Utrecht University

Faculty of Social and Behavioural Sciences

Network Analysis: A novel approach to the problem of nonresponse bias

By

SHANNON DICKSON

Methodology and Statistics for the Behavioural, Biomedical, and Social Sciences

Supervisors:
Dr Peter Lugtig
Dr Bella Struminskaya

Word count: 2409/2500

FETC Ethics Application Number: 22-2063

Student ID: 6369693

Target journal: Journal of Survey Statistics and Methodology

1 Introduction

Survey research is confronted with two paradoxical trends: Response rates are declining while estimates of nonresponse bias are stable (Lipps 2009; Andreski 2012; Schoeni et al. 2013; Czajka and Beyler 2016; Beullens et al. 2018). Nonresponse is one possible source of error in surveys and occurs when not all the sampled units respond. Immediate consequences of nonresponse include loss of precision due to reduced sample sizes and increased survey costs to boost responses. Nonresponse can also weaken the integrity of the data due to nonresponse bias. Since it would be unreasonable to expect all intended units to respond to our surveys, nonresponse is an inescapable fact of this kind of research. Rising nonresponse is well observed (Brick and Williams 2013). Nonresponse occurs for a variety of reasons of which some are situation specific. Face-to-face surveys attempt to contact potential respondents at home but may catch them at an inconvenient time or they might be unwilling to participate. Mail surveys may not reach the intended recipient due to an incorrect (email) address. Web surveys often fail to capture certain demographics such as older and/or rural recipients. Aside from the mode of the survey there are many more potential contributors to nonresponse. Survey length, purpose/saliency, incentivisation, source of sampling frame, and the type of question or statistic estimated and interactions between these variables may all contribute to nonresponse and increase the chance of bias. Prior investigations did not find convincing associations between these factors and survey nonresponse or nonresponse bias (Groves and Peytcheva 2008) but a more recent meta-analysis found a negative association between nonresponse and bias (Cornesse and Bosnjak 2018). Still, only a small magnitude of bias is detected across a range of surveys in health (Gundgaard et al. 2007), consumer research (Curtin, Presser, and Singer 2000), opinion polls (Keeter et al. 2000), and even simulations (Beullens and Loosveldt 2012). Interestingly, refusal to respond is the most commonly reported reason for high nonresponse across Western surveys (Brick and Williams 2013). This is promising if we believe that a potential respondent's willingness can be changed or manipulated by changing the survey environment. Immense effort is given to improving response rates (Sakshaug 2022) by increasing the contact attempts (Beebe et al. 2018), offering renumeration (Lipps 2010; McGonagle 2020), or defining the "hard-to-reach" for targeted follow-ups (Southern et al. 2008). Increasing effort is motivated by the belief that improving response rates reduces nonresponse bias. Empirically, it seems that the costs associated with these efforts is rarely justified by actual reductions in bias (Peytchev, Baxter, and Carley-Baxter 2009).

Should we abandon the response rate as an indicator of survey quality? We believe the extant literature does not offer enough evidence to conclusively answer this question. Groves and Peytcheva (2008)'s landmark study shows almost no relationship between nonresponse and nonresponse bias. Across 959 estimates pooled from 59 studies, there is only a weak correlation of around .20 and they conclude that a single indicator of survey quality such as response rates is misleading. Unfortunately, the veracity of their conclusion is undermined by a relatively small sample size that consequently limits them to univariate analyses. They claim further that most variation in bias occurs within-study without offering analysis of between-study variation. Actually, when Brick and Tourangeau (2017) re-analysed the data from Groves and Peytcheva (2008) they found a correlation of .49 between nonresponse and bias when the correlation is weighted by sample size. Another limitation of Groves and Peytcheva (2008) is that nonresponse is deemed a *cause* of nonresponse bias. Yet there is a disconnect between their causal theories of nonresponse bias and how the data is meta-analysed, as univariate analysis cannot capture complex interactions among survey characteristics).

Our study is unique in several aspects: (1) A greater sample size of empirical studies than previous studies; (2) More unique survey characteristics coded than previous studies; (3) Increased statistical power; and (4) a first application of network models that have a clearer link to causal theories. We consider how several survey characteristics influence nonresponse and bias and if these characteristics reveal how they are linked.

2 Background

2.1 Nonresponse Bias

Bias in an unadjusted mean or proportion is decided by the average response propensity and the correlation between individual response propensities and the variable of interest.

