SSOAR

Open Access Repository
www.ssoar.info

How to Optimize Online Mixed-Device Surveys: The Effects of a Messenger Survey, Answer Scales, Devices and Personal Characteristics

Menken, Caroline Marjanne; Toepoel, Vera

Veröffentlichungsversion / Published Version
Zeitschriftenartikel / journal article

Empfohlene Zitierung / Suggested Citation:
Menken, C. M., & Toepoel, V. (2023). How to Optimize Online Mixed-Device Surveys: The Effects of a Messenger Survey, Answer Scales, Devices and Personal Characteristics. *Methods, data, analyses : a journal for quantitative methods and survey methodology (mda)*, *17*(1), 47-70. https://doi.org/10.12758/mda.2022.08

Nutzungsbedingungen:
Dieser Text wird unter einer CC BY Lizenz (Namensnennung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier:
https://creativecommons.org/licenses/by/4.0/deed.de

Terms of use:
This document is made available under a CC BY Licence (Attribution). For more Information see:
https://creativecommons.org/licenses/by/4.0

gesis
Leibniz-Institut
für Sozialwissenschaften

Mitglied der
Leibniz-Gemeinschaft

methods, data, analyses | Vol. 17(1), 2023, pp. 47-70 DOI: 10.12758/mda.2022.08

How to Optimize Online Mixed-Device Surveys: The Effects of a Messenger Survey, Answer Scales, Devices and Personal Characteristics

Caroline Marjanne Menken & Vera Toepoel
University of Utrecht

Abstract

The goal of this research was to determine the best way to present mixed-device surveys. We investigate the effect of survey method (messenger versus regular survey), answer scale, device used, and personal characteristics such as gender, age and education on break-off rate, substantive answers, completion time and respondents' evaluation of the survey. Our research does not suggest that a messenger survey affects mixed-device surveys positively. Further research is necessary to investigate how to optimally present mixed-device surveys in order to increase participation and data quality.

Keywords: Mixed-device survey, messenger, answer scale, online survey, mobile friendly

© The Author(s) 2023. This is an Open Access article distributed under the terms of the Creative Commons Attribution 3.0 License. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

Online surveys are often used by researchers (Schlosser & Mays, 2018; Zhang, Kuchinke, Woud, Velten & Margraf, 2017). A traditional online survey is designed to be completed on a computer (Schlosser & Mays, 2018). However, since the use of mobile devices has grown, online surveys are also being completed on other devices such as mobile phones and tablets (De Bruijne & Wijnant, 2013; Mavletova, 2013; Millar & Dillman, 2012). Surveys that are being completed on different devices are called mixed-device surveys (Toepoel & Lugtig, 2015). What is the best way to present these mixed-device surveys?

Previous research has shown that there are several differences between devices, when it comes to response behavior (Couper & Peterson, 2017; Schlosser & Mays, 2018). For instance, the screen size of mobile phones is smaller than the screen size of computers and tablets (Schlosser & Mays, 2018), making it more difficult to answer questions on a mobile phone (De Bruijne & Wijnant, 2014; Mavletova, 2013). Especially close-ended questions with many answer options are not desirable on a small screen, because some answer options fall off screen (Couper & Peterson, 2017) and respondents need to scroll to see all the answer options. Alternatives are open-ended questions or close-ended questions with few answer options. However, research shows open-ended questions take a lot of effort to answer and result in higher (item) nonresponse (Couper & Peterson, 2017; De Bruijne & Wijnant, 2014; Couper & Peterson, 2017; Mavletova, 2013; Schlosser & Mays, 2018). Often, researchers make online surveys more suitable to be completed on mobile phones, for instance by making the design of the survey responsive. With a responsive design, the layout of the survey adapts to the device being used (Antoun, Katz, Argueta & Wang, 2018; De Bruijne & Wijnant, 2014; Mavletova, 2013). However, although a responsive design makes it easier for respondents to complete online surveys on mobile phones, it might still not be optimal for data quality (Antoun et al., 2018). A solution to increase the data quality of online surveys could be to make the survey more interactive by adding a conversational element (Kim, Lee & Gweon, 2019). Since mobile phones are mostly used for short messaging, a Whatsapp-type messenger survey could be a way to make an online survey more suitable for mobile phones. A research messenger survey is more similar to text messaging and adds a conversational element to the online survey.

