

The SOM Institute's Notes on Survey Methodology – 2022:1

The effects of questionnaire length on response rate, non-response bias, and data quality

Frida Sandelin

frida.sandelin@som.gu.se

The SOM Institute, University of Gothenburg

ABSTRACT

The questionnaire length or the number of questions in surveys is an area of debate related to the response burden, response rates, and quality of the data. This note describes the results of an experiment investigating the effect the length of a questionnaire has on response rate, non-response bias, and data quality in a mixed mode, paper-and-pencil mail-back or web survey, administered to a random sample of residents in the city of Gothenburg in Sweden in 2020. The results indicated, in a context of relatively high response rates, that a longer questionnaire (12 pages of questions and 205 questions) statistically significantly lowered the response rate by 2.7 percentage points compared to a shorter questionnaire (9 pages of questions and 149 questions). However, the length of the questionnaire seemed only to reduce response propensities among individuals who choose to complete and send in the paper-and-pencil version of the questionnaire. No difference in response rate was detected among individuals who completed the web questionnaire. In addition, the analyses indicated that the longer questionnaire increased non-response bias in terms of age, and geographical settlement, but did not decrease data quality. The results support previous findings that longer questionnaires lead to declined response rates and larger non-response bias. On the other hand, these results contribute to the less common conclusion that questionnaire length does not necessarily decrease data quality.

INTRODUCTION

One area in survey methodology is the discussion about response burden and motivation for survey participation and data quality. Questionnaire length is one aspect proposed to relate to response burden. Longer questionnaires have been found to negatively affect response rates in paper-and-pencil mail-back surveys (Heberlein & Baumgartner, 1978; Yammarino et al., 1991; Dillman et al., 1993; Bogen, 1996). Findings in literature have further shown larger measurement error for questions placed toward the end compared to when placed at the beginning (Herzog & Bachman, 1981; Krosnick et al. 2002; Olson & Peytchev, 2007), resulting in increased risk for poorer data quality and increased non-response bias. However, the effect the questionnaire length has on web surveys is less certain.

Several studies have found no relationship between questionnaire length and participation in web surveys when sample individuals received no information about the survey length (Sheehan, 2001; Cook et al., 2000; Koitsalu et al., 2018;). Other studies have found that longer expected response length (questionnaire length communicated as the approximated time to complete the survey) decreased response rates (Crawford et al., 2001; Galesic & Bosnjak, 2009). Furthermore, longer questionnaires have also been linked to decreased data quality, where respondents who answered the longer questionnaire took less time, provided fewer words to open-ended questions, and answered questions more uniformly toward the end of the questionnaire (Crawford et al., 2001; Galesic & Bosnjak, 2009).

Keeping the questionnaire length as short as possible to meet high response rates and achieve good-quality data is a trade-off when the objectives and purpose of the questionnaire are to ask many different questions. The main concern for the SOM-Institute's annual mixed mode surveys (paper-and-pencil mail-back and web) is the uncertainty of the consequences if the institute would increase the length of a questionnaire. The questionnaire in the local Gothenburg SOM-survey has since 2017 had 9 pages of questions and achieved an average response rate of 50 percent (RR5, AAPOR, 2016). However, an enhanced need for expanding the survey to 12 question pages motivates a solid experiment to assess the magnitudes of such a decision. In addition, since the share of respondents who opt for the web questionnaires continues to increase, it is vital to investigate the effects of the questionnaire length of the paper-and-pencil mail-back and the web questionnaire both separately and together.

This note reports the results from an experiment in a mixed mode survey (paper-and-pencil mail-back and web) assessing the effects of questionnaire length on response rates, non-response bias, and data quality. The experiment was administrated to a random sample of inhabitants of the city of Gothenburg in the fall of 2020. The sample was randomized to receive a paper-and-pencil questionnaire of either 12 question pages (205 questions) or 9 question pages (149 questions). All sample persons were allowed to fill out the questionnaire online. The purpose of the experiment was to assess the possible consequences of increasing the questionnaire length from 9 (149 questions) to 12 pages (205 questions).

HYPOTHESIS

Three hypotheses were assessed:

RESPONSE RATES

H1a: Individuals who receive a questionnaire with 12 question pages (205 questions) may be less likely to complete the questionnaire than individuals who receive a questionnaire with 9 question pages (149 questions).