Taking Bethlehem (2010) 's definition of nonresponse bias, we have:

$$NR_{bias}(\bar{y}_r) = \frac{\sigma_{p_i,\mu_i}}{\bar{P}}, \tag{1}$$

where $\bar{y}_r$ is the value of the outcome variable, σ_{p_i,μ_i} is the covariance of the outcome variable and individual response propensity, and $\bar{P}$ is overall response propensity.

Bethlehem (2010)'s definition highlights the theoretical simplicity of how nonresponse and bias are related. If respondents and nonrespondents are homogenous with respect to the outcome bias is low. If they are heterogenous bias is high. This quantifies nonresponse bias as a predominantly within-study phenomenon. Leverage-saliency (Groves and Peytcheva 2008; Groves 2006; Groves et al. 2006) and response-propensity theory (Brick and Tourangeau 2017) explain how between-study variation is also a factor.

2.1.1 Connecting Nonresponse to Nonresponse Bias

According to leverage-saliency theory, both individual differences and survey content affect participation. Some are motivated to respond to shorter surveys while others are motivated by the topic. Motivating factors increase the saliency of the survey and may have high leverage on the participation decision. If these factors are associated with the outcome variable, statistics calculate on it are biased.

A scenario under response-propensity theory known as *correlated propensities model* also produces nonresponse bias (Brick and Tourangeau 2017). If response propensities correlate with survey design features and individual characteristics of the respondents, certain groups become overrespresented and outcomes are biased to them [1].

2.1.2 Connecting Theories to Causal Mechanisms

Leverage-saliency and response-propensity theory posit that nonresponse and nonresponse bias arise from interaction of within- and between-study factors. **Figure 1** displays three causal models for these theories Groves (2006).

Response propensity, P, and the outcome, Y can have *separate causes*, Z and X respectively (**A**); a *common cause*, Z (**B**); or a *survey variable cause* of response propensity, Y (**C**). Common causes and survey variable causes both produce nonzero covariance matrices for bias, eliciting indirect and direct effects.

[1] See Brick and Tourangeau (2017) for details on all scenarios under response-propensity theory.

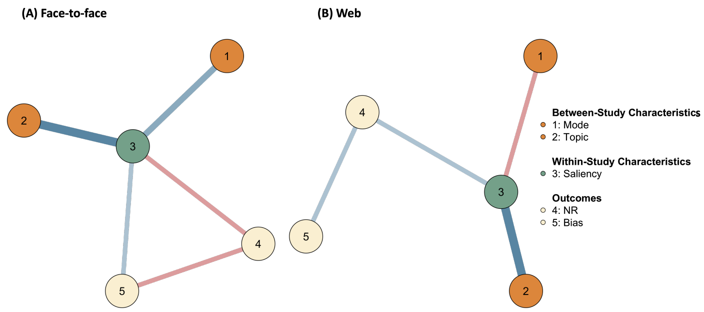

Figure 1: Causal models for nonresponse bias in surveys.

2.1.3 The Need for Nonresponse Analysis and A New Approach: Network Models

Declining response rates care a concern for survey researchers, who fear an accompanied rise in nonresponse bias. Yet previous studies fail to find this effect (Groves and Peytcheva 2008; Groves et al. 2006; Hedlin 2020; Beullens and Loosveldt 2012; Gundgaard et al. 2007). Nonresponse analysis is generally not well reported; just 30% reported it in one meta-analysis (Werner, Praxedes, and Kim 2007). This, together with the limitations of prior reports, indicate that we lack a full view on whether nonresponse leads to nonresponse bias. Traditional methods like multilevel regression or correlations don't have a clear link to the causal models in **Figure 1**.

In other fields, statistical network models are used to substantiate causal frameworks [2]. Networks visualize statistical conditional relations of many variables in a succinct and more informative way than regression can, identifying *potential* causal pathways and predictive relationships (Ryan, Bringmann, and Schuurman 2022; Epskamp et al. 2018). Assuming that the data generating process of nonresponse bias follows Bethlehem (2010)'s formula, we can estimate a network structure involving survey characteristics and the covariation to nonresponse and nonresponse bias. **Figure 2** illustrates this for survey topic and saliency in separate modes. Interviewers in face-to-face surveys can increase saliency by explicitly relating the topic to the individual, increasing response rates and reducing bias (**A**). There is no opportunity to increase saliency like this in web surveys, so re-

[2] Network models can only be considered explicitly as causal models be considered under strict circumstances (see Ryan, Bringmann, and Schuurman 2022).

sponses decline and bias increases (**B**). Expanding these examples to involve more survey characteristics will show us what common or survey causes affect response propensity and bias.

