Testing the Effects of Automated Navigation in a General Population Web Survey

Bakker, Jeldrik; Haan, Marieke; Schouten, Barry; Struminskaya, Bella; Lugtig, Peter; Toepoel, Vera; Giesen, Deirdre; Meertens, Vivian

Veröffentlichungsversion / Published Version
Zeitschriftenartikel / journal article

Empfohlene Zitierung / Suggested Citation:
Bakker, J., Haan, M., Schouten, B., Struminskaya, B., Lugtig, P., Toepoel, V., ... Meertens, V. (2022). Testing the Effects of Automated Navigation in a General Population Web Survey. *Methods, data, analyses : a journal for quantitative methods and survey methodology (mda)*, *16*(1), 107-128. https://doi.org/10.12758/mda.2021.10

methods, data, analyses | Vol. 16(1), 2022, pp. 107-128 DOI: 10.12758/mda.2021.10

Testing the Effects of Automated Navigation in a General Population Web Survey

Jeldrik Bakker[1,2], *Marieke Haan*[3], *Barry Schouten*[1,2], *Bella Struminskaya*[2], *Peter Lugtig*[2], *Vera Toepoel*[2], *Deirdre Giesen*[1] & *Vivian Meertens*[1]

[1] *Statistics Netherlands*

[2] *Utrecht University*

[3] *University of Groningen*

Abstract

This study investigates how an auto-forward design, where respondents navigate through a web survey automatically, affects response times and navigation behavior in a long mixed-device web survey. We embedded an experiment in a health survey administered to the general population in The Netherlands to test the auto-forward design against a manual-forward design. Analyses are based on detailed paradata that keep track of the respondents' behavior in navigating the survey. We find that an auto-forward design decreases completion times and that questions on pages with automated navigation are answered significantly faster compared to questions on pages with manual navigation. However, we also find that respondents use the navigation buttons more in the auto-forward condition compared to the manual-forward condition, largely canceling out the reduction in survey duration. Furthermore, we also find that the answer options 'I don't know' and 'I rather not say' are used just as often in the auto-forward condition as in the manual-forward condition, indicating no differences in satisficing behavior. We conclude that auto-forwarding can be used to reduce completing times, but we also advice to carefully consider mixing manual and auto-forwarding within a survey.

Keywords: mixed-device surveys, web surveys, auto-forward, paradata, usability

Web surveys are completed on a range of different devices: PCs, laptops, tablets, and smartphones. Since mobile devices vary in screen size and type of navigation, surveys designed for PCs and laptops tend to be more difficult to navigate on mobile devices. Survey designers have recognized this challenge and have adapted to the smaller screens and different mode of data entry used on smartphones. Nonetheless, even when surveys are "mobile-friendly", web surveys still take longer on smartphones compared to tablets and PCs (Couper, Antoun, & Mavletova, 2017; Couper & Peterson, 2017). Survey duration is an important factor to take into account, because it is a proxy for respondent burden (Zhang & Conrad, 2014). It is conjectured that the maximal duration of a survey that a respondent is willing to complete depends on the type of the device: respondents are less willing to complete longer surveys on smartphones (Hintze, Findling, Scholz, & Mayrhofer, 2014). Therefore, not accounting for survey duration when designing surveys for mixed-mode surveys can result in coverage errors, higher nonresponse, and lower data quality (Cook, 2014; Wells, Bailey, & Link, 2014; Struminskaya, Weynandt & Bosnjak, 2015).

Prior research shows that survey duration can be shortened by using an auto-forward design (Giroux, Tharp, & Wietelman, 2019; Selkälä & Couper, 2018; de Bruijne, 2016; Lugtig, Toepoel, Haan, Zandvliet, & Klein Kranenburg, 2019). In an auto-forward design, respondents automatically advance to the next question after an answer is given. This design feature can improve the survey experience in two ways. First, the required cognitive effort by respondents is reduced by adding smart navigation (i.e., to not have to decide whether the question was the last on the page and to not have to search for the 'next' button). Second, as auto-forward can increase the speed of the survey's advancement, the time spent on the survey is reduced. Respondents find surveys with auto-forward more enjoyable, more interesting, less difficult, and less lengthy compared to designs where manual-forwarding is the standard (Roberts, de Leeuw, Hox, Klausch, & de Jongh, 2012). Furthermore, auto-forwarding seems to decrease satisficing behavior (Selkälä, Callegaro, & Couper, 2020).

There are also potential disadvantages to using auto-forwarding (for an overview, see Giroux et al. 2019). Respondents may get confused because they are used to a page-by-page design in which they use navigation buttons which are often provided in web surveys (Bergstrom, Lakhe, & Erdman, 2016). This confusion may

Acknowledgments

Jeldrik Bakker and Marieke Haan contributed equally to this work and share the co-first authorship

Direct correspondence to

Jeldrik Bakker, Statistics Netherlands & Utrecht University
E-mail: j.bakker@cbs.nl

lead to the accidental skipping of questions resulting in higher item nonresponse (de Bruijne, 2016). Furthermore, the automated pace of the survey may discourage respondents to change answers by using navigation buttons which can lead to more suboptimal responses. Finally, many surveys include questions that are not fit for auto-forwarding, such as open answer questions or "select all that apply" questions. If some questions are auto-forwarded and others not, this may also confuse respondents. In this paper, we use paradata, more specifically we analyze the clicking and answering behavior and response timings between the manual- and auto-forward versions to better understand how auto-forwarding affects both response times and data quality. For this, we use an experimental design that was embedded in a health survey conducted among the general population in The Netherlands.