In this study, we randomly assigned respondents in the American Amazon MTurk Panel to a regular responsive online survey design and a messenger survey. In addition, we randomly assigned panel members to a closed-ended answer scale with many answer options, a closed-ended answer scale with few answer options (that would fit small screens of mobile phones), and an open-ended answer scale to investigate the effect of survey method and type of response format. We also

Direct correspondence to
Caroline Menken, University of Utrecht
E-mail: caroline-menken@hotmail.com

investigate the effect of the device used to complete the survey and the effect of personal characteristics such as age, gender and education. We compare break-off rate, substantive answers, completion time and respondents' evaluation of the survey to provide suggestions on how to optimally design mixed-device surveys. We conducted exploratory research.

Theoretical Background

Online Surveys

Since Internet has become more and more important in daily life, the use of online surveys has grown (Schlosser & Mays, 2018; Solomon, 2000; Zhang, Kuchinke, Woud, Velten & Margraf, 2017). Online surveys have a number of advantages and disadvantages. One of the advantages is that there is no need for interviewers (Tourangeau, Maitland, Rivero, Sun, Williams & Yan, 2017). Therefore, online surveys are anonymous, which reduces socially desirable responding (Tourangeau et al., 2017). Furthermore, online survey research takes less time and the costs of online survey research are typically low (Couper & Miller, 2008; Solomon, 2000; Wright, 2005; Zhang et al., 2017). For respondents it takes less effort to participate in a survey since respondents can complete online surveys in their own time and space (Solomon, 2000; Wright, 2005; Zhang et al., 2017). However, the absence of an interviewer also has disadvantages (Bowling, 2005; De Leeuw, 2008; Kim, Lee & Gweon, 2019). An interviewer can give the respondent additional instructions and can clarify the questions when needed (Bowling, 2005; De Leeuw, 2008; Harris & Brown, 2010). Besides, the presence of an interviewer leads to a higher response rate and a higher completion rate since an interviewer can convince and motivate reluctant respondents to participate and to finish the survey (De Leeuw, 2008; Heiervang & Goodman, 2011; Kim, Lee & Green, 2019). If an interviewer is present, a survey is more similar to a conversation (Bowling, 2005).

Mixed Devices

A traditional online survey is designed to be completed on a computer (Cunningham, Neighbors, Bertolet & Hendershot, 2013; Schlosser & Mays, 2018). However, nowadays online surveys are also being completed on other devices such as mobile phones and tablets (De Bruijne & Wijnant, 2013; Mavletova, 2013; Millar & Dillman, 2012). There are several differences between the different devices (Couper & Peterson, 2017; Schlosser & Mays, 2018). First of all, the screen of a mobile phone is much smaller than the screen of a computer (Maslovskaya, Smith & Durrant, 2020; Schlosser & Mays, 2018; Stapleton, 2013). So, it takes longer to complete a

survey on a mobile phone than on a computer. The time to complete a survey on a tablet is typically in between (Couper & Peterson, 2017). The Internet connection on mobile phones is often slower than on computers, also leading to a higher completion time (Couper & Peterson, 2017; Schlosser & Mays, 2018). Furthermore, respondents who complete a survey on a mobile phone or tablet are more likely to do this away from home, and could therefore be distracted (Couper & Peterson, 2017; Maslovskaya, Smith & Durrant, 2020; Schlosser & Mays, 2018). The higher completion time on mobile phones often leads to a higher break-off rate (Couper & Peterson, 2017; Cunningham et al., 2013; Mavletova, 2013; Schlosser & Mays, 2018). Respondents generally experience a higher difficulty to complete an online survey on a mobile phone (De Bruijne & Wijnant, 2013). Research shows that the response-rate of traditional, non-optimized online surveys on mobile phones is very low (Mavletova, 2013).

Additionally, the usage of mobile phones is different than the usage of computers and tablets. Mobile phones are mostly being used to communicate through messaging apps, especially Whatsapp (O'Hara, Massimi, Harper, Rubens & Morris, 2014). Whatsapp is being used for sharing information, images and videos and for ongoing conversations by sending short messages (Ahad & Lim, 2014; O'Hara et. al., 2014). Mobile phones are used for more casual conversations, whereas computers are generally being used for more formal communication (O'Hara et. al., 2014).