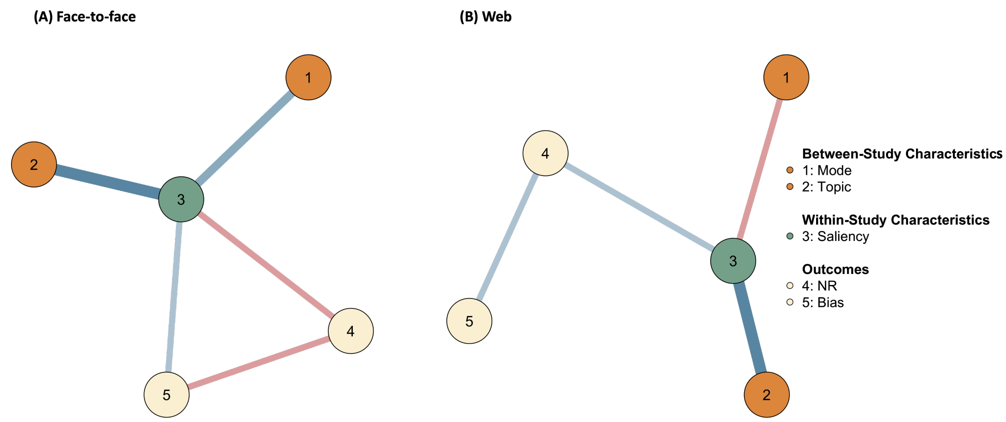

Figure 2: Example of how a network model could differ in face-to-face and web survey modes. Blue edges indicate positve edges and pink edges indicate negative edges.

3 Method

3.1 Aim

We aim to use statistical networks to describe the conditional relations between nonresponse and nonresponse bias, accounting for survey design. We initially conducted a systematic meta-analysis of literature reporting nonresponse and nonresponse bias in surveys. As it stands, this is the first application of a network approach to nonresponse analysis. Our results will hopefully inform survey designers who wish to minimize nonresponse bias.

3.2 Systematic Meta-Analysis

3.2.1 Eligibility Criteria and Search Strategy

Our search strategy was based on the strategy used in the original review by Groves and Peytcheva (2008). A list of key search terms was taken from Groves and Peytcheva

(2008) and expanded to include all relevant key words related to nonresponse rates and nonresponse bias. Some examples of these terms include "response bias", "nonresponse bias", "selection bias", "survey bias", and "participation rates". Articles were primarily identified from five electronic databases (Google Scholar, Web of Science, Scopus, Web Survey Methodology, and PsychInfo) in 2021. Searches of conference proceedings (e.g., American Statistical Association Survey Research Methods Section), articles cited in book chapters (e.g., the Wiley Series in Survey Methodology), and internet Google searches using the key search terms were also conducted to capture a wider range of articles and grey literature. Eligible articles met these criteria:
1. The study is empirical and about nonresponse rates 2. A probability-sample is drawn from a population 3. A survey is conducted 4. Sample frame information is collected on respondents and nonrespondents.

Articles originating from Groves and Peytcheva (2008) were automatically included. Panel studies were excluded in addition to articles that did not mention nonresponse rates and nonresponse bias. All articles were published between 1978 and 2020.

3.2.2 Screening phase

ASReview (v0.19.1) (Schoot et al. 2021) screened the literature from the initial search. A random sample of articles was coded as relevant or irrelevant by five independent coders. Coded articles served as training input to the ASReview which classified the remaining literature. ASReview results were divided across the coders who manually screened article abstracts or full-texts for quality until 25 consecutive articles were coded irrelevant. Articles incorrectly classified as irrelevant by ASReview were corrected before coding, resulting in 113 articles and 3301 estimates.

3.2.3 Coding phase

113 studies were divided across five independent coders instructed to record all relevant characteristics of the study. Discrepancies were resolved by discussion [3]. See Appendix: Table 1 for an overview.

[3] Kappa's Fleiss coefficient will be calculated when coding is complete to ensure interrator reliability.

3.2.4 Calculation of nonresponse rate

Nonresponse rate is defined as a percentage of all the potentially eligible units that do not respond to the items in a survey. Unit nonresponse rates were calculated in our study, where all data from a sampled respondent are missing.

Response rate is calculated as:

$$RR = 100 * \frac{n_R}{n_E}, \tag{2}$$

where n_R is the number of eligible units in the responding sample and n_E is the total number of eligible units invited to respond.

3.2.5 Calculation of nonresponse bias

Nonresponse bias consists of two components: nonresponse rates and the difference between respondents and nonrespondents on an estimate. We calculate the absolute relative bias in the same manner as Groves and Peytcheva (2008):

$$Bias(\hat{y}_r) = \left| \frac{100 * (\hat{y_r} - \hat{y_n})}{\bar{y_n}} \right|, \tag{3}$$

where $\hat{y_r}$ is the survey estimate for respondents, $\hat{y_n}$ is the survey estimate for nonrespondents, and p_n is the nonresponse rate.