The remainder of this paper is organized as follows: In section 2, we introduce our research questions and hypotheses. In section 3, we describe the data and methods. We discuss results in section 4. We end with conclusions and discussion in the last two sections.

Study Design and Research Questions

We build on earlier studies that used auto-forwarding design (for an overview see: Giroux et al., 2019). Most of these studies show that response times are generally shortened because of auto-forwarding (Hays et al., 2010; Roberts et al., 2012; Selkäla & Couper, 2018 – for PCs only; Lugtig et al., 2019), but some researchers also find no effects on completion times between auto-forward and manual-forward surveys (Arn et al., 2015; de Bruijne, 2015; Selkäla & Couper, 2018 – for smartphones only), or even longer completion times for auto-forward surveys (Roberts et al., 2013). In this paper, we focus on response times, respondent navigation behavior (i.e., mouse clicks or taps with a finger) and how often respondents answer 'I don't know' and 'I rather not say'. We answer four research questions: 1) Does auto-forwarding reduce response times?, 2) Does auto-forwarding lead to more efficient navigation through the survey?, 3) If so, is more efficient navigation independent of screen size?, and 4) Does auto-forwarding affect how often the answer options 'I don't know' and 'I rather not say' are used?

Our first research question comes from the hypothesis (H1) that auto-forwarding reduces the amount of time needed per survey question. We answer this question in the context of official general population surveys that often are, or were, interviewer-assisted and traditionally have a survey duration of 30 minutes and longer. Our second research question is, however, the most important: it concerns the actual effort needed by respondents to navigate through the survey. To investigate efficient navigation, we compare the number of clicks between an auto-forward version and a manual-forward version of a survey. A respondent is not efficiently

navigating through the survey when navigation buttons are used unnecessarily. We expect that auto-forwarding results in more efficient navigation (H2). The third research question is a follow-up question, which differentiates among smartphones, tablets and PCs. We expect to find more efficient completion on smaller screens (H3). The fourth research question is a first exploration into the impact of an auto-forward interface on item-nonresponse. Because almost all questions that are used in our survey are mandatory, the alternatives for item-nonresponse are: 'I rather not say', 'I don't know,' or selecting a random answer option. In line with the research of Selkälä et al. (2020) we expect that auto-forwarding decreases satisficing behavior, which we define in less 'I rather not say' and 'I don't know' responses (H4).

In order to investigate the four questions, we collected and analyzed audit trail paradata at the survey page-level (see Kreuter, 2013). The paradata we collected provide information about each page of the web survey and about each action requiring server contact (e.g., navigating to the next or the previous page, or start/quit the survey), including page-level response times. Our study will help to determine whether auto-forwarding should be used more widely in web surveys.

Method

Data Collection

Our experiment was linked to the Health Survey (HS) of Statistics Netherlands (SN), which is a repeated cross-sectional survey employing monthly simple random samples from the Dutch population register. The HS is a relatively long survey, with a median completion time of 29.2 minutes. It consists of 409 questions divided over 220 web pages, covering 48 topics, ranging from general health, visits to general practitioners and dentists, hospitalization, medicine use, to health-related behaviors such as smoking, food intake, and physical activity. Respondents have to go through all modules, but the number of questions per module varies based on their medical history and lifestyle. The survey had a predefined order and questions about the same topic were grouped together. The location of the auto-forward questions and manual-forward questions was almost randomly distributed over the survey, except for a block of questions about activities. This block primarily asked questions about either frequencies or duration of activities, and consisted almost solely of questions where auto-forwarding was not possible. The HS uses a sequential mixed-mode design with web followed by face-to-face interviewing. In this paper, we only use the web-administered part of the survey.

The HS auto-forwarding experiment employed a separate sample that ran parallel to the regular HS. The sampling frame was composed of earlier respondents to SN surveys of individuals aged 16 years and older that responded to at least

one of the surveys on a mobile device in the period of September 2016 to June 2017. The stratified simple random sample design with six strata was used: three age groups (16–29, 30–49, and 50 years and older) crossed with a type of device (smartphone, tablet). From each stratum the same number of sampling units was selected, leading to unequal sample inclusion probabilities. Thus, older respondents and respondents who previously used a tablet for survey completion have larger inclusion probabilities. We chose this sampling design in order to be reach higher statistical efficiency in testing the impact of device and age on response times and survey navigation. Sampled respondents were randomly allocated to one of the interface conditions: manual-forward and auto-forward (see section 3.2). Fieldwork took place in August–September 2017. Paradata on response times and navigation were collected using version 5.0.5 of the BLAISE computer-assisted interviewing system (Blaise, 2018).

Overall, 2098 individuals were sent an invitation letter by post and a maximum of two reminders in case they did not participate after one and two weeks. All sample members received a 5€ unconditional cash incentive. In total, 1535 sample units started the survey and 1461 sample units completed the survey with a response rate of 69.6% (AAPOR 2016, RR1). The high response rate can be partly explained by the sample composition of former respondents that completed at least one survey of SN on a smartphone or tablet. In total 74 respondents (4.8%) broke off the survey, 45.9% under the auto-forward condition and 54.1% under the manual-forward condition.