H1b: Individuals who receive a questionnaire with 12 question pages (205 questions) may be less likely to complete the questionnaire than individuals who receive a questionnaire with 9 question pages (149 questions) if they use the paper-and-pencil mail-back questionnaire.

H1c: Individuals who receive a questionnaire with 12 question pages (205 questions) may not be less likely to complete the questionnaire than individuals who receive a questionnaire with 9 question pages (149 questions) if they use the web questionnaire.

NON-RESPONSE BIAS

H2: Respondents and non-respondents may be less similar when they receive a questionnaire with 12 question pages (205 questions) than when they receive a questionnaire with 9 question pages (149 questions).

DATA QUALITY

H3: Individuals who receive a questionnaire of 12 question pages (205 questions) may be more likely to generate worse data quality than individuals who receive a questionnaire of 9 question pages (149 questions).

METHODS AND MATERIALS

SAMPLE

The sample was a simple random sample of 8,000 residents of the city of Gothenburg, 16-85 years old drawn by the Swedish Tax Authority on August 26, 2020. The experiment was part of a local survey (“The SOM-survey in Gothenburg 2020”) conducted by the SOM-Institute at the University of Gothenburg, Sweden.

PROCEDURE

Prior to being invited to complete the questionnaire, each respondent in the sample was randomly assigned into one treatment group ($n = 4\,000$) and one control group ($n = 4\,000$). The treatment group received a 12 pages long questionnaire (205 questions) while the control group received a 9 pages long questionnaire (149 questions). See Table 1 for an overview of the groups.

The questionnaire covered topics such as local political attitudes, leisure activities, and housing issues. The 12 question pages and the 9 questions pages included the same questions in the first part of the questionnaires, but the longer questionnaire included a few more questions about political attitudes and health in the middle part of the questionnaire and a longer section toward the end about parental issues and equality. Two pages of background questions finished both the shorter and the longer questionnaires.

The field period started on September 14, 2020, when a postal invitation to complete the survey through a self-administrated paper-and-pencil questionnaire or an online self-administrated questionnaire was sent to all sampled individuals. The full sample was also sent a pre-notification postcard a week prior to the postal invitation and a postcard thanking them for participating/reminding them to participate on field day 9. The thank you/reminder postcard included instructions on how to respond to the online self-administrated questionnaire. Individuals in both groups who had not yet responded or had refused to answer¹ were sent up to four postal reminders² (sent out on field days 24, 45, 59, and 77 respectively) and up to four reminders as text messages to their mobile telephones³ (sent out field day 17, 36, 52, 66). The content in the four postal reminders and the four text

¹ Individuals who refused to respond either informed the SOM-institute this by email or telephone, or by replying NO to text message two, three or four.

² The postal reminders included the survey, an information letter and a return envelope.

³ The first text message was sent out to 24 % of the sample.

messages was identical for both groups. The field period ended on January 11, 2021.

The approximate time to finish the survey (half an hour) was announced on the second page of the paper-and-pencil mail-back questionnaire and the first page after the respondent had logged in to the web survey. The announced time was identical regardless of the number of pages with questions.

The full sample was offered a scratch lottery ticket incentive (*Trisslott*, retail value 30 SEK) contingent upon completion of the survey. Since Swedish law prohibits sending lottery tickets to citizens under the age of 18, 16-17 years old were instead offered a gift card (*Sverigecheck*, retail value 50 SEK).

The experimental design is described in Table 1. The full field period is described in appendix 2.

Table 1. Experimental design

	Treatment <i>n</i> = 4 000	Control <i>n</i> = 4 000
Question pages	12 pages	9 pages
Pages in total	16 pages	12 pages
Number of questions	205 questions	149 questions

MEASURES AND ANALYSIS PLAN

RESPONSE RATES

To compare response rates between the groups, Response Rate 5 (RR5) was estimated (AAPOR, 2016). A two sample-test of proportion evaluating the difference in difference in RR5 between the treatment and control group was implemented to assess the effect of the longer questionnaire.