3.3 Network Analysis

Network models are families of probability distributions that satisfy several conditional (in)-dependency statements represented in an undirected graph (Haslbeck and Waldorp 2015). An undirected graph $G = (V, E)$ consists of nodes $V = \{1, 2, ..., p\}$ and edges $E \subseteq V * V$. Nodes represent variables (survey characteristics) and edges connecting nodes represent statistical relations.

3.3.1 Mixed Graphical Models

Mixed Graphical Models (MGMs; (Haslbeck and Waldorp 2015; Chen, Witten, and shojaie 2015; Yang et al. 2014)) are probabilistic graphical models reflecting the joint probability density of multiple variables that follow different distributions. Each node is therefore associated with a different conditional exponential family distribution. Specifically, the pairwise interaction between two continuous variables, s and r, is given by a single parameter, $\theta_{s,r}$. $\theta_{s,r}$ represents if the interaction is non-zero. Pairwise interactions between two categorical variables, m and u are given by $R = (m - 1) * (u - 1)$ parameters associated with corresponding indicator functions (dummy variables) associated with R states. Pairwise interactions between a continuous and a categorical variable is given by $R = 1 * (m - 1)$ parameters associated with $(m - 1)$ corresponding indicator functions for all R states (Haslbeck and Waldorp 2020).

MGMs are constructed by factoring n univariate conditional members of the exponential family to a joint distribution [4] :

$$P_{PL}(\mathbf{Y} = \mathbf{y}) = \prod_{i=1}^{n} P(Y_i / \mathbf{Y}_{\setminus i}) \tag{4}$$

where n is the number of nodes and $\mathbf{Y}_{\setminus i}$ is the set of nodes without node i.

There are two steps in estimating the parameters for the MGM. First, the conditional distribution of each node is estimated separately using generalized linear regression. Since estimation is univariate, there are two estimates for each node. Both estimates are averaged in a second step into one network structure. We will specify an *"AND"* rule to retain only edges where both estimates are non-zero. A regularization parameter to prevent overfitting will be selected by the Extended Bayesian Information Criterion (EBIC):

$$EBIC = -2LL(\hat{\theta}) + Jlog(n) + 2\gamma Jlog(p - 1), \tag{5}$$

where $LL(\hat{\theta}$ is the log likelihood of the model, J is the number of parameters, p is the

[4]Factorization only leads to well-defined joint probability distributions under certain conditions (see Chen, Witten, and shojaie (2015)).

number of variables, and γ is the hyperparameter weighting the extra penalty $2Jlog(p-1)$.

A different network will be estimated for each survey mode. Bootstrapping can estimated the accuracy of the edge-weights and test for differences between them through a bootstrapped difference test (Epskamp, Borsboom, and Fried 2018).

Software

All analysis is conducted using R version 4.1.2. Estimation is implemented in the `mgm` package (Haslbeck and Waldorp 2020), network visualization by the `qgraph` package (Epskamp et al. 2012), and bootstrapping by the `bootnet` package (Epskamp, Borsboom, and Fried 2018).

4 Results

4.1 Descriptive Analysis

Figure 3 and **Figure 4** show the weighted correlation between nonresponse rates and nonresponse bias overall (top) and between-study means (bottom). Studies originally included by Groves and Peytcheva (2008) are distinguished by the orange points. New studies included in our analysis are the green points. A linear regression line is fitted for the studies by Groves and Peytcheva (2008) separately from the new studies. The overall weighted correlation is $R_{wtd} = 0.20$ for the old studies and $R_{wtd} = 0.23$ for the new studies. Our analysis of the between-study correlation of nonresponse rates with the mean nonresponse bias found $R_{wtd} = 0.47$ and $R_{wtd} = 0.40$ in the olde and new studies, respectively.

Table 1 displays the mean absolute relative bias, response rates, and sample size of the coded survey characteristics. Notably, the absolute relative bias is lowest in the mode with the highest response rate (face-to-face) and more in the modes with lower response rates (i.e.paper drop and web).