Mobile Friendly Survey

Antoun and others (2018) created general guidelines to alter the design of a survey to make it suitable for mobile phones. First of all, it should be easy for respondents to read the questions and answer options. The font size should be large enough and answer options should be large enough to be easily selected by respondents with touch screen. In addition, the content of the survey should fit the width of the screen. If not all of the answer options fit on screen, the answer options should be presented vertically not horizontally. Besides, the features of the design should be simple, and respondents should be able to understand how to use them. At last, the design should work on different devices. A way to achieve this is by making the design responsive. With a responsive design the layout of the survey adapts to the device being used (Harb, Kapellari, Luong & Spot, 2011; Hussain & Mkpojiogo, 2015). The layout adapts to suit different screen sizes, larger buttons and texts are provided when using a mobile phone and non-essential elements are being hidden when the screen is small (Harb et al., 2011).

Several researchers have created a mobile friendly design to make a traditional online survey more suitable to be completed on mobile phones. For instance, De Bruijne & Wijnant (2013) made the content of their survey fit the width of the screen and made the font size larger. Moreover, the answer options were made wide

buttons in order to be easily selected with touch screen. The answer options were also presented vertically instead of horizontally. However, respondents who completed the survey with a mobile friendly design still reported a longer completion time than the respondents who completed a traditional online survey on a computer. This implies that completing a survey on a mobile phone still takes more effort and time, even when the survey is adapted to mobile phones. This conclusion is supported by other researchers who investigated the differences between a survey with a mobile friendly design and a traditional online survey (Couper & Peterson, 2017; Mavletova, 2013; Schlosser & Mays, 2018). Antoun and others (2018) also concluded that the guidelines might not be enough to make an online survey optimal for mobile phones. Therefore, more research into how to present an online survey on mobile phones is needed.

Answer Scales

Survey questions can have different types of answer scales. Questions could be open-ended, close-ended with many answer options or close-ended with few answer options. The non-response rate for open-ended questions is higher than for close-ended questions in online surveys (Reja, Manfreda, Hlebec & Vehovar, 2003) because it takes more effort and time for the respondents to answer the open-ended questions (Couper & Peterson, 2017). Close-ended questions have a number of answer options. Respondents tend to choose one of the answer options even if their true answer is not one of the options (Reja et al., 2003). Few answer options can give respondents too little information. The chance that their true answer is not one of the options is higher (Reja et al., 2003). However, too many options make the questions too complicated for respondents. This can lead to not considering all the options (Chung et al., 2010). The order in which the answer options are presented in close-ended questions also affects answers. Respondents are more likely to choose the answer options that are visualized on the screen than the answer options that fall off the screen (De Bruijne & Wijnant, 2013; Mavletova, 2013; Stapleton, 2013). Thus, too many answer options in mixed-device surveys might not be desirable.

In online surveys, the answers on open-ended questions completed on a mobile phone are shorter than the answers on open-ended questions completed on a computer (Mavletova, 2013). According to respondents, it is easier to type an answer on a computer keyboard (Mavletova, 2013). Therefore, it takes more time to answer open-ended questions on mobile phones than on a computer (Couper & Peterson, 2017). In addition, the completion time for close-ended questions with many options is higher than for close-ended questions with fewer answer options on a mobile phone. Research of De Bruijne & Wijnant (2014) has shown that the completion time for an 11-point answer scale was significantly higher than a 5-point or 7-point answer scale on mobile phones. This could also be explained by the fact that

the 5-point answer scale was visible for 99%, the 7-point answer scale for 94% and the 11-point answer scale only for 59%. According to Couper and Peterson (2017), the need to scroll on a mobile phone leads to a higher completion time. Moreover, because more answer options are off screen, the tendency of respondents to choose the visible answer options is especially a problem when the survey is completed on mobile phones (De Bruijne & Wijnant, 2013; Mavletova, 2013; Stapleton, 2013).

Personal Characteristics

Younger people are on average faster in completing a web survey than older people. An explanation could be that the working memory capacity of older people is reduced, which makes the web survey more difficult (Yan & Tourangeau, 2008). However, although younger people complete a web survey faster, the break-off rate is higher for younger people than for older people, possibly due to motivation issues (Peytchev, 2009). Furthermore, younger people use mobile devices, in particular mobile phones, more than older people (De Bruijne & Wijnant, 2013).

Research shows that the response rate of women in online surveys is higher than the response rate of men (Smith, 2008). However, the break-off rate of women is higher than the break-off rate of men (Steinbrecher, Roßmann & Blumenstiel, 2015). Research has shown that men use a smartphone more than women to complete an online survey (De Bruijne & Wijnant, 2013).