NON-RESPONSE BIAS

The parameters of four OLS-regressions equations were estimated and predicted the RR5 with the treatment variable, the independent variables of interest, and the interactions between them. A statistically significant negative parameter for the interaction means that the longer questionnaire increased non-response bias. The variables of interest were the individuals' sex, age, immigrant status (born in Sweden, born outside Sweden), and administrative area of living in Gothenburg. Males, younger cohorts, people born outside of Sweden, and residents in the administrative area of living Northeast (Nordost) and Hisingen have previously

been found to be less likely to complete the questionnaire (Falk et al. 2020). Hence, reducing the response rate in these groups will further exacerbate non-response bias.

DATA QUALITY

Data quality was assessed based on three concepts:

1. *Item non-response*: This indicator measured the proportion of missing answers in each questionnaire. Missing answers were defined as a lack of an answer to a specific question, whereas erroneous answers or non-sensical answers were still considered an answer.
2. *Breakoffs/Partials/Completes*: This indicator measured the proportion of answers of a maximum of 50 percent of all questions included in each questionnaire (*Breakoffs*), the proportion of answers between 50 and 79.9 percent of all questions included in each questionnaire (*Partials*) and the proportion of answers of 80 percent or more of all questions included in each questionnaire (*Completes*). As for item non-response, all answers from the survey were included in this analysis. A multinomial logistic regression with corresponding average marginal effects was estimated to evaluate if the probability to generate breakoffs and partials was higher with the longer questionnaire. In addition, the regressions were repeated separated between modes (paper-and-pencil mail-back and web).
3. *Concurrent validity*: This measure evaluated the relationship of one variable with another variable known to be related to each other. The pairs of indicators were chosen based on a known theoretical correlation and their location in the questionnaire. In each pair, one variable was chosen from the first or middle part of the questionnaire, and the other variable was chosen from the last part of the questionnaire. OLS regression models predicted if an expected correlation statistically significantly differed between the shorter and the longer questionnaire.

RESULTS

RESPONSE RATES

A two-sample test of proportions revealed that sample persons who received the questionnaire with 12 questions pages were statistically significantly less likely to complete the questionnaire (RR5 = 45.8%) than those who received 9 questions pages (RR5 = 48.5%) ($z = -2.41, p = 0.016$). Hence, the difference in response rate

between the 12 pages and the 9 pages questionnaire was 2.7 percentage points (see Table 2).

Table 2. Effects of the longer questionnaire on Response Rate (Proportions, Difference of Proportions)

	Proportion	<i>n</i>	Standard error	z-value	95% CI lower	95% CI higher
Treatment group	0.4585	3,978	.0079		.4430	.4740
Control group	0.4855	3,975	.0079		.4700	.5011
Difference of proportions	-0.0270*		.0112	-2.41	-.0051	-.0489

Note. Response rates were calculated in accordance with AAPOR 2016-standard (RR5). RR5 excludes some groups from the original sample which explains the somewhat reduced the sample sizes for the groups. The test was conducted by the two-sample test of proportion (prtest) in Stata 17 SE.

As previous findings in the literature suggested, it is motivated to analyze the response rates for individuals who used the mail-back questionnaire and the web questionnaire separately. Table 3 presents the results from the two different two-sample tests of proportions for each mode.

Table 3. Survey mode (paper or web) effects of the longer questionnaire compared on Response Rate (Proportions, Difference of proportions)

	Proportion	<i>n</i>	Standard error	z-value	95% CI lower	95% CI higher
Mail-back questionnaire						
Treatment group	.3469	3,978	.0077		.3321	.3617
Control group	.3753	3,975	.0075		.3603	.3904
Difference of proportions	-.0284*		.0108	-2.64	-.0073	.0495
Web questionnaire						
Treatment group	.1116	3,978	.0050		.1018	.1214
Control group	.1102	3,975	.0050		.1005	.1200
Difference of proportions	.0014		.0070	0.20	-.0152	.0124

Note. Response rates were calculated in accordance with AAPOR 2016-standard (RR5). The test has been conducted by the two-sample test of proportion (prtest) in Stata 17 SE.