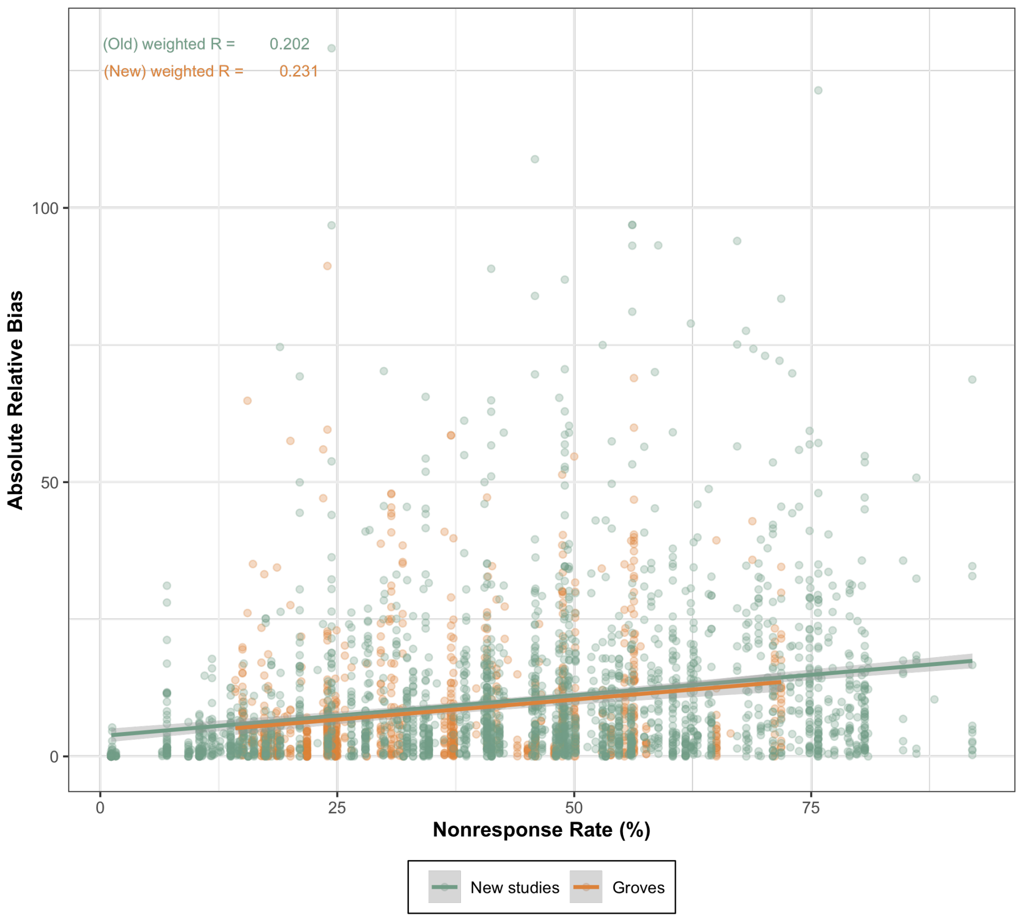

Figure 3: Correlation of nonresponse rates and nonresponse bias.

