
SHARE WORKING PAPER SERIES

Switching Survey Mode Between CAPI and CAWI: A Report from the Field

Marika de Bruijne and Adriaan Kalwij

Working Paper Series 93-2024

DOI: [10.17617/wp.932024](https://doi.org/10.17617/wp.932024)

SHARE-ERIC | Leopoldstr. 139 | 80804 Munich | Germany | share-eric.eu

About the SHARE Working Paper Series

The SHARE Working Paper Series started in 2011 and collects pre-publication versions of papers or book chapters, technical and methodological reports as well as policy papers based on SHARE data. The working papers are not reviewed by the publisher (SHARE-ERIC), layout and editing are not standardized. The publisher takes no responsibility for the scientific content of the paper. Working Papers can be updated – a version number is indicated on the front page. Previous versions are available upon request.

Acknowledgements

The SHARE data collection has been funded by the European Commission, DG RTD through FP5 (QLK6-CT-2001-00360), FP6 (SHARE-I3: RII-CT-2006-062193, COMPARE: CIT5-CT-2005-028857, SHARELIFE: CIT4-CT-2006-028812), FP7 (SHARE-PREP: GA N°211909, SHARE-LEAP: GA N°227822, SHARE M4: GA N°261982, DASISH: GA N°283646) and Horizon 2020 (SHARE-DEV3: GA N°676536, SHARE-COHESION: GA N°870628, SERISS: GA N°654221, SSHOC: GA N°823782, SHARE-COVID19: GA N°101015924) and by DG Employment, Social Affairs & Inclusion through VS 2015/0195, VS 2016/0135, VS 2018/0285, VS 2019/0332, and VS 2020/0313. Additional funding from the German Ministry of Education and Research, the Max Planck Society for the Advancement of Science, the U.S. National Institute on Aging (U01_AG09740-13S2, P01_AG005842, P01_AG08291, P30_AG12815, R21_AG025169, Y1-AG-4553-01, IAG_BSR06-11, OGHA_04-064, HHSN271201300071C, RAG052527A) and from various national funding sources is gratefully acknowledged (see www.share-project.org).

Switching survey mode between CAPI and CAWI: a report from the field

Marika de Bruijne¹, Adriaan Kalwij

Abstract

In this study, we investigate mode effects on response and data quality when switching between face-to-face interviews (CAPI) with self-completion web surveys (CAWI). We first provide an overview of earlier literature on the characteristics of these two survey modes regarding coverage, response, and survey measurement. Next, we report our findings from the field based on data from the Netherlands of the longitudinal Survey of Health, Ageing and Retirement in Europe. In the Netherlands, six of the of the first eight waves were conducted in CAPI mode and two were in CAWI mode. Our main findings are that the mode switch from CAPI to CAWI resulted in a lower unit response, especially in the oldest age group. Also, respondents provided more substantive answers and remained more motivated during the face-to-face interview than the web survey. Furthermore, the web survey respondents reported more often to have mental health issues, which suggests less socially desirable response behavior in the self-administered mode. These findings support the presence of survey mode effects on unit and item response and data quality.

Keywords: survey modes, face-to-face, web surveys, CAPI, CAWI, mode effects, response

1. Introduction

Traditional face-to-face data collections are increasingly often set against web surveys which are expected to have lower costs and quicker lead time for data collection. However, there are concerns on how a potential mode switch may influence research outcomes based on the data collected. In this study, therefore, we compare the two modes and their effects on response and data quality using data from the Survey of Health, Ageing and Retirement in Europe (SHARE) in the Netherlands. After five waves the survey mode for SHARE Netherlands was switched from face-to-face interviews to self-completion web interviews, and back to face-to-face interviews for Wave 8.

Already in the late 1970's, the term survey 'mode' became a point of interest in survey literature (Couper, 2011) and the potential differences in results caused by using different modes were acknowledged (Groves, 1979). When switching survey modes, it can for several reasons affect

¹ Corresponding author: marika.de.bruijne@centerdata.nl

the results of data analyses. The response rate can suffer or the net sample distribution can differ, the latter becoming a concern when the answers to survey questions of subgroups differ. Furthermore, respondents can provide different answers merely due to the differences in the presentation of the survey and the interview situation. To investigate what kind of biases can be expected when switching from personal interviews (i.e. face-to-face interviews or CAPI: Computer-Assisted Personal Interviewing) to self-completion web interviews (i.e. CAWI: Computer-Assisted Web Interviewing), we first discuss the mechanisms of these modes and their differences in Section 2 and, next, survey measurement in Section 3 for these modes. In Section 4 we report our findings from the field based on data from SHARE for the Netherlands. Section 5 concludes.

2. Reaching the respondent: coverage and response

Face-to-face and web surveys can differ in the extent to which researchers are able to reach respondents and motivate them to participate. Although face-to-face interviewing can be very costly, this mode has the highest potential regarding coverage and sampling (De Leeuw, 2012). In web surveys, coverage error is still a concern in many countries where internet access coverage is low, because those without internet access are overrepresented in specific subgroups of the population such as older people, people with lower education or income, and minorities (De Leeuw, Hox, & Dillman, 2012). One solution to the coverage problem with web surveys is to provide respondents without internet access with computers and internet access, as has been done in some probability-based web panels (e.g. Scherpenzeel & Das, 2010), but this approach is financially costly.

Even when people have access to the Internet, some may still be unwilling to respond to web questionnaires (Millar, O'Neill, & Dillman, 2009). While web interviewing brings about advantages in terms of cost savings and timeliness (Jäckle, Lynn, & Burton, 2015), face-to-face interviewing generally still results in a higher response rate (Burton & Jäckle, 2020). Earlier studies have also identified a generation gap in preferences for modes. It has been reported that older individuals are more likely to prefer interviewer-administered modes over the self-administered modes of mail and web (Smyth, Olson and Millar, 2014). Further, many studies have shown that especially the younger respondents prefer the web as the response mode (Diment & Garret-Jones, 2007; Millar et al., 2009; Smyth et al, 2014; Revilla, 2015) and that older respondents are more likely to prefer traditional modes and personal interviews than younger respondents (Mulder & de Bruijne, 2019). This is likely because of familiarity with and access to the mode, factors that have been found to be strong predictors of mode preference (Smyth et al, 2014). As internet usage continues to rise and the digitally engaged generations grow older, this age bias may abate over time when entire populations reach internet maturity.