The share who sent in and completed the paper-and-pencil mail-back questionnaire respondents was 34.7% in the treatment group and 37.5% in the control group ($z = -2.64$, $p = 0.008$). In other words, the longer questionnaire statistically significantly decreased the response rate among paper-and-pencil

respondents by 2.8 percentage points. In contrast, among the individuals who completed the web questionnaire, no difference was detected between the treatment group and the control group. The share who completed the questionnaire online in the treatment group was 11.1% and in the control group 11.0% ($z = 0.20, p = 0.840$). Hence, the difference in response rate between the 12 question pages and the 9 question pages appeared to only be related to the paper-and-pencil questionnaire.

NON-RESPONSE BIAS

The estimations of non-response bias detected that the only subgroups that were not negatively affected by the longer questionnaire were 30–49 year olds and individuals living in the City Centre of Gothenburg (see Figure 1).

Figure 1. *Effects of the longer questionnaire on response rate and age and area of living (percentage)*

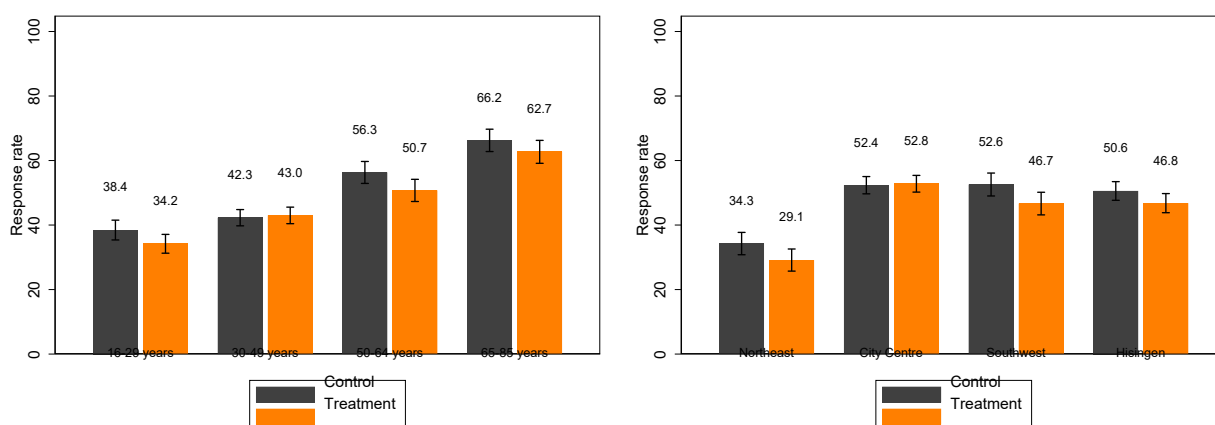


Table 4 further confirms these descriptive results with OLS regression estimates of the interactions between the treatment variable and the separate independent dummy variables. All interaction estimations were conducted by controlling for sex, age, immigrant status, and area of living. The estimations of the two-way interactions with age in Model 3 showed that using 30–49 as the reference dummy variable generated a statistically significant negative two-way interaction between the treatment group and the cohort 50–64 ($p = 0.036$) and for the two-way interaction between the treatment group and the cohort 16–29 ($p = 0.039$). The results for the area of living demonstrated the same pattern for the area of *the City Centre* used as a reference dummy variable in Model 5. The results showed marginally statistically significant two-way interactions between the treatment group and *Southwest* ($p = 0.09$), *Northeast* ($p = 0.074$), and *Hisingen* ($p = 0.08$). The estimations of the two-way interactions with sex and immigrant status were not

statistically significant. To conclude, as 30–49-year-olds and people living in the City Centre are not subgroups that experience low relative response rates in the survey in general, the longer questionnaire increased non-response bias.