In general, the time to complete a web survey is higher for respondents who did not complete high school than for respondents who did. The break-off rate of lower-educated respondents is higher than of higher-educated respondents (Peytchev, 2009; Yan & Tourangeau, 2008). In addition, higher-educated people use mobile phones more often than low-educated people (De Bruijne & Wijnant, 2013).

Innovative Ways to Conduct Surveys

Symon, Cassel and Dickson (2000) argue that there should be more alternative and innovative research methods. Online surveys are often seen by respondents as boring, which leads to reluctance to complete the survey (Dolnicar, Grün & Yanamandram, 2013). For this reason, researchers should look into new ways to present a survey (Dolnicar et al., 2013; Symon et al., 2000). One innovative research method is the gamification of online surveys. Gamification is the use of game design elements in non-game contexts (Harms, Seitz, Wimmer, Kappel & Grechenig, 2015). Gamification leads to more motivation of the respondents, a better user experience and positive feedback of the respondents (Harms et al., 2015). It is a way to make an online survey more interactive and dynamic. This can lead to a higher response rate and lower break-off rate of online surveys (Dolnicar et al., 2013). However, gamification of a survey takes effort (Seaborn & Fels, 2014).

Another way to make an online survey more interactive is a chatbot survey. Kim, Lee and Gweon (2019) used a text-based chatbot in their research to investigate the effect of a conversational element in an online survey. The researchers compared a chatbot survey with a regular web survey. The conversational style of the chatbot survey increased the differentiation in the responses of the respondents, leading to a higher quality of the response data. Furthermore, the respondents who completed the chatbot survey evaluated the survey more positively than respondents who completed the regular web survey (Kim, Lee & Gweon, 2019). The researchers suggest that the conversational style should be casual (Kim, Lee & Gweon, 2019). The research of Kim, Lee and Gweon (2019) did not focus on mixed-device surveys. As mentioned before, mobile phones mostly are used to communicate online by using short messaging apps, like Whatsapp (O'Hara et. al., 2014). A research messenger could be a way to make mixed-device surveys more interactive with a casual conversational style, since a research messenger is similar to short messages that are sent via Whatsapp. No previous research has been conducted to investigate the use of messenger type surveys in mixed-device research.

Methods

Respondents

Respondents could participate by completing an online survey, which was distributed among Amazon Mechanical Turk panel members in the United States of America. Amazon Mechanical Turk is a crowdsourcing marketplace that makes outsourcing of processes and jobs to a distributed workforce, which can perform these tasks virtually, easier. Participation in this research was possible from June to August in 2018. There were 2078 respondents in this research. However, 201 respondents did not complete a single question. In addition, 149 respondents did not complete the survey. The remaining 1728 respondents form the base of our analytic sample.

Survey

The respondents could self-select the device (computer, tablet or mobile phone) to complete the survey. At the beginning of the survey the respondents were randomly assigned to either a regular responsive survey design or a research messenger survey design. Appendix 1 shows images of both survey methods. In addition, respondents were randomly assigned to one of three response option conditions: a condition with open-ended questions; close-ended questions with few answer options; or close-ended questions with many answer options.

The survey consisted of four modules. Module A was about media use, model B about most important issues in the country, module C about politics and module D about sports. The order of the modules was randomized to avoid order effects. The order of the questions within the modules was not randomized. However, in this research only one question of each module is used. Therefore, the question order does not affect results. The respondents were assigned to the same response option condition and the same design in every module. After the four modules, the respondents had to answer questions about their background and their opinion about the survey (evaluation questions).

Analyses

The main goal of this research was to determine the best way to present mixed-device surveys. In order to do this, we investigated if there are differences between three different types of answer scales, two survey methods and the different devices used to complete the survey. Since there were only 104 respondents who completed the survey on a tablet, we decided not to treat tablets as a separate group. Tablets are sometimes grouped with mobile phones, because they are both mobile devices (De Bruijne & Wijnant, 2014). Some researchers group tablets together with computers because tablets are more similar to computers to complete a survey on, for instance both tablets and computers have large screen sizes (Couper & Peterson, 2017). We decided to group tablets with computers, because both devices are in general not used for Whatsapp type of messaging. However, we checked if the results are different when we group tablets with mobile phones. We added the personal characteristics age, gender and education to all analyses as control variables. Despite the low theoretical evidence for the effects of these personal characteristics, we checked if these variables affect the results of this research. To analyze the data we used IBM SPSS Statistics, 26.0.0.