Table 1 shows the choice of device of respondents by age group and highest-attained educational level. The break-off rates per device varied very little and are not shown.

Table 1

Device use by age

	Smartphone		Tablet		PC		Unknown		Total		RR1
Age	n	(%)	n	(%)	n	(%)	n	(%)	n	(%)	(%)
16-29	211	(48.8)	104	(24.1)	117	(27.1)	0	(0)	432	(100)	61.9
30-49	179	(37.1)	204	(42.2)	99	(20.5)	1	(0.2)	483	(100)	69.0
50+	126	(23.1)	276	(50.5)	142	(26.0)	2	(0.4)	546	(100)	78.0
Total	516	(35.3)	584	(40.0)	358	(24.5)	3	(0.2)	1461	(100)	69.6

Device use by education level

	Smartphone		Tablet		PC		Unknown		Total	
Education	n	(%)	n	(%)	n	(%)	n	(%)	n	(%)
Low	48	(30.4)	84	(53.2)	26	(16.5)	0	(0)	158	(100)
Middle	198	(39.4)	186	(37.0)	119	(23.7)	0	(0)	503	(100)
High	241	(35.2)	252	(36.8)	189	(27.6)	2	(0.3)	684	(100)
Other	29	(25.0)	62	(53.4)	24	(20.7)	1	(0.9)	116	(100)
Total	516	(35.3)	584	(40.0)	358	(24.5)	3	(0.2)	1461	(100)

Design of the Survey Interface

At the start of the survey, respondents were randomized into one of two interface conditions:

1) In the manual-forward version, respondents had to navigate between survey web pages using 'previous' and 'next' buttons (the default design in surveys fielded by SN.
2) In the auto-forward version, respondents were auto-forwarded to the subsequent survey web page when they answered the last question, unless that last question was a 'check all that apply' question or an open-ended question.

Within the auto-forward condition, auto-forwarding was applied for 75.5% of the pages. For 24.5% of the pages which contained 'check all that apply' questions or open questions, manual-forwarding was applied. Respondents were required to answer every question within the survey except for questions about sexuality.

The auto-forward interface included 'previous' and 'next' buttons and was completely similar in the visual design to the manual-forward interface (see Figure A1 in the Appendix). Respondents could thus navigate backward and forward in the auto-forward condition when they, for example, wanted to correct an answer provided earlier in the survey or review a previous question. We decided to include the 'next' button in the auto-forward condition to avoid confusion between pages where auto-forward was possible and those where it was not. Respondents were not informed about the auto-forward design prior to the survey start.

Data Preparation

Before we move to the analysis methods, we first describe the data preparation. The data preparation consisted of three steps: selection of complete responses, processing of paradata, and omission of outliers.

As a first step, we selected only those cases with complete data. We removed the 74 sample units who broke off as they provided only partial information on response times. Given the small size of this group, we decided not to complicate our analyses by including censored data. After the selection, we had 713 respondents in the auto-forward condition and 748 respondents in the manual-forward condition.

As a second step, we translated the web survey paradata to meaningful features and variables. We coded the device that respondents used to complete the survey using user agent strings. Whenever a person accesses any website, the website receives information. This information is referred to as the user agent string and contains characteristics of the device in order for the website to be able to adapt to the device. These strings have a known format and allow one to derive the type of

device. For three respondents, the user agent string showed that a mobile device was used, but it was unclear whether it was a smartphone or a tablet. We excluded these three respondents from the analysis. Some respondents (n=53) switched between devices during the survey. In the analysis, these respondents are allocated to the device in which they answered the majority of the questions. Next, we processed the survey web page response times. The page-level response time was calculated as the difference between the time stamp of entering a page and the time stamp of leaving the page. The total response time (i.e., respondent-level) was calculated by summing up the page-level response times for a respondent. Since both respondent-level and page-level response times are right-skewed, we applied a log transformation to the response times.

In the third step, we removed outliers at the respondent level and at the page level. We applied the interquartile rule for outliers for both respondent-level and page-level outliers. We calculated the interquartile range (IQR) for the data, multiplied the IQR by 1.5, and added this to the third quartile (Upton & Cook, 1996). A log-transformed response time was marked as an outlier if it was larger than the third quartile plus 1.5 times the IQR. At the respondent level, 14 respondents were removed based on the interquartile rule, leading to 1,444 respondents (705 in the auto-forward design and 739 in the manual-forward design). At the page level, about three percent of the log-transformed response times were removed (i.e., 4,589 out of 152,423 log-transformed response times).

In the following sections, all response times are transformed back from the log scale to aid interpretation.

Analysis

We answer the four research questions through three analyses. We use multi-level analysis to answer the first research question on response times. We use standard regression analysis explaining the numbers of navigational actions to answer the second and third research questions. We use Chi-square tests to answer the final research question on the choice of 'I don't know' and 'I rather not say' responses. All analyses were conducted in R version 3.6.2 (R Core Development Team, 2019).