Table 4. Effects of a longer questionnaire on response and sex, age, immigrant status, and area of living (OLS regression coefficients)

	Response				
	Baseline	Sex- interaction	Age- interaction	Immigrant- interaction	Area of living interaction
	Model 1	Model 2	Model 3	Model 4	Model 5
Treatment (reference: control)	-.03* (.01)	-.02 (.02)	.01 (.02)	-.03* (.02)	.00 (.02)
Treatment * Female (reference: male)		-.00 (.02)			
Treatment * Age: 16- 29 (reference: 30-49)			-.06* (.03)		
Treatment * Age: 50- 64 (reference: 30-49)			-.06* (.03)		
Treatment * Age: 65- 85 (reference: 30-49)			-.04 (.03)		
Treatment * Born outside Sweden (reference: Born in Sweden)				.00 (.02)	
Treatment * Area of living: The Northeast (reference: The City Centre)					-.06+ (.03)
Treatment * Area of living: The Southwest (reference: The City Centre)					-.05+ (.03)
Treatment * Area of living: Hisingen (reference: The City Centre)					-.05+ (.03)
Constant	.39*** (.02)	.39*** (.02)	.47*** (.02)	.38*** (.02)	.42*** (.03)
Observations	7,945	7,945	7,945	7,945	7,945
R^2	.11	.11	.11	.11	.11

Note. All models were estimated with sex, age, immigrant status and area of living as control variables but excluded in the table for readability. For the full tables, see the Appendix. Unstandardized OLS regression coefficients with standard errors in parentheses.

+ $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$.

DATA QUALITY

To investigate if the longer questionnaire had an effect on the quality of the data *Item non-response*, *Breakoffs/Partials/Completes*, and *Concurrent validity* were evaluated.

1. *Item non-response*

Table 5 presents the results from an OLS regression predicting the effect of the linear variable *item non-response*, that is, the proportion of missing answers for each response in the treatment group and the control group. The results showed that the proportion of missing answers was not statistically significantly different in the longer questionnaire compared to the shorter questionnaire (Model 1). Adding an interaction of mode made no difference to these results (Model 2).

Table 5. Effects of item non-response (OLS regression coefficients)

	Baseline	Mode control
	Model 1	Model 2
Treatment (reference: control)	.003 (.003)	.010 (.007)
Paper (reference: web)		.039*** (.006)
Treatment * Paper (reference: web)		-.008 (.008)
Constant	.949*** (.002)	.916*** (.005)
Observations	3,956	3,956
R^2	.000	.020

Note. + $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$.

2. *Breakoffs/Partials/Completes*

Respondents missing answers can further be analyzed as different proportions of answers in each questionnaire in terms of breakoffs, partials, and completes. Table 6 presents the results from a multinomial logistic regression predicting the probability of the treatment group yielding a higher share of breakoffs, partials, and completes respectively. The results showed that the share of breakoffs, partials, and completes were not statistically significantly different in the longer questionnaire compared to the shorter questionnaire.

To control for possible differences in probabilities of breakoffs/partial/completes between the longer and shorter questionnaire and mode (paper-and-pencil mail-back and web) multinomial logistic regressions were conducted for each of the two different modes. The separate analyses for mode detected no differences

between breakoffs/partial/completes across the modes. The table with the separate regressions for mode can be found in Appendix (Table A1).

Table 6. Effects of a longer questionnaire on the share of breakoffs and partials (Multinomial logistic regression coefficients)

	dy/dx	z-value	95% CI lower	95% CI higher
Breakoffs	.0013 (.0040)	0.32	-.0066	.0092
Partials	.0062 (.0058)	1.07	-.0052	.0176
Completes	-.0075 (.0070)	-1.07	-.0213	.0062

Note. Average marginal effects of the treatment group on breakoffs, partials and completes with standard errors in parentheses. The main effects of the multinomial logistic regression breakoff, partial and complete variables were included in the regression but were dropped from the table for readability.

3. Concurrent validity

The criterion variable was yearly household income, a question that demanded both time and cognitive ability to answer accurately. The chosen independent variables that should be related to the dependent variable yearly household income were visiting restaurants in the evening time (*Visit restaurants*), life satisfaction (*Life satisfaction*), and being married or living together with a partner (*Married/living with a partner*). The results are presented in Table 7 and indicated that the longer questionnaire did not decrease concurrent validity.

Table 7. Effects of a longer questionnaire on concurrent validity (OLS regression coefficients)

	Visit restaurants	Life satisfaction	Married/ living with a partner
	Model 1	Model 2	Model 3
Household income	.09*** (.01)	-.05*** (.00)	.07*** (.00)
Treatment	.07 (.12)	-.08+ (.04)	.00 (.03)
Treatment * Household income	-.00 (.02)	.01 (.01)	-.00 (.00)
Constant	3.31*** (.09)	2.02*** (.03)	.17*** (.02)
Observations	3,535	3,536	3,549
R^2	.03	.06	.21

Note. Effects of a longer questionnaire on concurrent validity through the variables visit restaurants, life satisfaction and married/living with a partner on the interaction between the treatment group and the variable household income (OLS regression coefficients). + $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$.