Many experiments have been conducted to combine face-to-face and web modes, with the expectation of having the benefits of both. The results have been mixed. For example, in an experiment in the Understanding Society Innovation Panel, an existing longitudinal panel in which respondents had been interviewed face-to-face, a mixed-mode design of offering web

first and only then face-to-face, resulted in a smaller proportion of households fully responding, lower individual response rates, and higher item nonresponse rates (Jäckle et al., 2015). In another experiment, of the European Social Survey (ESS), the existing face-to-face strategy was compared to a sequential mixed-mode strategy (web first and follow-up by face-to-face) in two countries. While no significant differences in the final response rates were found in Estonia, in the UK the response rate was significantly higher in the face-to-face mode (Revilla, 2015).

When switching from one survey mode to another, the mere switching can threaten the response rate regardless of the modes. When people have completed a questionnaire via a specific mode and thus are more experienced in this specific mode, they are more likely to prefer that mode also in the future (Millar, O'Neill, and Dillman, 2009). Early research has reported that especially respondents with in-person interviews are largely satisfied with that same mode. When asked how they would prefer to answer questions, nearly four out of five people preferred keeping the same mode (i.e. face-to-face) over switching to other modes, like telephone or mail (Groves, 1979).

3. During the interview: survey measurement

Regarding the interview situation, a survey mode can be considered as a multidimensional continuum of several characteristics (Couper, 2011). These different dimensions of survey modes are defined as 1) the degree of interviewer involvement, 2) the degree of contact with the respondent, 3) the channel of communication (aural/visual), 4) the locus of control over the interview, 5) the degree of privacy, and 6) the degree of technology use (e.g. Burton & Jäckle, 2020; Couper, 2011). Below, we discuss the main differences between the face-to-face and web modes based on these six survey dimensions (Table 1).

Table 1: Six survey dimensions

| Survey dimension | Face-to-face | Web |
|---|---|---|
| 1. Interviewer involvement | Fully interviewer-administered | Fully self-administered |
| 2. Contact with the respondent | Direct, verbal, and nonverbal cues | Indirect |
| 3. Communication channel (aural/visual) | Aural, supplemental visual material | Visual |
| 4. Locus of control over the interview | Interviewer | Respondent, software |
| 5. Privacy | Medium | High |
| 6. Technology use | No computer skills required from the respondent | Computer skills required from the respondent (or proxy) |

Firstly, a survey may be administered by an interviewer or by the respondent himself. In face-to-face interviews, the 1) interviewer involvement is naturally high as the survey is administered by the interviewer. This can have both positive and negative effects on the responses. Ideally, the interviewer takes a neutral position during the interview and helps the respondent to understand and answer the questions correctly without affecting the respondent's responses (e.g. Fowler & Mangione, 1990). For example, for more complex open questions, it has been found that the assistance and probing of an interviewer leads to more detailed responses (De Leeuw, 2012). Less ideally, the interviewer may influence the research outcomes. The mere presence of the interviewer can influence the respondent, especially at sensitive questions, but also the variance in interviewer skills may cause an effect (De Leeuw, 2012). Generally, the interviewer's presence has been found to cause social desirability - respondents are more likely to give answers they think are expected - and acquiescence - respondents are more likely to agree with interviewers than to disagree (Dillman, 2009). Also, extremeness, i.e., the tendency to choose scale endpoints, which can be hard to disentangle from acquiescence, has been studied in the context of mode effects but with mixed results (Groves et al., 2009). There are some indications that web surveys, like mail surveys, result in less extremeness than for example telephone surveys (De Leeuw et al., 2012).

In self-administered interviews such as web surveys, the response process neither benefits nor suffers from interviewer involvement, as the dialog takes place directly between the survey and the respondent.

Secondly, 2) during personal interviews, the interviewer has direct contact with the respondent and the communication relies not only on verbal but also on nonverbal cues that may help to understand each other Schwarz et al., 1991. During face-to-face interviews, it is also possible to use additional written materials such as showcards to facilitate responding, collect interviewer observations, or perform physical measurements (Couper, 2011). In web surveys, researchers have only indirect contact with the respondent. The questionnaire itself has to include all the cues and support to make the respondent understand the questions and provide the answers in the correct format.

Third, 3) the response process in face-to-face interviews relies mainly on aural communication, as the respondent hears the questions spoken out by the interviewer, except for questions with visual aids such as showcards. Web surveys, however, rely on visual processing of the questions. One reported difference between these information processing channels is that aural (auditory) modes are associated with recency effects, while visual modes on the contrary are associated with primacy effects (Schwarz, Knäuper, Oyserman, & Stich, 2012), although the evidence is mixed (De Leeuw et al., 2012). According to the theory, in aural modes, when a respondent hears a question, he or she cannot start processing an answer until the interviewer has read the whole question. When the respondent hears a long list of response options, he or she is likely to recall the last items of the list better than the first ones, because those were the last thing heard. An opposite effect is thought to occur when the respondent reads the response options (visual mode): the respondent thinks about the options in the order that they are presented and is therefore more likely to choose options in the beginning of a list than at the end of a list

(Schwarz et al., 2012). This latter effect has been found, e.g., in check-all-that apply questions in web surveys (Dillman & Smyth, 2007) and has also been explained as response strategy called satisficing: respondents may choose the first option that they encounter and somewhat applies, just to finish the task instead of putting the full effort to provide optimal responses (Krosnick, 1991; Dillman & Smyth, 2007).

The 4) locus of control over the interview rests mostly with the interviewer in face-to-face surveys: the interviewer defines the pace and flow of the interview (Couper, 2011). In web surveys, on the other hand, the respondent decides when and where to fill out the survey, and at which pace. This could be beneficial especially for working people. Furthermore, the interviewer is not there to control possible misunderstandings, although researchers can try to compensate this loss of control by adding checks and limits to response options in web surveys (Couper, 2011). Moreover, self-administered modes such as web surveys provide the respondent with unrestricted opportunity to commit to satisficing and response behaviors such as straightlining (providing identical answers to a battery of question items) or speeding (answering the survey very quickly), styles that have also been found to be correlated (Zhang & Conrad, 2014).