Number of Completes

First, we conducted a simple binary logistic regression analysis to investigate if there is a difference in the proportion of completes between the types of answer scale, the survey methods and the devices used. We added age, gender and education to the analysis as control variables. The analysis has been done to investigate the break-off rate of the respondents. A survey is complete if all the questions of the survey were completed. We used the data of the 1877 respondents who started the survey. However, for fifteen respondents the device used to complete the survey could not be determined, so we did not include these respondents in the analysis. So, we used the data of 1862 respondents. We conjecture that the number of completes of respondents who completed the regular survey on a mobile phone is lower

than the number of completes of respondents who completed the research messenger survey on a mobile phone and the number of completes of respondents who completed the survey on a computer. Furthermore, we conjecture that the number of completes of respondents with open-ended questions is lower.

Substantive Answers

We choose one question per module to investigate if answers differ between the types of answer scale, survey methods, devices, and personal characteristics. We used the first question of the modules media use and sports. From the module about the most important issues in the country we used the only question which had the answer option "other, please write" so that all respondents could give the same answer despite the different types of answer scale. The module about politics did not contain an experiment with answer scales, therefore we did not use a question of this module. The questions we used are in Appendix 1.

The question about media use was "On a typical day, about how much time do you spend watching, reading or listening to news about politics and current affairs?". Respondents in the short scale got five answer options, respondents in the long scale eight. Respondents in the open format had to give their answer in hours and minutes. We dichotomized answer options; values lower than 75 minutes were coded as 0 and values of 75 minutes and higher were recoded as 1 (75 minutes was about the median time). Don't know answers were treated as non-substantive answer options. We dichotomized the answer options to make the answers comparable between the different conditions. The closed-ended questions have nominal time categories as answer options, not single numerical values. These time categories differ between the short scale and long scale, because the long scale has more answer options. The difference in mean could be caused by these different answer scales, therefore we dichotomized the answer options. The question about the most important issues in the country was "Which people or organizations you think have the most influence on the actions of the American government?". The close-ended questions had nine or twelve answer options for the short and long scale, respectively. The question about sports was "What sport or physical activity do you take part in most frequently?". Respondents in the short scale received thirteen answer options and respondents in the long scale twenty. In the analyses, the answers on the questions about the most important issues and sports were adjusted to the closed answer scale with many options. The answers of respondents with the short scale that answered "other" and the answers of respondents with the open answer scale were recoded manually to the long list of the closed answer scale with many options. After that, the questions were dichotomized; answer options that were only options in the long list received the value 1.

We perform simple binary logistic regressions with answer scales, survey methods, devices used, and personal characteristics in order to investigate if respondents

in different conditions give different answers. Interaction effects between answer scale, survey method, device, gender, age and education were also investigated, by adding the interaction terms to the regressions. We used the data of the 1728 respondents who completed the survey. However, of fifteen respondents it could not be determined which device was used to complete the survey. Furthermore, seven respondents choose the answer option "Other" or "Would rather not say" on the question "What is your gender?" and nineteen respondents did not fill in their age or answered with an invalid number. These respondents are not included in the analyses, so we used the data of the remaining 1687 respondents. We conjecture that there is a difference in answers on the questions between the different answer scale conditions. We also conjecture that respondents who completed the survey on a computer choose more often an answer option that is only in the long list than respondents who completed the survey on a mobile phone.

Completion Time

We use simple multiple regression to investigate if there is a difference in completion time between the types of answer scale, the survey methods, the devices used and age, gender and education. We also investigated interaction effects between answer scale, survey method, device, gender, age and education, by adding the interaction terms to the regression. The completion time is the time it took the respondents to complete the survey, so the time between the start and the end of the survey, and it is measured in seconds. We used the log of the completion time, because the distribution of the completion time is right-skewed. The data of the 1687 respondents of whom we had all the data was used. We conjecture that respondents with the research messenger have a higher completion time than respondents with the regular survey (due to the conversational element). Additionally, we conjecture that respondents with open-ended questions have a higher completion time than respondents with close-ended questions. Furthermore, we conjecture that respondents who completed the regular survey on a mobile phone have a higher completion time than respondents who completed the regular survey on a computer.