In addition, the three regression models were further analyzed separated by mode, with the same null-finding. These tables can be found in Appendix (Table A2). In other words, the concurrent validity measure indicated no inferior data quality for the longer questionnaire compared to the shorter.

CONCLUSION

This note described the results from an experiment investigating the effect of questionnaire length on response rate, non-response bias, and data quality. The SOM Institute has for several years been using different lengths of questionnaires in their different mixed-mode surveys (paper-and-pencil mail-back and web questionnaire). However, it has not yet been possible to draw any clear conclusions about deciding to increase the length of the questionnaires. To evaluate the effects of the difference in questionnaire length between 9 question pages (149 questions) and 12 question pages (205 questions), an experiment was administrated in the fall of 2020. The results showed that the response rate (RR5) among those who got the longer questionnaire was 45.8% compared to 48.5% among those who got the shorter questionnaire, a statistically significant decreased response rate of 2.7 percentage points.

Furthermore, the effect of the reduced response rate appeared to be subject only to the paper-and-pencil mail-back questionnaire as no difference in response rate between the length of the questionnaires was detected among web respondents.

In addition, the results suggested support for an increased non-response bias among respondents of the longer questionnaire. Individuals aged 30-49 years were not statistically significantly negatively affected in terms of response rate and people living in the central area of the city of Gothenburg were marginally statistically significantly not affected, two groups that already were more represented among respondents than their comparative counterparts. On the other hand, immigrant status and sex were not statistically significant factors in terms of reduced representativeness. Regarding data quality, no measures used in this note indicated any decline in data quality in the longer questionnaire compared to the shorter questionnaire.

In part, the results are similar to the literature in the sense that longer questionnaires seemed to result in a lower response rate. Since the decrease in response rate was only related to the paper-and-pencil mail-back respondents, this might be related to the fact that respondents in this mode could grasp the extent of the length and effort to respond by holding the questionnaire in their hands. At the same time, in contrast to what the literature suggests, the data quality was not

affected, regardless of in which mode the respondent chose to complete the questionnaire. In conclusion, although shorter surveys and fewer questions resulted in better general outcomes, there will always be a tradeoff with the need and demands for survey practitioners.

REFERENCES

- Bogen, K. (1996). The effect of questionnaire length on response rate: A review of the literature. United States Census Bureau.
- Cook, C., Heath, F., & Thompson, R. L. (2000). A Meta-Analysis of Response Rates in Web- or Internet-Based Surveys. *Educational and Psychological Measurement*, 60(6), pp. 821–836.
- Crawford, S. D., Couper, M. P., & Lamias, M. J. (2001). Web Surveys: Perceptions of Burden. *Social Science Computer Review*, 19(2), 146–162.
- Diehr, P., Chen, L., Patrick, D et al. (2008). Reliability, effect size, and responsiveness of health status measures in the design of randomized and cluster-randomized trials. *Contemp Clin Trials*, 26 (2005), pp. 45-58.
- Dillman, D., Sinclair, M.D. & Clark., J.R (1993). Effects of questionnaire length, respondent friendly design, and a difficult question on response rate for occupant-addressed census mail surveys. *Public Opinion Quarterly* 57:3, pp. 289-304.
- Falk, E., Sandelin, F. & Weissenbilder, M. (2020). SOM-undersökningen i Göteborg 2019 – En metodöversikt. University of Gothenburg. The SOM-Institute.
- Galesic, M. & Bosnjak M. (2009). Effects of Questionnaire Length on Participation and Indicators of Response Quality in a Web Survey. *Public Opinion Quarterly* 73:2.
- Heberlein, T. A., & Baumgartner, R. (1978). Factors affecting response rates to mailed questionnaires: A quantitative analysis of the published literature. *American Sociological Review*, 43(4), pp. 447–462.
- Herzog, A. R., & Bachman, J. G. (1981). Effects of questionnaire length on response quality. *Public Opinion Quarterly*, 45(4), 549–559.
- Koitsalu, M., Eklund, M., Adolfsson, J., Grönberg H & Brnadberg Y. (2018). Effects of pre-notification, invitation length, questionnaire length and reminder on participation rate: a quasi-randomised controlled trial. *BMC Medical Research Methodology* 18:3.