As for 5) privacy, interviewer-administered surveys can be considered to offer a medium level of seclusion (Couper, 2011). The interviewer is obviously always present in face-to-face surveys and when the interview takes place at home, other family members may be listening to the interview as it is not always possible to arrange for a private space. In theory, web surveys offer the possibility for full privacy, although in practice the respondent may still choose otherwise. A setting with a high degree of privacy is especially important when answering sensitive questions or questions about persons that are close to the respondent.

Finally, face-to-face and web surveys differ in the use of 6) computer technology. During face-to-face interviews, it is the interviewer that administers the survey on the computer, often either laptop or tablet. The interviewers are trained to do this and receive technical support from the fieldwork agency when needed. Further, all interviewers of a fieldwork agency often use the same type of or similar devices. Self-completion web surveys, however, require the respondent to have access to an input device - whether a desktop, laptop, tablet, or a smartphone - as well as an internet connection, and to be competent to use the device. In cross-national studies like SHARE, the internet penetration may vary largely between countries. Moreover, when conducting a web survey, there can be a great variety in the input devices, operating systems, and browsers.

All the abovementioned differences in interview characteristics should be accounted for when switching modes. In the following, we look at how these differences played a role in the case study of SHARE Netherlands where the usual face-to-face interview was conducted as a web survey in Waves 6 and 7.

4. Case: SHARE Netherlands

SHARE (Survey on Health, Ageing and Retirement in Europe)² is a large longitudinal study on the life of people aged 50 years and older in Europe. The survey covers key areas of the respondents' life such as health, socioeconomic status, and social and family networks (Börsch-Supan et al., 2013).

Our study is based on the data from the Netherlands, one of the countries that have participated in SHARE since the first wave in 2004. The Netherlands participated in the biennial face-to-face interviews until Wave 5. In Wave 6, the Netherlands was unable to participate in the harmonized SHARE survey, using the CAPI mode, due to insufficient funding (Das, Bruijne, de, Janssen, & Kalwij, 2017). Instead, the Dutch survey was fielded as a mixed-mode survey: the survey was offered as a self-administered web survey, and for those who refused because they were unable to complete the survey online, a follow-up telephone survey was offered.

In the following, we investigate the consequences of the mode switch on the Dutch net sample and data quality. The regular SHARE Wave 5 face-to-face interviews were conducted in 2013 (Malter and Börsch-Supan, 2015) and the Wave 6 web survey in the fall of 2015, with the telephone follow-up in the beginning of 2016 (Das et al., 2017). For our analyses on mode-selection effects (sections 4.1 and 4.2), we look at the longitudinal sample that participated in the face-to-face interviews in Wave 5 and the degree to which those respondents participated in the mixed-mode survey in Wave 6³. For our investigation on mode measurement effects (see section 4.3), we conduct within-respondent analyses among those respondents who participated in both Wave 5 and Wave 6.

Since the different modes were used sequentially in separate measurements, it is strictly speaking not possible to disentangle real changes over time in the respondent and changes caused by the mode switch (De Leeuw, Hox, & Dillman, 2012). However, after another web survey in Wave 7, the Dutch data collection was conducted in face-to-face again in Wave 8. Although Wave 8 fieldwork had to be halted prematurely when the Corona pandemic started, the longitudinal character of the study provides an opportunity to test whether the effects we may find on data quality were caused by the mode or by calendar time. If the same change between waves is found when switching back from web to face-to-face, this demonstrates a real

² For more information on SHARE, see Börsch-Supan, Brandt, Hunkler, Kneip, Korbmacher, Malter, Schaan, Stuck, & Zuber (2013) and Börsch-Supan, A., T. Kneip, H. Litwin, M. Myck, G. Weber (2015). For information on SHARE response and retention rates, see Bergmann, Kneip, Luca, & Scherpenzeel (2019).

³ In the Wave 6 mixed-mode measurement, the longitudinal SHARE sample was supplemented with a refreshment sample from the LISS panel, a probability-based online panel (Das et al., 2017), but since this group was obviously more likely to participate via web and already used to online surveys, it was excluded from the analysis.

change over time. If the opposite effect is found when switching back, this is a strong indication that the change is caused by the mode.

4.1. Coverage

The survey dimension of computer technology and in specific, the requirement of adequate computer skills and internet access to fill out the web survey, represented one of the main challenges for the mode switch. It formed a possible cause for coverage and nonresponse error. This was especially so because of the research project's target age group of 50 years and older, even if the internet penetration is relatively high in the Netherlands. The fieldwork for the web survey in Wave 6 was conducted in 2015. That year, 94% of the Dutch as of 12 years old had access to internet facilities (Statistics Netherlands, 2019). Until the age group of 75 years, the internet penetration rate remained over 90%, but as of the age group of 75 years and older, it declined rapidly to 59% (Statistics Netherlands, 2019).

The SHARE survey includes questions about internet usage. When during the face-to-face interviews in Wave 5, the respondents were asked whether they had used the internet in the past 7 days, 75% reported to have done so. But also in the SHARE sample there was a drop in internet usage to 39% for the age group of 75 years and older. As for self-evaluated computer skills, about 45% of the respondents considered themselves to have good to excellent computer skills, but only 19% thought so in the age group of 75 years and older.

4.2. Response

Among the total longitudinal SHARE sample, i.e. respondents who had participated in at least one of the earlier waves, 4,791 respondents were invited to participate in the web survey in Wave 6 (Das et al., 2017). Those who refused were approached by telephone (when telephone number was known) and offered to participate by telephone (CATI, Computer-Assisted Telephone Interviewing). In total, 2,071 (43%) longitudinal respondents participated in the main interview⁴. Only a small fraction of them, 130 individuals, participated by telephone.

When looking at retention of the original (Wave 1) panel sample, 47% of those who participated in Wave 5 also participated in Wave 6, which was clearly lower than the earlier wave-to-wave participation rates of the original sample that varied between 62% at the first transition after the baseline interview and 85% at the fourth transition between Waves 4 and 5 (Bergmann et al., 2019). When also including the original sample members who did not participate in Wave 5 but could be recovered in Wave 6, the retention rate rises to 57%, but remains lower than in earlier waves (62% - 94%) (Bergmann et al., 2019).