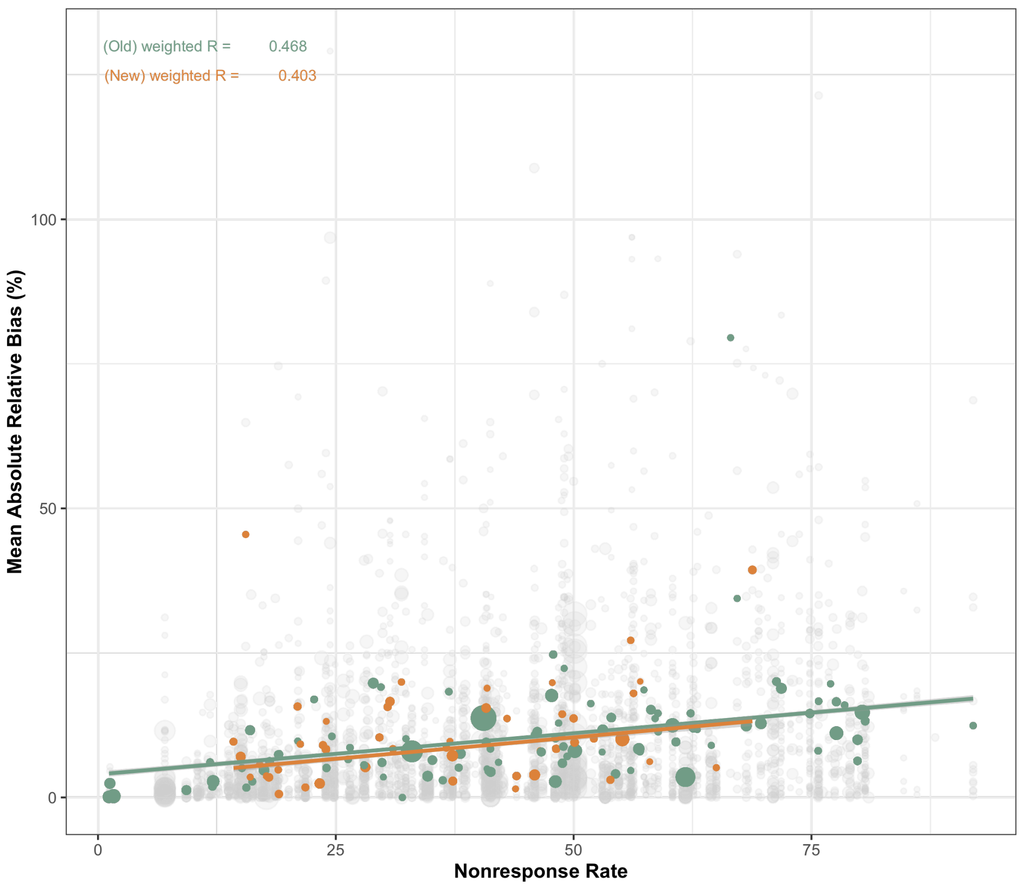

Figure 4: Between-Study correlation of nonresponse rates and nonresponse bias.

Table 1: Absolute Relative Bias and Nonresponse Rates across Survey Characteristcs

ID	Absolute Relative Bias	Response Rate	N
Mode			
Paper drop	11.13	44.77	14681
F2F	8.44	73.50	3324
Mail	8.98	58.92	11831
Telephone	11.14	57.43	15822
Web	12.67	47.95	30615
NA	10.49	47.83	13007
Topic			
Consumer satisfaction	11.69	52.03	7888
Crime	10.81	48.81	24981
Education	15.81	38.32	2218
Employment	10.42	50.47	8017
Finances	10.21	64.44	176274
General attitudes	5.60	65.41	1477
Health	9.40	59.99	15603
Living	12.82	64.53	1579
Omnibus	6.10	88.21	1506
Parenthood	13.21	72.77	2006
Safety	7.50	62.30	1851
Special interests	14.95	42.66	4291
Travel	2.96	63.75	3777
Voting	10.67	53.02	3724
NA	8.69	53.00	3318
Saliency			
Omnibus/Undetermined	8.85	64.95	6058
No	11.08	59.46	15871
Yes	9.19	58.10	12097
Incentives			
Yes for some respondents	13.11	45.87	34670
Yes	8.33	59.23	15072
No	9.92	59.02	12768
NA	10.84	49.91	9396
Reminders			
No	9.90	64.97	3941
Unknown	8.81	53.79	24152
Yes	9.82	58.04	13636
NA	11.69	46.94	10212
Statistic			
Mean	7.17	51.01	15762
Median	10.60	50.08	58124
Other	4.61	38.25	143731
Proportion	8.78	64.51	13061
Total	12.36	51.22	11554
Source			
Follow up	7.16	70.30	6495
Frame	11.38	52.17	16178
Intention to respond	7.06	67.19	851
Screener	8.84	55.09	2289

5 References

Andreski, Katherine A. Mcgonagle, Frank Stafford. 2012. "Response Rates in National Panel Surveys - Robert F. Schoeni, Frank Stafford, Katherine A. Mcgonagle, Patricia Andreski, 2013." *The ANNALS of the American Academy of Political and Social Science*, November. https://journals.sagepub.com/doi/full/10.1177/0002716212456363.

Beebe, Timothy J., Robert M. Jacobson, Sarah M. Jenkins, Kandace A. Lackore, and Lila J. Finney Rutten. 2018. "Testing the Impact of Mixed-Mode Designs (Mail and Web) and Multiple Contact Attempts Within Mode (Mail or Web) on Clinician Survey Response." *Health Services Research* 53 (Suppl Suppl 1): 3070–83. https://doi.org/10.1111/1475-6773.12827.

Bethlehem, Jelke. 2010. "Selection Bias in Web Surveys." *International Statistical Review* 78 (2): 161–88. https://doi.org/10.1111/j.1751-5823.2010.00112.x.

Beullens, Koen, and Geert Loosveldt. 2012. "Should High Response Rates Really Be a Primary Objective?" *Survey Practice* 5 (3): 1–5. https://doi.org/10.29115/SP-2012-0019.