- Krosnik et al. (2002). The Impact of "No Opinion" Response Options on Data Quality: Non-Attitude Reduction or an Invitation to Satisfice? *Public Opinion Quarterly*, 66:3, pp. 371-403.
- Olson, K., & Peytchev, A. (2007). Effect of interviewer experience on interview pace and interviewer attitudes. *Public Opinion Quarterly*, 71(2), 273–286.
- Sheehan, K.B. (2001). E-mail Survey Response Rates: a Review, *Journal of Computer-Mediated Communication*, 6:2 2.
- Snyder, C.F., Watson, M.E., Jackson J.D et al. (2007). Patient-reported outcome instrument selection: designing a measurement strategy. *Value Health*, 10 (Suppl. 2) (2007), pp. S76-S85.
- Yammarino, Francis J., Skinner, Steven J & Childers, Terry L. (1991). Understanding mail survey response behaviour a meta-analysis. *Public Opinion Quarterly*, 55:4, pp. 613-639.

APPENDIX

Table A1. *Effects of a longer questionnaire on the share of breakoffs and partials (Multinomial logistic regression coefficients on mail back and web questionnaire separately)*

	dydx	n	z-value	95% CI lower	95% CI higher
Mail back questionnaire					
Breakoffs	.0023 (.0025)	13	0.92	-.0027	.0073
Partials	.0077 (.0067)	104	1.14	-.0055	.0209
Completes	-.0100 (.0071)	2,872	-1.40	-.0241	.0040
Web questionnaire					
Breakoffs	-.0047 (.0145)	52	-0.32	-.0332	.0238
Partials	.0017 (.0117)	33	.15	-.0212	.0246
Completes	.0029 (.0182)	882	.16	-.0327	.0387

Note. Average marginal effects of the treatment group on breakoffs, partials and completes with standard errors in parentheses. The effects were predicted with the treatment group as value 1 and control group as value 0. The main effects of the multinomial logistic regression breakoff, partial and complete variables were included in the regression but were dropped from the Table for readability.

+ $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$.

Table A2a (baseline). *Effects of a longer questionnaire on response and sex, age, immigrant status, and area of living (OLS regression coefficients)*

	Baseline
	Model 1
Treatment (reference: control)	.03* (.01)
Female (reference: male)	.03*** (.01)
Age: 16-29 (reference: 30-49)	.10*** (.01)
Age: 50-64 (reference: 30-49)	.19*** (.02)
Age: 65-85 (reference: 30-49)	.27*** (.02)
Born outside Sweden (reference: Born in Sweden)	-.25*** (.01)
Area of living: The Northeast (reference: The Centre)	-.10*** (.02)
Area of living: The Southwest (reference: The Centre)	-.04** (.01)
Area of living: Hisingen (reference: The Centre)	-.01 (.01)
Constant	.45*** (.02)
Observations	7,945
R^2	.11

Note. Unstandardized OLS regression coefficients with standard errors in parentheses.
+ $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$.

Table A2b (sex interaction). *Effects of a longer questionnaire on response and sex, age, immigrant status, and area of living (OLS regression coefficients)*

	Sex
	Model 2
Treatment (reference: control)	.03* (.01)
Treatment * Female (reference: male)	-.00 (.02)
Female (reference: male)	.04** (.01)
Age: 16-29 (reference: 30-49)	.10*** (.01)
Age: 50-64 (reference: 30-49)	.19*** (.02)
Age: 65-85 (reference: 30-49)	.27*** (.02)
Born outside Sweden (reference: Born in Sweden)	-.25*** (.01)
Area of living: The Northeast (reference: The Centre)	-.10*** (.02)
Area of living: The Southwest (reference: The Centre)	-.04** (.01)
Area of living: Hisingen (reference: The Centre)	-.01 (.01)
Constant	.45*** (.02)
Observations	7,945
R^2	.11

Note. Unstandardized OLS regression coefficients with standard errors in parentheses.
+ $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$.