Multi-level analysis of log response times. Similar to Antoun and Cernat (2019), the page-level log-transformed response times form the dependent variable in the analysis which are clustered by adding a level for the respondent and a level for the page. The respondent-specific influence and the page-specific influence are entered as a random effect. We include experimental condition, age, education and type of device as explanatory variables at the respondent-level and include respondent random effects that vary across age and device groups.

Regression analysis of navigation behavior by clicks and taps (from here on called clicks). We first investigated the clicks between conditions with descriptive

statistics using normalized data, meaning the number of clicks was divided by the number of respondents in each group.

Secondly, we conducted a regression analysis where we included all the previous button clicks as well as the unnecessary use of the 'next' button (i.e., failed attempts to proceed to the next page). To minimize item non-response, all survey questions - except the questions about sexuality - were mandatory. Clicking the 'next' button without answering the question thus resulted in a warning message that a question was left unanswered preventing moving forward to the next page. The unnecessary clicks were all caused by manually clicking the 'next' button while not having answered all of the questions on a page.

For more insight, we followed-up with an investigation of the 10 pages where differences in clicks between the two conditions were the largest. The difference in clicks was calculated by taking the absolute difference between the number of clicks per page per type of navigation button in the manual-forward condition and the auto-forward condition.

Chi-square tests for the answer options 'I don't know' and 'I rather not say.' For each answer option, a Chi-square test is conducted to test in which condition this type of answer is used the most. To simplify the analysis, we compared respondents that never chose such answers to respondents that chose such answers at least once.

Results

Does auto-forwarding reduce response times?

Table 2 shows several models to explain the variance in the log-transformed response time. In the empty model (i.e., the model with no predictors), 60% of the variance in the log response time was explained by the page and 10% by the respondent. The full model only included variables related to the respondent and this model explained 24% of the respondent variance.

These results confirm our first hypothesis (H1) that auto-forwarding reduces the total response times. When correcting for education, age, device, and including the interaction of device and age, respondents in de auto-forward condition required on average 0.65 seconds less per page than respondents in the manual-forward condition (10.97 vs. 11.61 seconds). The survey consisted of an average of 106.6 pages, which, thus, translates to an average 68.9 seconds reduction of the total completion time.

Table 2 Log-transformed response time per page predicted by type of navigation, age, education, device for the total data and split between pages with single choice & matrix questions (only pages with auto-forward) and open ended & check all that apply questions (only pages with manual-forward)

Model	Empty model		+ navigation		+ age and education		+ device		+ device * age		Only pages with auto- forward		Only pages with manual- forward	
	Coeff.	(s.e.)	Coeff.	(s.e.)	Coeff.	(s.e.)	Coeff.	(s.e.)	Coeff.	(s.e.)	Coeff.	(s.e.)	Coeff.	(s.e.)
Fixed part														
Intercept	2.45	(.04) ***	2.48	(.04) ***	2.46	(.04) ***	2.45	(.04) ***	2.45	(.04) ***	2.32	(.04) ***	2.82	(.07) ***
Navigation (Ref. = Manual-forward)														
Auto-forward			-0.05	(.01) ***	-0.06	(.01) ***	-0.06	(.01) ***	-0.06	(.01) ***	-0.07	(.01) ***	0.02	(.01)
Education (Ref. = Low)														
Middle					-0.05	(.02) **	-0.04	(.02) *	-0.04	(.02) *	-0.05	(.02) ***	-0.03	(.02)
High					-0.13	(.02) ***	-0.12	(.02) ***	-0.12	(.02) ***	-0.12	(.02) *	-0.09	(.02) ***
Other					0.01	(.03)	0.01	(.03)	0.01	(.03)	0.01	(.03) ***	0.00	(.03)
Age (Ref. = 16-29)														
30-49					0.08	(.01) ***	0.07	(.01) ***	0.06	(.02) **	0.05	(.02) **	0.12	(.02) ***
50+					0.21	(.01) ***	0.20	(.01) ***	0.20	(.02) ***	0.18	(.02) ***	0.25	(.03) ***
Device (Ref. = Smartphone)														
Tablet							0.05	(.01) ***	0.05	(.02) *	0.05	(.02) *	0.09	(.03) **
PC							-0.04	(.01) ***	-0.06	(.02) **	-0.06	(.02) *	-0.03	(.03)
*Device * Age*														
Tablet * 30-49									-0.00	(.03)	-0.01	(.03)	-0.05	(.04)
PC * 30-49									0.04	(.03)	0.04	(.03)	-0.08	(.04)
Tablet * 50+									-0.01	(.03)	-0.01	(.03)	-0.01	(.04)
PC * 50+									0.02	(.03)	0.02	(.03)	-0.02	(.04)

