at 0.86. The discrimination value for each was acceptable, in each case greater than 1, and the test information plot suggested that the test provided most information at around $\theta = 0$. Unidimensionality was investigated using parallel analysis in the `psych` (Revelle 2022) package, which suggested one factor/dimension. Local independence was investigated using Yen's Q3 (Yen 1993). The largest Q3 value is -0.388. For reasons discussed above, this would seem an acceptable value, given the short scale. Model fit was evaluated using empirical plots, which suggested acceptable fit. The measured θ values ranged from -1.611 to 0.635, with a mean of 0.

5 Climate change knowledge scale

The data set used to construct the climate change knowledge scale was derived from the survey "Public Perceptions of Climate Change across Four European Countries: United Kingdom, France, Germany and Norway" (Pidgeon 2016). 0.5% of the data was missing across the variables used (education, age, gender, and the three knowledge items). These were imputed using multiple imputation in R's `Hmisc` package (Harrell 2022). The knowledge scale was constructed using `mirt` (Chalmers 2012). The three items loaded well onto the factors: CC1 at 0.938, CC2 at 0.718, and CC3 at 0.437. The discrimination value for each was acceptable, in each case greater than 1, except for COV3, which had a discrimination value of 0.827. The test information plot suggested that the test provided most information at around θ = -1.5. Unidimensionality was investigated using parallel analysis in the `psych` (Revelle 2022) package, which suggested one factor/dimension. Local independence was investigated using Yen's Q3 (Yen 1993). The largest Q3 value was -0.435. For reasons discussed above, this would seem an acceptable value, given the short scale. Model fit was evaluated using empirical plots, which suggested acceptable fit. The measured θ values ranged from -1.737 to 0.661, with a mean of 0.

6 COVID knowledge scale

The data set used to construct the COVID knowledge scale was derived from a separate survey administered by the authors in connection with a separate project in July 2020. The scale was constructed using `mirt` (Chalmers 2012), and the three items loaded well onto the factors: COV1 at 0.589, COV2 at 0.628, and COV3 at 0.618 The discrimination value for each was acceptable, in each case greater than 1, except for COV3, which had a discrimination value of 0.827. The test information plot suggested that the test provided most information at around θ = -2. Unidimensionality was investigated using parallel analysis in the `psych` (Revelle 2022) package, which suggested one factor/dimension. Local independence was investigated using Yen's Q3 (Yen 1993). The largest Q3 value is -0.192. Model fit was evaluated using empirical plots, which suggested acceptable fit. The measured θ values ranged from -1.863 to 0.478, with a mean of 0.

References

ANES. 2019. "2019 Pilot Study." 2019. `https://electionstudies.org/data-center/2019-pilot-study/`.

Ansolabehere, Brian F. Schaffner, Stephen, and Sam Luks. 2021. "Cooperative Election Study, 2020: Common Content." Cambridge, MA: Harvard University. 2021. `http://cces.gov.harvard.edu/`.

Ayala, R J de. 2009. *The Theory and Practice of Item Response Theory*. Methodology in the Social Sciences. New York, NY: Guilford Publications.

Chalmers, R. Philip. 2012. "Mirt: A Multidimensional Item Response Theory Package for the r Environment." *Journal of Statistical Software* 48 (6): 1–29. `https://doi.org/10.18637/jss.v048.i06`.

Fieldhouse, J. Green, E. 2020. "British Election Study Internet Panel Wave 17." 2020.

Harrell, Frank. 2022. *Hmisc: Harrell Miscellaneous*. https://CRAN.R-project.org/package=Hmisc.

Meltzer, Christine, Jakob-Moritz Eberl, Nora Theorin, Fabienne Lind, Tobias Heidenreich, Sebastian Galyga, Hajo G. Boomgaarden, Jesper Strömbäck, and Christian Schemer. 2020. "REMINDER: Online-Panel Study on Migration and Mobility Attitudes 2017-2018 (SUF edition)." AUSSDA. https://doi.org/10.11587/LBSMPQ.

Pidgeon, N. 2016. "Public Perceptions of Climate Change Across Four European Countries: United Kingdom, France, Germany and Norway, 2016." 2016. https://beta.ukdataservice.ac.uk/datacatalogue/doi/?id=8325#!#1.

Revelle, William. 2022. *Psych: Procedures for Psychological, Psychometric, and Personality Research*. Evanston, Illinois: Northwestern University. https://CRAN.R-project.org/package=psych.

Yen, Wendy M. 1993. "Scaling Performance Assessments: Strategies for Managing Local Item Dependence." *Journal of Educational Measurement* 30 (3): 187–213. https://doi.org/https://doi.org/10.1111/j.1745-3984.1993.tb00423.x.