Evaluation Questions

Finally, we use simple multiple regression analyses to investigate if the answers on three evaluation questions differ between the types of answer sale, survey methods, the devices used and age, gender and education. Interaction effects between answer scale, survey method, device, gender, age and education were also investigated, by adding the interaction terms to the regressions. We used the data of the 1687 respondents of whom all the variables are known. The questions were: "Was it difficult to answer the questions?", "Did you enjoy answering the questions?" and "Was the subject interesting?". Answers on the evaluation questions were investigated in

order to determine the preferences of respondents with regard to presentation of mixed-device surveys. We conjecture that respondents who completed the research messenger evaluated the survey more positively than respondents who completed the regular survey, in particular respondents who completed the survey on a mobile phone. In addition, we conjecture that respondents in the open answer scale condition evaluated the survey more negatively than respondents in a closed answer scale condition.

Results

Descriptives

Table 1 shows the number of respondents per device, survey method, type of answer scale, gender and education. For fifteen respondents the device used to complete the survey could not be determined. Seven respondents choose the answer option "Other" or "Would rather not say" on the question "What is your gender?". Moreover, table 1 shows the minimum and maximum age of the respondents, the mean age, the mean completion time and the recoded binary variables from the different modules.

Table 1 Descriptive statistics

	n	%
Mobile phone (0)	538	31.4
PC (1)	1071	62.5
Tablet (1)	104	6.1
Total	1713	
Research messenger (RM) (0)	871	50.4
Regular survey (1)	857	49.6
Total	1728	
Open	580	33.6
Closed with few options	574	33.2
Closed with many options	574	33.2
Total	1728	
Female (0)	1157	67.2
Male (1)	564	32.8
Total	1721	

	n	%
Less than high school/ high school graduate (0)	1015	58.7
Some college/ college graduate (1)	713	41.3
Total	1728	
	min	max
Age in Years	18	84
	M	SD
Age in Years	34.87	10.119
Completion time in seconds	764.39	451.853
%	0	1
Question Media use *1:75 minutes or more*	51.6	48.4
Question Important issue *1: answers only in the long list*	82.8	17.2
Question Sports *1: answers only in the long list*	91.1	8.9

Completes

A binary logistic regression analysis is conducted to predict the proportion of completes. Table 2 shows that device, survey method, type of answer scale, age, gender and education do not significantly predict if the respondent completed the survey. The regression model is also not significant. However, as expected a lower proportion of respondents in the open answer condition completed the survey compared to respondents in a closed-ended condition. Furthermore, a higher proportion of respondents in the regular survey condition than in the research messenger condition completed the survey.

Table 2 Results binary logistic regression analysis predicting completes

	B	Exp(B)
Device (0: mobile phone, 1: tablet/PC)	.038	1.038
Survey method (0: RM, 1: regular survey)	-.495	.609
Open	-.820	.440
Closed few	-.350	.705
Closed many (ref.)		
Age	-.004	.996
Gender (0: female, 1: male)	.285	1.330
Education (0: ≤ high school graduate, 1: ≥ college)	-.437	.646
Nagelkerke R^2	.026	
n	1862	

Substantive Answers

We conducted binary logistic regression analyses to predict the answers on the questions about media use, important issues and sports from device used, survey method, answer scale, age, gender and education. Table 3 shows that respondents that completed the survey on a mobile phone significantly chose more often the response options in the long list on the question about most important issues. However, on the other questions respondents that completed the survey on a computer or tablet chose more often an answer option of the long list. The survey method did not have a significant effect on responses. Respondents with the short answer scale gave significantly different answers compared to the respondents with the long answer scale. In addition, in two out of three questions (media use and sports), responses from respondents in the open condition were significantly different than respondents in the long answer scale condition. There were two significant interaction effects, suggesting that respondents with high education that completed the survey on their computer/tablet reported less time in the media question; while men in the open format also reported to spend less time on media use. The model predicts 5.1% of the answers on the question about media use, 24.1% of the answers on the question about important issues and 11.1% of the answers on the question about sports.

Table 3 Results binary logistic regression analyses predicting answers on survey questions

	Media use		Important Issues		Sports	
	B	Exp(B)	B	Exp(B)	B	Exp(B)
Device (0: mobile phone, 1: tablet/PC)	.173	1.189	-.429	.651*	.209	1.233
Survey method (0: RM, 1: regular survey)	-.005	.995	.093	1.097	.224	1.252
Open	.380	1.462*	-.049	.952	-.513	.598*
Closed few	.646	1.908**	-4.595[1]	.010**	-2.537	.079**
Closed many (ref.)						
Age	.020	1.020**	.001	1.001	.001	1.001
Gender (0: female, 1: male)	.400	1.492*	-.530	.588**	.162	1.176
Education (0: ≤ high school graduate, 1: ≥ college)	.203	1.225	-.075	.928	-.044	.957
Education*device	-.498	.608*				
Gender*open	-.513	.599*				
Nagelkerke R^2	.051**		.241**		.111**	
n	1646		1687		1687	

*p< .05, **p< .001
1: the sample size is small.