⁴ In addition, during the fieldwork, 38 respondents were reported to have passed away and for 19 of them, an End-of-Life interview about their last year could be finished by a proxy.

When looking at the wave-to-wave retention rate among the respondents who had completed the Wave 5 main interview (4,165), we see that 43% participated also in the Wave 6 Main interview: 40% via web and 3% by telephone (Table 2). The retention rate via web was higher among those who considered themselves more skilled using the computer (χ^2 (5, $N=4,148$)=269.07, $p<.001$; Cramer's $V=.25$, $p<.001$) and those who had used the internet in the past 7 days (χ^2 (1, $N=4151$)=301.56, $p<.001$; Cramer's $V=.27$, $p<.001$) in Wave 5. As the oldest respondents were found to be less computer-savvy and less often using the internet, it is not surprising that we found the retention rate to drop in the oldest age group of 75 years and older, where it was only 26% via web. In comparison, the share of taking part via telephone interviews rises by age.

Table 2: Retention (%) of Wave 5 face-to-face participants in Wave 6 in mixed-mode

| | Age (as in 2015) | | | Total |
|----------------------|------------------|---------------|------------|-------|
| | <65 years | 65 – 74 years | >=75 years | |
| Completed interviews | 44 | 49 | 33 | 43 |
| Via web | 43 | 47 | 26 | 40 |
| Via telephone | 1 | 2 | 7 | 3 |
| Nonresponse | 56 | 51 | 67 | 57 |
| <i>N</i> | 1,589 | 1,494 | 1,082 | 4,165 |

χ^2 (4, $N = 4,165$) = 201.26, $p < .001$.

Other background characteristics could also play a role in the willingness to participate in the mixed-mode measurement in Wave 6. In addition to age and internet usage, we analyzed the likeliness to participate by gender, education level, length of participating in the SHARE study, and earlier respondent willingness level during the interview in Wave 5. Further, because the first Wave 6 advance letter was sent to the household respondent when multiple respondents had participated in the previous wave (Das et al., 2017), we investigated whether this had an effect. We also looked at the type of housing (house, apartment, or elderly or nursing home), as this had earlier been found to affect the unit response rate in the Dutch SHARE sample (Kalwij, 2010). We found no effect by gender (χ^2 (1, $N=4165$)=1080, $p=.299$; Cramer's $V=.02$, $p=.299$), but saw that the participation rate grew by education level (χ^2 (2, $N=4080$)=63.63, $p<.001$; Cramer's $V=.13$, $p<.001$), was higher among the household respondents (χ^2 (1, $N=4165$)=43.95, $p<.001$; Cramer's $V=.10$, $p<.001$), people living in houses (χ^2 (2, $N=4131$)=48.38, $p<.001$; Cramer's $V=.11$, $p<.001$) and when people had already participated in several waves (χ^2 (4, $N=4165$)=35.41, $p<.001$; Cramer's $V=.09$, $p<.001$) or when they were (very) willing to participate in the interview in Wave 5 (χ^2 (5, $N=4139$)=31.27, $p<.001$; Cramer's $V=.09$, $p<.001$). These outcomes apply both for the total mixed mode survey and response via web only.

To ascertain the role of these different background variables on the web participation in a multivariate setting, we conducted a logistic regression analysis (Table 3). As independent variables, we included the age grouped into younger or older than 75 years, gender, education level (low, middle, high), usage of the internet in the past 7 days (yes/no), being the household

respondent, housing type (house, apartment, or elderly or nursing home), participating longer than since the previous wave, and high willingness to answer questions during the interview in Wave 5 (yes/no). The results confirm that internet usage was strongly associated with the likeliness to participate via web. Also, people under the age of 75, respondents with higher education, respondents who had participated longer in SHARE, and the household respondents were more likely to participate. Further, respondents who lived in a house, or apartment, were more likely to participate than those living in an elderly or nursing home. The model was statistically significant ($\chi^2(10, N = 4,035) = 431.09, p < .001$), explained 13.7% (Nagelkerke R²) of the variance, and correctly classified 63.7% of cases.

Table 3. Logistic regression predicting likelihood of responding via web

| | <i>B</i> | S.E. | Wald | <i>df</i> | <i>p</i> | Odds ratio | 95.0% C.I. for Odds ratio | |
|--|----------|------|--------|-----------|----------|------------|---------------------------|-------|
| | | | | | | | Lower | Upper |
| High age: 75 years or older (no/yes) | -.37 | .09 | 16.64 | 1 | <.001 | .69 | .58 | .82 |
| Used internet in past 7 days (no/yes) | 1.22 | .10 | 138.68 | 1 | <.001 | 3.39 | 2.77 | 4.15 |
| Gender (male) | .06 | .07 | .73 | 1 | .39 | 1.06 | .93 | 1.21 |
| Education: low | | | 9.49 | 2 | .01 | | | |
| Education: middle | .19 | .09 | 5.02 | 1 | .03 | 1.21 | 1.02 | 1.43 |
| Education: high | .25 | .09 | 8.36 | 1 | <.01 | 1.28 | 1.08 | 1.51 |
| Housing: elderly or nursing home | | | 8.83 | 2 | .01 | | | |
| Housing: apartment | .53 | .23 | 5.35 | 1 | .02 | 1.70 | 1.09 | 2.67 |
| Housing: house | .63 | .22 | 8.31 | 1 | <.01 | 1.88 | 1.22 | 2.89 |
| Started in SHARE before previous wave (no/yes) | .39 | .07 | 30.80 | 1 | <.001 | 1.47 | 1.28 | 1.69 |
| Household respondent (no/yes) | .47 | .08 | 38.89 | 1 | <.001 | 1.60 | 1.38 | 1.86 |
| High willingness to answer questions in previous wave (no/yes) | .14 | .08 | 3.01 | 1 | .08 | 1.15 | .98 | 1.36 |
| <i>Constant</i> | -2.70 | .25 | 120.72 | 1 | <.001 | .07 | | |

N=4,035

4.3. Mode effects during the interview

To examine possible mode effects during the interview, we investigated the data quality based on the framework of the different survey dimensions, expect technology use which was discussed in the chapter about coverage. Our aim was to find out whether the data showed any indications of measurement error caused by mode effects. Measurement error occurs at the level of the question: while for some questions in a survey there is no mode effect, other questions in the same survey may show a strong mode effect (Burton & Jäckle, 2020). For each mode effect, we selected questions that were expected to be sensitive for the associated effect. For the analyses, we used within-person comparisons of responses to questions that were included both in the face-to-face interviews in Wave 5 and the web survey in Wave 6. The telephone responses were excluded from the analysis because these can influence the mode effects of switching from face-to-face to web interviews. In those cases that we found a

statistically significant effect, we repeated the analysis among the respondents that completed both Wave 7 in web and Wave 8 in face-to-face.