Model	Empty model		+ navigation		+ age and education		+ device		+ device * age		Only pages with auto- forward		Only pages with manual- forward	
	Coeff.	(s.e.)	Coeff.	(s.e.)	Coeff.	(s.e.)	Coeff.	(s.e.)	Coeff.	(s.e.)	Coeff.	(s.e.)	Coeff.	(s.e.)
Random part														
$\sigma^2_{respondent}$	0.06	(.24)	0.06	(.24)	0.04	(.21)	0.04	(.21)	0.04	(.21)	0.04	(.21)	0.05	(.22)
σ^2_{page}	0.28	(.53)	0.28	(.53)	0.28	(.53)	0.28	(.53)	0.28	(.53)	0.21	(.46)	0.23	(.48)
$\sigma^2_{residuals}$	0.14	(.37)	0.14	(.37)	0.14	(.37)	0.14	(.37)	0.14	(.37)	0.13	(.36)	0.17	(.42)
Marginal R^2	0.00		0.00		0.03		0.03		0.03		0.04		0.04	
Conditional R^2	0.71		0.71		0.71		0.71		0.71		0.68		0.63	
Deviance	131773		131757		131393		131337		131335					
AIC	131781		131767		131419		131367		131373					
BIC	131821		131817		131547		131516		131562					

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

The reduction in response time was only observed for pages where an auto-forward functionality could be applied (i.e., pages with only single choice or matrix questions). On these pages, the auto-forward functionality resulted in a 0.72 second or 7.1% reduction in response time (10.16 vs. 9.44 seconds); $t(1,417) = -6.64$, $p < .001$. On the other pages (i.e., pages with open-ended and check-all-that-apply questions), we observed a 0.28 seconds increase in response time (16.85 vs. 17.13 seconds). The latter difference is not significant; $t(1,420) = 1.27$, $p = .20$.

As for education, higher-educated respondents completed the survey faster than lower-educated respondents; $t(1,426) = -5.98$, $p < .001$. Furthermore, older respondents needed more time to complete the survey than the other age groups, with the youngest respondents being the fastest; $t(1,776) = 8.41$, $p < .001$. Tablet users needed more time to complete the survey than smartphone users: $t(1,672) = 2.04$, $p = .04$, while PC users needed less time: $t(2,055) = -2.86$, $p = .004$. Finally, we did not find interaction effects between age and device type.

Does auto-forwarding lead to more efficient navigation through the survey, and, if so, is any improvement related to type of device?

Contrary to our hypothesis (H2), auto-forwarding led to less efficient navigation through the survey. When looking at all navigations (i.e., automated navigations and the manual clicks), auto-forwarding increased the average number of clicks to the previous page by 1.0 (auto-forward: M = 2.9, SD = 6.4; manual-forward: M = 1.9, SD = 2.9) and the (attempted) navigations to proceed to the next page increased by 16.0 (auto-forward: M = 137.3, SD = 23.2; manual-forward: M = 121.3, SD = 9.6). The unnecessary clicks, which are all caused by manual clicking, account for 16.0% of the total next-page navigations and are also more frequent in the auto-forward condition (auto-forward: M = 28.1, SD = 20.8; manual-forward: M = 13.5, SD = 5.3).

The results presented in Table 3 confirm that both buttons (i.e., all 'previous' button clicks and unnecessary 'next' button clicks) are used significantly more often in the auto-forward condition. An effect for device was only apparent for respondents aged 50 and older, who used the navigation buttons less when using a PC than when using a mobile device (i.e., a tablet or a smartphone). This finding is in the opposite direction of hypothesis (H3). Furthermore, we found fewer clicks for the higher-educated respondents.

To understand these results better, we examined pages where differences in clicks between the two conditions were the strongest. Tables A1 and A2 (see Appendix) provide an overview of the pages with the largest difference in clicks per type of navigation, including the difference in the number of clicks between the conditions.

As Table A1 shows (see the Appendix), the 'previous' button is used most in the auto-forward condition when questions are cognitively demanding, when a new

Table 3 Regression analyses with number of clicks per person as a dependent variable

	Estimate (B)	SE	t
Intercept	13.51 ***	0.93	14.48
Button (Ref. = Next)			
Previous	-11.51 ***	0.58	-20.00
Condition (Ref. = Manual-forward)			
Auto-forward	14.47 ***	0.58	24.82
Device (Ref. = Smartphone)			
Tablet	1.43	0.94	1.52
PC	-0.19	0.91	-0.21
Age (Ref. = 16-29)			
30-49	0.02	0.80	0.03
50+	1.81 *	0.91	2.00
Education (Ref. = Low)			
Middle	-0.36	0.74	-0.49
High	-1.52 *	0.73	-2.09
Other	-1.05	0.97	-1.08
Button * Condition			
Previous * Auto-forward	-13.62 ***	0.82	-16.54
Device * Age			
Tablet * 30-49	0.72	1.23	0.58
PC * 30-49	-0.97	1.33	-0.73
Tablet * 50+	0.01	1.27	0.01
PC * 50+	-4.01 **	1.33	-3.02
$R^2 = 0.48$			

*$p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

topic is introduced, or when respondents think they might have answered a question already (i.e., respondents check the previous question because of similarities in question wordings). Within the auto-forward condition, we do not find increased use of the previous button between pages with automated navigation and pages with manual navigation, indicating respondents are not confused by this transition; $t(356) = 0.43$, $p = .67$.

Respondents unnecessarily use the 'next' button most in the auto-forward condition. This finding is most apparent on pages with multiple questions (see Table 4, Table A2 in the Appendix). On those pages, multiple single-choice questions were presented.