Note: all the other interaction effects are not significant and were therefore not included in this model. The first question has fewer cases because of the DK option that is omitted from the analysis.

Completion Time

A multiple regression analysis is conducted to predict completion time from device used, survey method, type of answer scale, and personal characteristics. Table 4 shows that the time to complete the survey on a mobile phone was shorter than on a computer or tablet. The research messenger took significantly longer to complete than the regular survey method. Furthermore, the time to complete the survey with a closed answer scale with few options was significant higher than the time to complete a survey with another answer scale condition. There was no significant difference in completion time between the open answer scale and the closed answer scale with many options. Older respondents had a significantly higher completion

time than younger respondents. Women had a significant higher completion time than men. There was no difference in completion time between different levels of education. Although older respondents and respondents on a computer/tablet took longer to complete the survey, the interaction effect shows that older respondents on a computer/tablet took less time to complete the survey. Note that 6% of the variance in completion time can be predicted by the regression model.

Table 4 Results multiple regression analysis predicting the log completion time

	Beta
Device (0: mobile phone, 1: tablet/PC)	.330**
Survey method (0: RM, 1: regular survey)	-.114**
Open	.022
Closed few	.137**
Closed many (ref.)	
Age	.546**
Gender (0: female, 1: male)	-.073*
Education (0: ≤ high school graduate, 1: ≥ college)	.042
Age*device	-.472**
R^2	.086
F	19.666**
n	1686

* p< .05, ** p< .001

Note: One outlier is removed. Other interaction effects are not significant and were therefore not included in this model.

Evaluation Questions

To predict the three evaluation questions from device used, survey method, type of answer scale and personal characteristics, we conducted multiple regression analyses. Table 5 shows that respondents who completed the survey on a computer or tablet answered the evaluation questions significantly more negatively than the

respondents who completed the survey on a mobile phone. There is no significant difference in answers on the evaluation question between the survey methods nor types of answer scale. Older respondents enjoyed the survey significantly more and evaluated the survey as significantly more interesting than younger respondents. Men found the survey significantly more difficult than women. There is a significant interaction effect of gender and education on the difficulty of the survey (men with a high education found the survey more difficult) and a significant interaction effect of gender and survey method on how interesting the respondents evaluated the survey (men that received the regular survey design found the survey less interesting).

Table 5 Results multiple regression analyses predicting answers on evaluation questions

	Difficulty	Enjoyment	Interesting
	Beta	Beta	Beta
Device (0: mobile phone, 1: tablet/PC)	.043	-.066*	-.050*
Survey method (0: RM, 1: regular survey)	-.019	-.009	.024
Open	-.001	-.024	-.042
Closed few	-.023	-.025	-.000
Closed many (ref.)			
Age	-.012	.107**	.102**
Gender (0: female, 1: male)	.064*	-.001	.029
Education (0: ≤ high school graduate, 1: ≥ college)	-.055	.062*	.044
Gender*education	.084*		
Gender*survey			-.084*
R^2	.015	.015	.014
F	4.228**	4.663**	4.080**
n	1687	1687	1687

* p< .05, ** p<.001

Note: Other interaction effects are not significant and were therefore not included in this model.

Discussion and Conclusion

The goal of this research was to determine the best way to present mixed-device surveys. In order to do this, we investigated the differences between three different types of answer scales, two survey methods and the devices used.