Interviewer involvement - social desirability

To study mode effects deriving from interviewer involvement during the interview, we compared the rate of reporting on sensitive topics. In general, interviewer-assisted modes can lead to more socially desirable answers (e.g. De Leeuw, 2012). Underreporting sensitive topics demonstrates socially desirable behavior and therefore we expected the web mode to yield higher reporting on such topics. We chose ten questions from the Mental Health module that were asked in both waves 5 and 6 and could be asked in both modes without changes in the wording of the questions. The items that were included in the analysis and their outcomes are given in Table 4a.

Table 4a. Comparison of reporting sensitive items in Wave 5 (face-to-face) and Wave 6 (web)

| Item | Label | Face-to-face (% positive responses) | Web (% positive responses) | p-value |
|---|---------------------------------|-------------------------------------|----------------------------|---------|
| <i>McNemar's test (Response options: Yes/No)</i> | | | | |
| MH002 | Sad or depressed last month | 27 | 31 | .005 |
| MH004 * | Wished to be dead last month | 3 | 4 | .086 |
| MH007 | Trouble sleeping recently | 24 | 30 | <.001 |
| MH010 | Irritability recently | 14 | 14 | .765 |
| MH013 | Fatigue last month | 25 | 26 | .311 |
| MH014 * | Difficulties with concentration | 9 | 15 | <.001 |
| MH017 | Cried last month | 26 | 32 | <.001 |
| <i>Wilcoxon's test (Response options: 1. Often; 2. Some of the time; 3. Hardly ever or never. Options 1 and 2 are counted as positive responses.)</i> | | | | |
| MH034 | Feels lack of companionship | 21 | 27 | <.001 |
| MH035 | Feels left out | 13 | 20 | <.001 |
| MH036 | Feels isolated from others | 13 | 22 | <.001 |

N=1,560

*) In face-to-face interviews, the interviewer interpreted and coded the answer to Yes or No.

The outcomes suggest that the respondents are more likely to report sensitive issues when completing a self-administered web survey, in line with earlier literature (e.g. De Leeuw, 2012). For seven of the ten items, the web mode results in significantly higher reporting of the mental health issues than the face-to-face survey. For the remaining three items, mental health problems are reported equally often in both modes. When testing this outcome with the opposite mode switch between Wave 7 and 8, we found that apart from one item (feeling depressed), the rate of reporting mental health issues remains the same or even declines again when the mode changes back to face-to-face (Table 4b).

Table 4b. Comparison of reporting sensitive items in Wave 7 (web) and Wave 8 (face-to-face)

| Item | Label | Web (% positive responses) | Face-to-face (% positive responses) | p-value |
|---|---------------------------------|----------------------------|-------------------------------------|---------|
| <i>McNemar's test (Response options: Yes/No)</i> | | | | |
| MH002 | Sad or depressed last month | 31 | 38 | <.001 |
| MH004 * | Wished to be dead last month | 4 | 4 | 1.000 |
| MH007 | Trouble sleeping recently | 30 | 27 | .045 |
| MH010 | Irritability recently | 13 | 13 | .819 |
| MH013 | Fatigue last month | 25 | 24 | .539 |
| MH014 * | Difficulties with concentration | 15 | 10 | <.001 |
| MH017 | Cried last month | 29 | 26 | .128 |
| <i>Wilcoxon's test (Response options: 1. Often; 2. Some of the time; 3. Hardly ever or never. Options 1 and 2 are counted as positive responses.)</i> | | | | |
| MH034 | Feels lack of companionship | 32 | 30 | .424 |
| MH035 | Feels left out | 17 | 14 | .006 |
| MH036 | Feels isolated from others | 19 | 13 | <.001 |

N=1,098

Contact with the respondent - understanding survey questions

One of the benefits of direct contact with the respondent during the interview is that the interviewer is able to help the respondent to understand the questions, whereas this possibility is lacking when filling out a web survey. At the end of the SHARE survey, it is asked whether the respondent understood the questions on a scale from 1 = never to 6 = always. In the face-to-face survey, this was evaluated by the interviewer and in the web survey, directly by the respondent (Table 5a and 5b). Only a good half (53%) of the respondents were reported to have understood all the questions of the Wave 5 face-to-face interview, in the web survey even less than half (47%) said to have understood all the questions in the web survey. The differences between the two modes were statistically significant at the 0.05 level (Wilcoxon signed-rank test, $Z=-2.20$, $p=.03$). When switching back from web to face-to-face, the mode effect is reverted (Wilcoxon signed-rank test, $Z=-3.05$, $p=.002$), again more respondents reported to always have understood the questions in the face-to-face interview.