Table 4 Regression analyses with the frequency of using the 'next' button unnecessarily per page as a dependent variable

	Estimate (B)	SE	t
Intercept	3.10 ***	0.29	10.64
Condition (Ref. = Manual-forward)			
Auto-forward	0.93 *	0.40	2.36
Number of questions (Ref. = 1)			
2	1.96 ***	0.25	7.68
> 2	2.02 ***	0.32	6.30
Question type (Ref. = open/check-all that apply)			
Single-choice or matrix	-0.86 **	0.28	-3.12
Number of questions * condition			
2 questions * Auto-forward	-1.17 ***	0.35	-3.37
> 2 questions * Auto-forward	-0.74	0.44	-1.68
Question type * Condition			
Single-choice or matrix * auto-forward	0.83 *	0.38	2.17
$R^2 = 0.27$			

*$p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Does auto-forwarding affect how often the answer options 'I don't know' and 'I rather not say' are used?

An auto-forward functionality had no effect on how often respondents gave either an 'I rather not say' or an 'I don't know' answer. Contrary to our expectations (H4), these two answer options were used just as often in the auto-forward condition as in the manual-forward condition. The answer 'I rather not say' was given at least once by 79.0% of the respondents in the manual forward condition and by 80.1% in the auto-forward condition; $\chi 2(1, N=1,444) = 0.28$, $p=.59$. The answer 'I don't know' was given at least once by 30.6% of the respondents in the manual-forward condition and by 33.3% in the auto-forward condition; $\chi 2(1, N=1444) = 1.29$, $p=.26$.

Conclusion

In this study, we randomly assigned respondents to an auto-forward design or a manual-forward design in a long mixed-device web survey on health. We compare these two conditions across devices used for survey completion (PC, tablet, and smartphone). We find slightly shorter completion times for all devices in the

auto-forward design compared to the manual-forward design. Results also show that questions on pages with automated navigation are answered significantly faster than the questions on pages with manual navigation (i.e., where respondents needed to use the navigation buttons).

However, the difference in completion times between the conditions is relatively small. Therefore, we used paradata to investigate how respondents navigated the survey. Analyses of clicks on the 'previous' button show that it is used more often in the auto-forward condition compared to the manual-forward condition. Such increased use might be explained by the novelty of the design and its pace: respondents may not be used to automated navigation within a survey. Within the auto-forward condition, we do not find more use of the 'previous' button between pages with automated navigation and pages with manual navigation, indicating that respondents are not confused by this transition.

We also find that respondents in the auto-forward condition unnecessarily use the 'next' button (i.e., failed attempts to proceed to the next page). This finding may be explained by the following reasons: 1) respondents wish to navigate faster than the pace of the automated navigation of the survey, 2) respondents are not used to an auto-forward design and use the 'next' button as a common habit, 3) respondents did not notice that a new page with a new question has finished loading, or 4) respondents who mistakenly missed a question might think they should click the next button because they are not taken to the next page. However, in reality, these respondents forgot to fill in a question and for that reason they do not automatically go to the next page. Only after filling in the overlooked question, they will automatically be forwarded to the next page (i.e., the next button should not be used in this situation).

Contrary to our expectation, we found no significant difference for the use of 'I don't know' and 'I rather not say' answers between the auto-forward and manual-forward condition.

This result deviates from the outcomes of Selkälä et al.'s study (2020). The difference between their study and ours is that we used a long survey with different types of questions with a mix of auto-forward and manual-forward which may have affected answering behavior differently.

Discussion

Overall, we conclude that auto-forwarding can be used to reduce completion times. Since it is difficult to include auto-forwarding with check-all-that-apply, open and numerical questions we advise to carefully consider mixing manual and auto-forwarding within one survey. Ideally, survey layout and navigation should be predictable within a survey and across devices (Antoun, Katz, Argueta, & Wang, 2018).

In line with the recommendations of Giroux et al. (2019), we advise to include clear instructions to inform respondents about their navigation possibilities within the survey. A particular challenge for future research is how to implement auto-forwarding in surveys that include different types of questions.

Our study has some limitations. The main limitation, as mentioned above, is that our survey contained questions in which auto-forward cannot be applied. Future research should replicate our design in a long survey where auto-forward can be applied to all questions. A second limitation is the self-selection of respondents to complete the survey on a mobile device. Random assignment of respondents to a certain device leads to issues of respondent noncompliance (de Bruijne & Wijnant, 2013; Mavletova, 2013; Wells, Bailey, & Link, 2014). Therefore, our sample was composed of earlier respondents to SN individual surveys that responded at least once with a mobile device. Those respondents are likely to be more motivated than a freshly recruited cross-section.

Another further step would be to examine the quality of answers provided to different auto-forward interface conditions in more detail. We only explored the impact of auto-forwarding on item nonresponse. Furthermore, we advise to evaluate users' experience of the auto-forward interface in more detail pre- or post-survey, for example, by conducting semi-structured open interviews and adding open-ended evaluation questions.

Data Availability

The data are available on site or by means of remote access. This can be requested by contacting the corresponding author at j.bakker@cbs.nl .

Software Information

We used R version 3.6.2 (R Core Development Team, 2019). The R-script can be requested by contacting the corresponding author at j.bakker@cbs.nl . Paradata were collected using Version 5.0.5 of the BLAISE computer-assisted interviewing system (Blaise, 2018).