Beullens, Koen, Geert Loosveldt, Caroline Vandenplas, and Ineke Stoop. 2018. "Response Rates in the European Social Survey: Increasing, Decreasing, or a Matter of Fieldwork Efforts?" *Survey Methods: Insights from the Field (SMIF)*, April. https://doi.org/10.13094/SMIF-2018-00003.

Brick, J. Michael, and Roger Tourangeau. 2017. "Responsive Survey Designs for Reducing Nonresponse Bias." *Journal of Official Statistics* 33 (3): 735–52. https://doi.org/10.1515/jos-2017-0034.

Brick, J. Michael, and Douglas Williams. 2013. "Explaining Rising Nonresponse Rates in Cross-Sectional Surveys." *The ANNALS of the American Academy of Political and Social Science* 645 (1): 36–59. https://doi.org/10.1177/0002716212456834.

Chen, Shizhe, Daniela M. Witten, and Ali shojaie. 2015. "Selection and Estimation for Mixed Graphical Models." *Biometrika* 102 (1): 47–64. https://doi.org/10.1093/biomet/asu051.

Cornesse, Carina, and Michael Bosnjak. 2018. "Is There an Association Between Survey Characteristics and Representativeness? A Meta-Analysis." *Survey Research Methods* Vol 12 (April): 1–13 Pages. https://doi.org/10.18148/SRM/2018.V12I1.7205.

Curtin, Richard, Stanley Presser, and Eleanor Singer. 2000. "The Effects of Response Rate Changes on the Index of Consumer Sentiment." *Public Opinion Quarterly* 64 (4): 413–28. https://doi.org/10.1086/318638.

Czajka, John L, and Amy Beyler. 2016. "Declining Response Rates in Federal Surveys: Trends and Implications." Office of the Assistant Secretary for Planning; Evaluation.

Epskamp, Sacha, Denny Borsboom, and Eiko I. Fried. 2018. "Estimating Psychological Networks and Their Accuracy: A Tutorial Paper." *Behavior Research Methods* 50 (1): 195–212. https://doi.org/10.3758/s13428-017-0862-1.

Epskamp, Sacha, Angélique O. J. Cramer, Lourens J. Waldorp, Verena D. Schmittmann, and Denny Borsboom. 2012. "**Qgraph** : Network Visualizations of Relationships in Psychometric Data." *Journal of Statistical Software* 48 (4). https://doi.org/10.18637/jss.v048.i04.

Epskamp, Sacha, Lourens J. Waldorp, René Mõttus, and Denny Borsboom. 2018. "The Gaussian Graphical Model in Cross-Sectional and Time-Series Data." *Multivariate Behavioral Research* 53 (4): 453–80. https://doi.org/10.1080/00273171.2018.1454823.

Groves, R. M. 2006. "Nonresponse Rates and Nonresponse Bias in Household Surveys." *The Public Opinion Quarterly* 70 (5): 646–75. https://www.jstor.org/stable/4124220.

Groves, R. M., Mick P. Couper, Stanley Presser, Eleanor Singer, Roger Tourangeau, Giorgina Piani Acosta, and Lindsay Nelson. 2006. "Experiments in Producing Nonresponse Bias." *Public Opinion Quarterly* 70 (5): 720–36. https://doi.org/10.1093/poq/nfl036.

Groves, R. M., and E. Peytcheva. 2008. "The Impact of Nonresponse Rates on Nonresponse Bias: A Meta-Analysis." *Public Opinion Quarterly* 72 (2): 167–89. https://doi.org/10.1093/poq/nfn011.

Gundgaard, J., O. Ekholm, E. H. Hansen, and N. Kr. Rasmussen. 2007. "The Effect of Non-Response on Estimates of Health Care Utilisation: Linking Health Surveys and Registers." *The European Journal of Public Health* 18 (2): 189–94. https://doi.org/10.1093/eurpub/ckm103.

Haslbeck, Jonas M. B., and Lourens J. Waldorp. 2015. "Structure Estimation for Mixed Graphical Models in High-Dimensional Data." arXiv. http://arxiv.org/abs/1510.05677.

———. 2020. “Mgm: Estimating Time-Varying Mixed Graphical Models in High-Dimensional Data.” arXiv. http://arxiv.org/abs/1510.06871.

Hedlin, Dan. 2020. “Is There a ’Safe Area’ Where the Nonresponse Rate Has Only a Modest Effect on Bias Despite Non-Ignorable Nonresponse?” *International Statistical Review* 88 (3): 642–57. https://doi.org/10.1111/insr.12359.

Keeter, Scott, Carolyn Miller, Andrew Kohut, R. M. Groves, and Stanley Presser. 2000. “Consequences of Reducing Nonresponse in a National Telephone Survey.” *The Public Opinion Quarterly* 64 (2): 125–48. https://www.jstor.org/stable/3078812.