Table 5a. Understanding survey questions in Wave 5 face-to-face and Wave 6 web

| IV008 Respondent understood questions | Wave 5 face-to-face (%) | Wave 6 web (%) |
|--|--------------------------------|-----------------------|
| 1. Never | 1 | 0 |
| 2. Almost never | 0 | 0 |
| 3. Now and then | 1 | 2 |
| 4. Often | 15 | 16 |
| 5. Very often | 29 | 35 |
| 6. Always | 53 | 47 |
| Total | 100 | 100 |

N=1,557

Table 5b. Understanding survey questions in Wave 7 web and Wave 8 face-to-face

| IV008 Respondent understood questions | Wave 7 web (%) | Wave 8 face-to-face (%) |
|--|-----------------------|--------------------------------|
| 1. Never | 0 | 2 |
| 2. Almost never | 0 | 0 |
| 3. Now and then | 1 | 1 |
| 4. Often | 10 | 8 |
| 5. Very often | 33 | 20 |
| 6. Always | 56 | 68 |
| Total | 100 | 100 |

N=1,106

Contact with the respondent – item nonresponse

Another benefit of the interviewer presence is that they may prompt for answers when the respondent is hesitating. Moreover, while it was possible to answer 'Don't know' or 'I refuse to say' in the face-to-face interview, these options were not actively offered to the respondent, but were only entered by interviewers when spontaneously mentioned by respondents. In the web survey, on the other hand, the 'Don't know' and 'Refusal' options were provided as visible buttons on screen, but only for questions when deemed necessary, e.g. at knowledge questions. To compare item nonresponse (non-substantive responses), we selected questions from the Housing and Consumption modules about monetary amounts for which it was possible to answer 'Don't know' or 'Refuse to answer' in both modes, and that were presented to an adequate number of respondents in both waves.

In line with the earlier literature (e.g. De Leeuw, 2012, Jäckle et al., 2015), Table 6a shows a higher rate of item nonresponse in the self-administered web survey than face-to-face interviews. All the investigated questions show a significant difference between modes when switching from face-to-face to web: in the web survey more people provide a non-substantive answer compared to the face-to-face interview. This pattern is confirmed by the return to face-to-face interviews with a shrinking share of non-substantive answers (Table 6b).

Table 6a. Comparison of providing non substantive (Don't know or Refusal) answers between Wave 5 face-to-face and Wave 6 web surveys (McNemar's test)

| Item | Label | N | Percentage of non-substantive answers | | p-value |
|--------|---|------|---------------------------------------|------------|---------|
| | | | Wave 5 face-to-face | Wave 6 web | |
| HO008e | Amount charges and services | 206 | 10 | 23 | <.001 |
| HO015e | Amount still to pay on mortgage or load | 587 | 21 | 33 | <.001 |
| HO024e | Value of property | 847 | 5 | 23 | <.001 |
| HO067e | Amount similar dwelling today's market | 860 | 14 | 34 | <.001 |
| CO002e | Amount spend on food at home | 1118 | 9 | 25 | <.001 |
| CO003e | Amount spend on food outside the home | 1118 | 4 | 20 | <.001 |
| CO011e | Value of home produced food | 111 | 22 | 41 | .001 |

Table 6b. Comparison of providing non substantive (Don't know or Refusal) answers between Wave 7 web and Wave 8 face-to-face surveys (McNemar's test)

| Item | Label | N | Percentage of non-substantive answers | | p-value |
|--------|---|-----|---------------------------------------|---------------------|---------|
| | | | Wave 7 web | Wave 8 face-to-face | |
| HO008e | Amount charges and services | 109 | 19 | 12 | .170 |
| HO015e | Amount still to pay on mortgage or load | 260 | 26 | 16 | .002 |
| HO024e | Value of property | 442 | 20 | 3 | <.001 |
| HO067e | Amount similar dwelling today's market | 450 | 33 | 6 | <.001 |
| CO002e | Amount spend on food at home | 596 | 26 | 7 | <.001 |
| CO003e | Amount spend on food outside the home | 596 | 21 | 3 | <.001 |
| CO011e | Value of home produced food | 65 | 33 | 9 | <.001 |

Communication channel - primacy and recency effects

The SHARE survey does not include many scalar questions and when it does, showcards are often used to support the response process in the face-to-face interviews. To investigate the primacy and recency effects, we looked for scalar questions where no showcards were used in the face-to-face interview and the response process thus relied purely on the auditory mode. When comparing the rates of selecting first or last options, we find little evidence on primacy or recency effects between the modes (Table 7), in contrast to what could be expected based on the literature (see e.g. Schwarz et al., 2012). At two of the five investigated items, the first option was in fact selected less often in the visual web mode, which could be an indication of less extremeness in the web mode, but that effect cannot be concluded as most of the items reveal no difference. It should be mentioned that the used scales are relatively short (4 or 5-points), making them relatively easy to process in an auditory mode and as such, less vulnerable to mode effects by design.

Table 7. Investigating primacy and recency in face-to-face and web (McNemar's test)

| Item | Label | N | Selected | F-to-f (%) | Web (%) | p |
|-------|--|------|--------------|------------|---------|-------|
| PH003 | Health in general | 1628 | First option | 15 | 9 | <.001 |
| | | | Last option | 3 | 3 | .70 |
| PH043 | Eye sight distance | 1620 | First option | 26 | 24 | .15 |
| | | | Last option | 1 | 2 | .62 |
| PH044 | Eye sight reading | 1620 | First option | 20 | 16 | .01 |
| | | | Last option | 5 | 3 | <.01 |
| HC125 | Satisfaction with coverage in basic health insurance | 1559 | First option | 11 | 9 | .24 |
| | | | Last option | 2 | 1 | .11 |
| CO007 | Is household able to make ends meet | 1115 | First option | 3 | 3 | 1.00 |
| | | | Last option | 59 | 40 | <.001 |

Locus of control - straightlining and speeding

Response behavior such as straightlining and speeding (Zhang & Conrad, 2014) are used to examine data quality but can also indicate lack of (interviewer) control during the interview. To investigate the presence of straightlining, we looked at the battery of twelve questions in the Activities module (AC014 – AC25), which could only be completed by the respondent without a proxy and was answered on a scale from 1 = Often to 4 = Never. As in CAPI, these items were presented each on a separate screen in the web survey. We investigated whether all the items were answered with the same response option and found no indication of straightlining in either modes. For the within-person comparison group ($N=1475$), only two respondents in the face-to-face and one (other) respondent in the web mode gave identical responses to all the items. Further, we compared the share of respondents who had completed the whole survey in less than 30 minutes. Only a few cases in the web survey (0.3%) matched this criterion based on a within-respondent analysis. Among the face-to-face interviews, the rate was somewhat higher

(3.3 %). The latter, however, also included several unrealistic short completion times that could be due to technical inconsistencies in recording timestamps.