Table A2c (age interaction). *Effects of a longer questionnaire on response and sex, age, immigrant status, and area of living (OLS regression coefficients)*

	Age
	Model 3
Treatment (reference: control)	.01 (.02)
Treatment * Age: 16-29 (reference: 30-49)	-.06* (.03)
Treatment * Age: 50-64 (reference: 30-49)	-.06* (.03)
Treatment * Age: 65-85 (reference: 30-49)	-.04 (.03)
Female (reference: male)	.03** (.01)
Age: 16-29 (reference: 30-49)	-.07*** (.02)
Age: 50-64 (reference: 30-49)	.12*** (.02)
Age: 65-85 (reference: 30-49)	.19*** (.02)
Born outside Sweden (reference: Born in Sweden)	-.27*** (.01)
Area of living: The Northeast (reference: The Centre)	-.10*** (.02)
Area of living: The Southwest (reference: The Centre)	-.04** (.01)
Area of living: Hisingen (reference: The Centre)	-.01 (.01)
Constant	.53*** (.02)
Observations	7,945
R^2	.11

Note. Unstandardized OLS regression coefficients with standard errors in parentheses.
+ $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$.

Table A2d (immigrant interaction). *Effects of a longer questionnaire on response and sex, age, immigrant status, and area of living (OLS regression coefficients)*

	Immigrant status
	Model 4
Treatment (reference: control)	-.03* (.01)
Treatment * Born outside Sweden (reference: Born in Sweden)	.00 (.02)
Female (reference: male)	.03** (.01)
Age: 16-29 (reference: 30-49)	.10*** (.01)
Age: 50-64 (reference: 30-49)	.19*** (.02)
Age: 65-85 (reference: 30-49)	.27*** (.02)
Born outside Sweden (reference: Born in Sweden)	-.25*** (.02)
Area of living: The Northeast (reference: The Centre)	-.10*** (.02)
Area of living: The Southwest (reference: The Centre)	-.04** (.01)
Area of living: Hisingen (reference: The Centre)	-.01 (.01)
Constant	.46*** (.02)
Observations	7,945
R^2	.11

Note. Unstandardized OLS regression coefficients with standard errors in parentheses.
+ $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$.

Table A2e (area of living interaction). Effects of a longer questionnaire on response and sex, age, immigrant status, and area of living (OLS regression coefficients)

	Area of living
	Model 5
Treatment (reference: control)	-.00 (.02)
Treatment * Area of living: The Northeast (reference: The Centre)	-.06+ (.03)
Treatment * Area of living: The Southwest (reference: The Centre)	-.05+ (.03)
Treatment * Area of living: Hisingen (reference: The Centre)	-.05+ (.03)
Female (reference: male)	.03** (.01)
Age: 16-29 (reference: 30-49)	.10*** (.01)
Age: 50-64 (reference: 30-49)	.19*** (.02)
Age: 65-85 (reference: 30-49)	.27*** (.02)
Born outside Sweden (reference: Born in Sweden)	-.25*** (.01)
Area of living: The Northeast (reference: The Centre)	-.07** (.02)
Area of living: The Southwest (reference: The Centre)	-.01 (.02)
Area of living: Hisingen (reference: The Centre)	-.02 (.02)
Constant	.44*** (.02)
Observations	7,945
R^2	.11

Note. Unstandardized OLS regression coefficients with standard errors in parentheses.
+ $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$.



The SOM Institute is an academic organization located at the University of Gothenburg, Sweden. Since 1986 the SOM Institute conduct annual cross-sectional surveys among the Swedish population with a focus on Society, Opinion and Media, as well as administering the web panel called the Swedish Citizen Panel. The annual surveys and the web panel both function as infrastructures, enabling researchers and public organizations to effectively collect research and opinion data in collaboration with researchers at the SOM Institute.

In order to strengthen contemporary research on Swedish society, as well as to contribute to international methodological development, the SOM Institute frequently publish notes on methodological research.

These are made available for scholars and the public alike at

<https://www.gu.se/en/som-institute/publications/research-on-survey-methodology>

For more information, please contact us at: info@som.gu.se