Lipps, Oliver. 2009. “Attrition of Households and Individuals in Panel Surveys.” *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1367371.

———. 2010. “Effects of Different Incentives on Attrition and Fieldwork Effort in Telephone Household Panel Surveys.” *Survey Research Methods* 4 (2): 81–90. https://doi.org/10.18148/srm/2010.v4i2.3538.

McGonagle, Katherine A. 2020. “The Effects of an Incentive Boost on Response Rates, Fieldwork Effort, and Costs Across Two Waves of a Panel Study.” *Methoden, Daten, Analysen* 14 (2): 241–50. https://doi.org/10.12758/mda.2020.04.

Peytchev, Andy, Rodney K. Baxter, and Lisa R. Carley-Baxter. 2009. “Not All Survey Effort Is Equal: Reduction of Nonresponse Bias and Nonresponse Error.” *The Public Opinion Quarterly* 73 (4): 785–806. http://www.jstor.org/stable/40467642.

Ryan, Oisín, Laura F. Bringmann, and Noémi K. Schuurman. 2022. “The Challenge of Generating Causal Hypotheses Using Network Models.” *Structural Equation Modeling: A Multidisciplinary Journal* 29 (6): 953–70. https://doi.org/10.1080/10705511.2022.2056039.

Sakshaug, Joseph W. 2022. “Reducing Nonresponse and Data Linkage Consent Bias in Large-Scale Panel Surveys.” *Forum for Health Economics and Policy* 25 (1-2): 41–55. https://doi.org/10.1515/fhep-2021-0060.

Schoeni, Robert F., Frank Stafford, Katherine A. McGonagle, and Patricia Andreski. 2013. “Response Rates in National Panel Surveys.” *The Annals of the American Academy of Political and Social Science* 645 (1): 60–87. https://doi.org/10.1177/0002716212456363.

Schoot, Rens van de, Jonathan de Bruin, Raoul Schram, Parisa Zahedi, Jan de Boer, Felix

Weijdema, Bianca Kramer, et al. 2021. "An Open Source Machine Learning Framework for Efficient and Transparent Systematic Reviews." *Nature Machine Intelligence* 3 (2): 125–33. https://doi.org/10.1038/s42256-020-00287-7.

Southern, Danielle A., Steven Lewis, Colleen J. Maxwell, James R. Dunn, Tom W. Noseworthy, Gail Corbett, Karen Thomas, and William A. Ghali. 2008. "Sampling 'Hard-to-Reach' Populations in Health Research: Yield from a Study Targeting Americans Living in Canada." *BMC Medical Research Methodology* 8 (1): 57. https://doi.org/10.1186/1471-2288-8-57.

Werner, Steve, Moira Praxedes, and Hyun-Gyu Kim. 2007. "The Reporting of Nonresponse Analyses in Survey Research." *Organizational Research Methods* 10 (2): 287–95. https://doi.org/10.1177/1094428106292892.

Yang, Eunho, Yulia Baker, Pradeep Ravikumar, Genevera Allen, and Zhandong Liu. 2014. "Mixed Graphical Models via Exponential Families." In *Proceedings of the Seventeenth International Conference on Artificial Intelligence and Statistics*, 1042–50. PMLR. https://proceedings.mlr.press/v33/yang14a.html.

6 Appendix

Table 2: Overview of the survey characteristics coded in the meta-analysis

Variable	Description
N. Questions	Indicates the number of questions as an measure of survey length
Reminders	Indicates if the respondents received a reminder to participate (Yes/No)
Gender	Indicates the gender of the respondent (male/female)
Prenotification	Indicates if the the targeted population are notified about the survey prior to receiving it
Source	Indicates the source of the information on nonrespondents (sample frame, supplemental information, screener, intention to respond)
Max NC attempts	Indicates the number of attempts made to contact members of the target population who remained uncontactable
Populaton Type	Indicates if the population is special (pre-specified) or not (yes/no)
Sponsorship	Indicates If the survey is sponsorship by a third party (yes/no)
Incentives	Indicates If the respondents are offered an incentive for participating
Mode	Indicates the mode of the Survey (face-to-face, telephone, (e)mail, web)
Statistic type	Indicates if the outcome of interest is a mean, proportion, total, or median
Year	Indicates the year the corresponding article was published
Topic	Indicates the topic of the survey topic (collapsed to health/other)
Topic Saliency	Indicates If the topic is saliency to the respondent (yes/no)
Question type	Indicates if the question is about demographics, is observable, or is unobservable