Locus of control - willingness to answer questions

Further, at the end of the SHARE interview, a question is asked about the willingness to answer questions in the survey. In the face-to-face interview, the willingness was evaluated by the interviewer and in the web survey, by the respondent themselves. While the willingness to answer of almost all respondents within the analysis group in the Wave 5 face-to-face interview was evaluated to be very good or good, this was less the case for the Wave 6 self-evaluated web survey (Table 8a): 17% of the respondents reported to be fairly willing to answer the web survey and 10% reported that their willingness got worse during the survey. The differences between the two modes were statistically significant (Wilcoxon signed-rank test, $Z=-28.80$, $p<.001$). As a reason for the declining willingness during the web survey it was often mentioned that the respondent was losing interest or concentration, which could be caused by the length of the survey. This finding of a lower willingness to answer in web interviews, compared to face-to-face interviews, is confirmed when the mode switching back from web to face-to-face in Wave 8 (Table 8b; Wilcoxon signed-rank test, $Z=-24.79$, $p<.001$).

Table 8a. Willingness to answer questions in Wave 5 face-to-face and Wave 6 web

| IV004 Willingness to answer | Wave 5 face-to-face | Wave 6 web |
|---|----------------------------|-------------------|
| 1 Very good | 80 | 17 |
| 2 Good | 19 | 55 |
| 3 Fair | 1 | 17 |
| 4 Bad | 0 | 0 |
| 5 Good in the beginning, got worse during the interview | 0 | 10 |
| 6 Bad in the beginning, got better during the interview | 0 | 1 |

$N=1,557$

Table 8b. Willingness to answer questions in Wave 7 web and Wave 8 face-to-face

| IV004 Willingness to answer | Wave 7 web | Wave 8 face-to-face |
|---|-------------------|----------------------------|
| 1 Very good | 19 | 85 |
| 2 Good | 56 | 14 |
| 3 Fair | 16 | 1 |
| 4 Bad | 1 | 0 |
| 5 Good in the beginning, got worse during the interview | 8 | 0 |
| 6 Bad in the beginning, got better during the interview | 1 | 0 |

$N=1,106$

Privacy during the interview

In theory, the web mode offers a more private means for filling out a survey, whereas during a face-to-face interview the respondent is never truly alone. However, in the SHARE survey it is possible to ask a proxy to help with responding, which was also encouraged for the web survey if the respondent lacked skills to handle the web questionnaire. Of the people who answered both Waves 5 and 6 ($N=1,557$), there were only a few (0.3%) respondents using a proxy in the face-to-face interviews, but 11% using a proxy in the web survey. The proxies were mostly partners or children of the respondents. While there were more respondents using a proxy when looking at the whole net sample of the face-to-face interviews in Wave 5 (1.5%), almost none of these people took part in the Wave 6 web survey. People who already need the help of a proxy to participate in a face-to-face interview seem to be less likely to take part in a web survey.

When asked whether someone else, except for the proxy, was present during the interview, this was more often the case during the Wave 5 face-to-face interview (32%) than during the Wave 6 web survey (19%) (McNemar $X^2(1) = 74.98, p < .001$). During the face-to-face interviews, these third persons also more often intervened with the interview (Wilcoxon signed-rank test, $Z = -2.34, p = .02$).

The differences between the modes were smaller but in line with the earlier outcomes when moving back from the Wave 7 web survey to the Wave 8 face-to-face survey, except that there was no significant difference regarding the presence of third parties.

All in all, the results show a mixed picture regarding privacy: during the web survey there were more often proxies that helped, which is likely due to a lack of technical skills when answering a web survey, but there were less third-persons intervening when answering the questions.

5. Conclusions

Switching from face-to-face interviews to web interviews can lead to a different net sample composition and can affect response behavior. In our longitudinal sample of respondents 50 years of age and older, the response to a web survey with a telephone follow-up was substantially lower than in the earlier face-to-face survey. The response was especially lower in the age group of 75 years and older and higher among those who used the internet in the past seven days.

Further, respondents understood the questions better and were more willing to answer the entire survey in face-to-face interviews than in web interviews. There were also less non-substantive answers (don't know or refusal) provided during face-to-face interviews. In web interviews, however, sensitive issues related to mental health were reported more often than in face-to-face interviews. Finally, there was no indication of primacy and recency effects by the two modes, nor of straightlining for a battery of questions with responses on a four-points Likert scale.

To conclude, our findings from the field support the presence of survey mode effects on unit and item response and data quality.

Used datasets

Börsch-Supan, A. (2022). *Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 5*. Release version: 8.0.0. SHARE-ERIC. Data set. DOI: <http://dx.doi.org/10.6103/SHARE.w5.800>

De Bruijne, Marika; Das, Marcel; Janssen, Josette; Kalwij, Adriaan; Oudejans, Marije et. al. (2017): Dutch Mixed Mode Experiment. Version: 1.0.0. SHARE-ERIC. Dataset. <https://doi.org/10.6103/SHARE.w6NLmmExp.100>

Börsch-Supan, A. (2022). Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 5 & 6. NL Interview Duration. Release version: 8.0.0. SHARE-ERIC. Internal data set.

Contributions

A special thank to Tim Birkenbach for the (internal) Dutch data release. Many thanks also to Michael Bergmann for his review and advice that helped us to improve the manuscript

Acknowledgement

This paper uses data from SHARE Waves 1, 2, 3, 4, 5, 6, 7, and 8 (DOIs: 10.6103/SHARE.w1.800, 10.6103/SHARE.w2.800, 10.6103/SHARE.w3.800, 10.6103/SHARE.w4.800, 10.6103/SHARE.w5.800, 10.6103/SHARE.w6.800, 10.6103/SHARE.w7.800, 10.6103/SHARE.w8.800), see Börsch-Supan et al. (2013) for methodological details. (1)The SHARE data collection has been funded by the European Commission, DG RTD through FP5 (QLK6-CT-2001-00360), FP6 (SHARE-I3: RII-CT-2006-062193, COMPARE: CIT5-CT-2005-028857, SHARELIFE: CIT4-CT-2006-028812), FP7 (SHARE-PREP: GA N°211909, SHARE-LEAP: GA N°227822, SHARE M4: GA N°261982, DASISH: GA N°283646) and Horizon 2020 (SHARE-DEV3: GA N°676536, SHARE-COHESION: GA N°870628, SERISS: GA N°654221, SSHOC: GA N°823782, SHARE-COVID19: GA N°101015924) and by DG Employment, Social Affairs & Inclusion through VS 2015/0195, VS 2016/0135, VS 2018/0285, VS 2019/0332, and VS 2020/0313. Additional funding from the German Ministry of Education and Research, the Max Planck Society for the Advancement of Science, the U.S. National Institute on Aging (U01_AG09740-13S2, P01_AG005842, P01_AG08291, P30_AG12815, R21_AG025169, Y1-AG-4553-01, IAG_BSR06-11, OGHA_04-064, HHSN271201300071C, RAG052527A) and from various national funding sources is gratefully acknowledged (see www.share-project.org).