References

American Association for Public Opinion Research (2016). Standard definitions: Final dispositions of case codes and outcome rates for surveys. Ann Arbor, MI: AAPOR.

Antoun, C., & Cernat, A. (2019). Factors affecting completion times: A comparative analysis of smartphone and PC web surveys. *Social Science Computer Review*, *38*, 477-489. https://doi.org/10.1177/0894439318823703

Antoun, C., Katz, J., Argueta, J., & Wang, L. (2018). Design heuristics for effective smartphone surveys. *Social Science Computer Review*, *36*, 557-574. https://doi.org/10.1177/0894439317727072

Arn, B., Klug, S., & Kolodziejski, J. (2015). Evaluation of an adapted design in a multi-device online panel: A DemoSCOPE case study. *Methods, data, analyses*, *9*, 185-212. https://doi.org/10.12758/mda.2015.011

Bergstrom, J. C., Lakhe, S., & Erdman, C. (2016). Navigation buttons in web-based surveys: Respondents' preferences revisited in the laboratory. *Survey Practice*, *9*. https://doi.org/10.29115/SP-2016-0005

Blaise (2018). https://www.blaise.com (accessed: August 2018).

Cook, W. A. (2014). Is mobile a reliable platform for survey taking? Defining quality in online surveys from mobile respondents. *Journal of Advertising Research, 54*, 141–148. https://doi.org/10.2501/JAR-54-2-141-148

Couper, M. P., Antoun, C., & Mavletova, A. (2017). Mobile web surveys: A total survey error perspective. In P. Biemer, S. Eckman, B. Edwards, E. de Leeuw, F. Kreuter, L. Lyberg, C. Tucker, & B. West (Eds.), *Total Survey Error in Practice*, (pp. 133-154). New York, NY: Wiley.

Couper, M. P., & Peterson, G. J. (2017). Why do web surveys take longer on smartphones? *Social Science Computer Review*, *35*, 357-377. https://doi.org/10.1177/0894439316629932

De Bruijne, M.A. (2015). Designing web surveys for the multi-device internet. PhD thesis., The Netherlands: Tilburg University: Center for Economic Research.

De Bruijne, M.A. (2016). Online vragenlijsten en mobiele devices (Online questionnaires and mobile devices). *Jaarboek van de Marktonderzoekassociatie*, 137-151. https://adoc.pub/9-online-vragenlijsten-en-mobiele-devices.html

De Bruijne, M.A., & Wijnant, A. (2013). Comparing survey results obtained via mobile devices and computers: An experiment with a mobile web survey on a heterogeneous group of mobile devices versus a computer-assisted web survey. *Social Science Computer Review*, 31(4), 482-504. https://doi.org/10.1177/0894439313483976

Giroux, S., Tharp, K., & Wietleman, D. (2019). Impacts of implementing an automatic advancement feature in mobile and web surveys. *Survey Practice*, *12*. https://doi.org/10.29115/SP-2018-0034

Hays, R. D., Bode, R., Rothrock, N., Riley, W., Cella, D., & Gershon, R. (2010). The impact of next and back buttons on time to complete and measurement reliability in computer-based surveys. *Quality of Life Research*, *19*, 1181-1184. https://doi.org/10.1007/s11136-010-9682-9

Hintze, D., Findling, R, D., Scholz, S., & Mayrhofer, R. (2014, December). Mobile device usage characteristics: The effect of context and form factor on locked and unlocked usage. In *Proceedings of the 12th International Conference on Advances in Mobile Computing and Multimedia* (pp. 105-114). ACM. https://doi.org/10.1145/2684103.2684156

Kreuter, F. (2013) Improving surveys with paradata: Introduction. In F. Kreuter (Ed.)*, Improving surveys with paradata: Analytic uses of process information*, (pp. 1-9). New York, NY: Wiley.

Lugtig, P., Toepoel, V., Haan, M., Zandvliet, R., & Klein Kranenburg, L. (2019). Recruiting hard-to-reach groups into a probability-based online panel by promoting smartphone use. M*ethods, data, analyses*, *13*, 291-306. https://doi.org/10.12758/mda.2019.04

Mavletova, A. (2013). Data quality in PC and mobile web surveys. *Social Science Computer Review*, *31*, 725-743. https://doi.org/10.1177/0894439313485201

R Core Development Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/

Roberts, A., de Leeuw, E. D., Hox, J., Klausch, T., & de Jongh, A. (2012). Leuker kunnen het wel maken. Online vragenlijst design standaard matrix of scrollmatrix. *In het 38e jaarboek van de MOA : Developments in Market Research*, (pp. 133-148). http://dspace.library.uu.nl/handle/1874/291084

Selkälä A., Callegaro M., & Couper M.P. (2020) Automatic Versus Manual Forwarding in Web Surveys - A Cognitive Load Perspective on Satisficing Responding. In: Meiselwitz G. (eds). *Social Computing and Social Media. Design, Ethics, User Behavior, and Social Network Analysis. HCII 2020. Lecture Notes in Computer Science*, vol 12194. Springer, Cham. https://doi.org/10.1007/978-3-030-49570-1_10

Selkälä, A., & Couper, M.P. (2018). Automatic versus manual forwarding in web surveys S*ocial Science Computer Review*, *36*, 669-689. https://doi.org/10.1177/0894439317736831

Struminskaya, B., Weyandt, K., & Bosnjak, M. (2015). The effects of survey completion using mobile devices on data quality - Evidence from a probability-based general population panel. *Methods, data, analyses*, *9*, 261–292. https://doi.org/10.12758/mda.2015.014

Upton, G. & Cook, I. (1996). Understanding Statistics. Oxford: Oxford University Press.