Furthermore, this paper makes use of internal data (SHARE Wave 5 & 6. NL Interview Duration. Release version: 8.0.0.

References

- Bergmann, M., T. Kneip, G. De Luca, and A. Scherpenzeel (2019). Survey participation in the Survey of Health, Ageing and Retirement in Europe (SHARE), Wave 1-7. Based on Release 7.0.0. SHARE Working Paper Series 41-2019. Munich: MEA, Max Planck Institute for Social Law and Social Policy.
- Burton, J., & Jäckle, A. (2020). Mode effects. *ISER: Understanding Society Working Paper Series*, (2020-05).
- Börsch-Supan, A., T. Kneip, H. Litwin, M. Myck, G. Weber (Eds.) (2015). *Ageing in Europe - Supporting Policies for an Inclusive Society*. Berlin: De Gruyter.
- Börsch-Supan, A., M. Brandt, C. Hunkler, T. Kneip, J. Korbmacher, F. Malter, B. Schaan, S. Stuck, S. Zuber (2013). Data Resource Profile: The Survey of Health, Ageing and Retirement in Europe (SHARE). *International Journal of Epidemiology*. DOI: 10.1093/ije/dyt088.
- Das, M., Bruijne, M. de, Janssen, J., & Kalwij, A. (2017). Experiment: Internet interviewing in the sixth wave of SHARE in the Netherlands. In: Malter, F., & Börsch-Supan, A. (Eds.), *SHARE Wave 6: Panel innovations and collecting Dried Blood Spots* (pp. 151-162). Munich: Mea, Max Planck Institute for Social Law and Social Policy, ISBN: ISBN 978-3-00-058556-2.
- Couper, M. P. (2011). The future of modes of data collection. *Public Opinion Quarterly*, 75(5), 889-908.
- Dillman, D. A., & Smyth, J. D. (2007). Design effects in the transition to web-based surveys. *American journal of preventive medicine*, 32(5), (pp. S90-S96), doi:10.1016/j.amepre.2007.03.008
- Diment, K., & Garrett-Jones, S. (2007). How demographic characteristics affect mode preference in a postal/web mixed-mode survey of Australian researchers. *Social Science Computer Review*, 25(3), 410-417.
- De Leeuw, E. D. (2012). Choosing the method of data collection. In *International handbook of survey methodology* (pp. 113-135). Routledge.
- De Leeuw, E. D., Hox, J. J., & Dillman, D. A. (2012). Mixed-mode Surveys: When and Why. In *International handbook of survey methodology* (pp. 299-316). Routledge.
- De Leeuw, E. D. (1992). *Data quality in mail, telephone and face to face surveys*. TT Publikaties, Amsterdam.

- Fowler, F. J., & Mangione, T. W. (1990). Standardized survey interviewing; minimizing interviewer-related error. Sage Publications. March 12, 2024.
- Groves, R. M. (1979). Actors and Questions in Telephone and Personal Interview Surveys. *The Public Opinion Quarterly*, 43, 2, 190-205.
- Groves, R. M., Fowler Jr, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2009). Survey methodology. John Wiley & Sons.
- Jäckle, A., Lynn, P., & Burton, J. (2015). Going online with a face-to-face household panel: effects of a mixed mode design on item and unit non-response. In *Survey Research Methods* (Vol. 9, No. 1, pp. 57-70).
- Kalwij, A. (2010). An empirical analysis of the association between neighborhood income and unit non-response in the Survey Of Health, Ageing, and Retirement in Europe. *Review of Income and Wealth*, 56(2), 351-365.
- Krosnick, J. A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology*, 5(3), 213-236.
<https://doi.org/10.1002/acp.2350050305>
- Malter, F. and A. Börsch-Supan (Eds.) (2015). SHARE Wave 5: Innovations & Methodology. Munich: MEA, Max Planck Institute for Social Law and Social Policy.
- Millar, M. M., O'Neill, A. C., & Dillman, D. A. (2009). Are mode preferences real. Pullman: Washington State University, 9, 2-41.
- Mulder, J., & Bruijne, de, M. (2019). Willingness of Online Respondents to Participate in Alternative Modes of Data Collection. *Survey Practice*, 12, 1, 1-11.
- Revilla, M. (2015). Comparison of the quality estimates in a mixed-mode and a unimode design: an experiment from the European Social Survey. *Quality & Quantity: International Journal of Methodology*, 49(3), 1219-1238. <https://doi.org/10.1007/s11135-014-0044-5>
- Scherpenzeel, A. C., Das, M. (2010). "'True" Longitudinal and Probability-Based Internet Panels: Evidence From the Netherlands. In *Social and behavioral research and the Internet* (pp. 77-104). Routledge.
- Schwarz, N., Knäuper, B., Oyserman, D., & Stich, C. (2012). The psychology of asking questions. In *International handbook of survey methodology* (pp. 18-34). Routledge.

Schwarz, N., Strack, F., Hippler, H.-J., & Bishop, G. (1991). The impact of administration mode on response effects in survey measurement. *Applied Cognitive Psychology*, 5(3), 193–212.

<https://doi.org/10.1002/acp.2350050304>

Statistics Netherlands (8 October 2019). Internet; toegang, gebruik en faciliteiten; 2012-2019.

Statline. <https://opendata.cbs.nl/statline/#/CBS/nl/dataset/83429NED/table?ts=1661864025129>

Zhang Chan and Frederick G. Conrad. (2014). Speeding in web surveys: The tendency to answer very fast and its association with straightlining. *Survey Research Methods*, 8: 127-35.