Wells, T., Bailey, J. T., & Link, M. W. (2014). Comparison of smartphone and online computer survey administration. *Social Science Computer Review, 32*, 238-255. https://doi.org/10.1177/0894439313505829

Zhang, C. & Conrad, F. (2014). Speeding in web surveys: The tendency to answer very fast and its association with straightlining. *Survey Research Methods*, *8*, 127-135. https://doi.org/10.18148/srm/2014.v8i2.5453

Appendix

Table A1 Top 10 pages with the largest difference in click behavior of the 'previous' button between the auto-forward and manual-forward condition.

	Auto-forward		Manual-forward		Difference			
Page nr*	n	%	n	%	n	%	Topic	Remarks/difficulties
220	38	1.9	7	0.5	31	3.6	Occupational accident	Intro page + long text
214	55	2.7	27	1.9	28	3.2	Eating vegetables	Similar question block to previous** + specifications
60	35	1.7	10	0.7	25	2.9	Education	Similar question block to previous** + specifications
81	30	1.5	10	0.7	20	2.3	Education finished	-
184	31	1.5	11	0.8	20	2.3	Blood sugar	-
218	23	1.1	4	0.3	19	2.2	Eating fruit	Relates to the previous page
255	33	1.6	14	1.0	19	2.2	Alcohol consumption	Very similar to the previous question (4 vs. 6 glasses)
49	28	1.4	10	0.7	18	2.1	Paid work	Numeric question + conditional 2nd quest.
219	25	1.2	8	0.6	17	1.9	Eating fish	Similar questions to previous[1] + specifications
225	23	1.1	6	0.4	17	1.9	Accidents	Intro page + long text + possible sensitive subject
Total	321	15.7	107	7.6	214	24.6		

*The page number is the raw numbering according to the programming of the survey. Many pages are not shown to respondents due to routing.
**Similar question block to the previous question block refers to the almost exact same wording of the blocks.

Table A2 Top 10 pages with the largest difference in unnecessary click behavior of the 'next' button* between the auto-forward and manual-forward condition

	Auto-forward		Manual-forward		Difference				
Page nr**	n	%	n	%	n	%	Topic	Remarks/difficulties	# Questions
156	321	1.6	36	0.4	285	2.8	Chronic disease	Matrix questions	10
158	288	1.5	24	0.2	264	2.6	Psychological health	Matrix questions	5
155	321	1.6	79	0.8	242	2.4	Chronic disease	Matrix questions	11
159	266	1.3	30	0.3	236	2.3	Acute illness	Matrix questions	6
257	256	1.3	44	0.4	212	2.1	Narcotic use	Matrix questions	11
59	359	1.8	150	1.5	209	2.0	Education	Intro page + long introduction	1
11	842	4.3	634	6.4	208	2.0	Household info	Multiple questions on screen	3
157	263	1.3	102	1.0	161	1.6	Chronic disease	2nd question on same page is conditional	2
148	170	0.9	25	0.3	145	1.4	Diabetes	2nd and 3rd questions are conditional	3
160	154	0.8	9	0.1	145	1.4	Pain	2nd question on same page is conditional	2
Total	3,240	16.4	1,133	11.4	2,107	20.5			

* Failed attempts to proceed to the next page.
**The page number is the raw numbering according to the programming of the survey. Many pages are not shown to respondents due to routing.

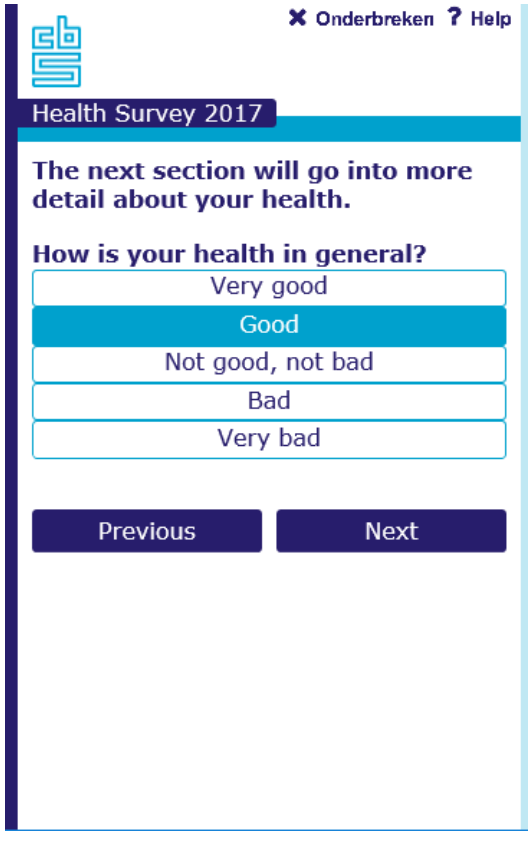

Figure A1 Screenshot of the survey layout