Respondents with the close-ended answer scale with few options choose less often an answer option in the long list (e.g. by choosing other please specify) than respondents in other answer scale conditions. The completion time of the respondents who completed the survey with the close-ended answer scale with few options is also longer. An explanation for this could be that the true answer of the respondents is not in the list with few options. Previous research has shown that respondents tend to choose one of the options of the list even if it is not their true answer (Chung, Boyer & Han, 2010; Reja et al., 2003). Therefore, the respondents have to think longer about their answers and eventually choose an answer that is one of the options. Especially on the question about the most important issue, a very small number of respondents with the close-ended answer scale with few options have chosen an answer option that was only in the list of the close-ended answer scale with many options. This question does not have one possible answer, because respondents could think more issues are important. Therefore, it is likely that the respondents tend to choose an answer that is available rather than "other, please specify". For the question about sports the answers of respondents with the open-ended answer scale are also different than the answers of respondents with the close-ended answer scale with many options. Respondents that choose an answer option that was only in the longer list is in general low. This could be explained by the fact that the sports that were only in the longer list are in general not popular sports. For example, bowling was popular in America in 1960 but that popularity has declined (McIntosh, 2011). The difference in answers on the media question between the respondents in different answer scale conditions might have been caused by the difference in the scales instead of only by the difference in the presentation of the scales. As shown in appendix 1, the answer options are different, for example the first category of the closed-question with few options is "less than an hour" while the first category of the closed-question with many options is "less than half an hour". The answer options can be suggestive and serve as anchors for the respondents (Desai & Reimers, 2019). Therefore, the different scales with different categories could explain the difference in answers between respondents with few answer options and respondents with many answer options.

The time to complete the research messenger survey is longer than the time to complete the regular survey. The conversational element in the research messenger survey takes more time, because the respondents have to wait for the next question. However, the respondents who completed the research messenger survey did not answer the evaluation questions of the survey more negative. The break-off rate of

respondents who completed the research messenger survey is also not higher. So, the respondents with the research messenger survey did not seem to mind that it took longer to complete. However, the respondents did not evaluate the research messenger more positive compared with the regular survey. Which indicates that a research messenger survey might not be a better way to present mixed-device surveys.

In contrast to other studies and our assumptions, the respondents who completed the survey on a computer or tablet evaluated the survey more negative than the respondents who completed the survey on a mobile phone. However, the effects are small and did not depend on survey method.

In general, respondents can self-select the device to complete an online survey on. Therefore, in this research the respondents also could self-select the device. However, we expect that the effect of mobile phones would have been greater if respondents were assigned to a device, because then there were also respondents in the mobile phone condition who are less experienced with using a mobile phone. Furthermore, only single items were used in this research. We expect that the effects would be greater if we analyzed rating scales that encompass multiple items, since the effect on multiple items would be measured instead of the effect on a single item.

The completion time of older respondents was higher than of younger respondents. Despite the higher completion time, older respondents found the survey more enjoyable and interesting. Also, women had a significant higher completion time than men. However, men evaluated the survey as more difficult than women.

The survey was distributed among Amazon Mechanical Turk panel members in the United States of America. These panel members received a compensation for completing the survey. This could be a reason for the low break-off rate of the survey. Furthermore, the panel members might evaluate the research messenger survey more negatively than respondents who are not in a panel, because they are used to completing regular online surveys. Moreover, since the panel members are trained in completing surveys, they might have less problems with open-ended questions or closed-ended questions with many answer options, such as a longer completion time. Although the respondents are not representative of the population, the sample was heterogeneous. The sample is heterogenous, though not representative of the population. Especially female respondents are overrepresented. Future research using a probability-based sample should be used to replicate our results and test robustness.

In conclusion, based on our results we recommend to use a close-ended answer scale with many options or an open-ended answer scale since a closed-ended answer scale with few options results in other frequencies hence outcomes. The research messenger survey did not seem to be a better method to present mixed-device surveys than a regular survey. Further research is necessary to investigate

how to present mixed-device surveys in order to increase participation and data quality in mixed-device surveys.

References

Ahad, A. D., & Lim, S. M. A. (2014). Convenience or Nuisance?: The 'WhatsApp' dilemma. *Procedia–Social Behavioral Sciences, 155*, 189-196. DOI:10.1016/j.sbspro.2014.10.278

Antoun, C., Katz, J., Argueta, J., & Wang, L. (2018). Design heuristics for effective smartphone questionnaires. *Social Science Computer Review, 36*(5), 557-574. https://doi.org/10.1177/0894439317727072

Bowling, A. (2005). Mode of questionnaire administration can have serious effects on data quality. *Journal of Public Health, 27*(3), 281-291. https://doi.org/10.1093/pubmed/fdi031

Brant, J. M., Haas-Haseman, M. L., Wei, S. H., Wickham, R., & Ponto, J. (2015). Understanding and evaluating survey research. *Journal of the Advanced Practitioner in Oncology, 6*(2), 168–71. DOI:-

Chung, C., Boyer, T., & Han, S. (2010). How many choice sets and alternatives are optimal? Consistency in choice experiments. *Agribusiness, 27*(1), 114-125. https://doi-org.proxy.library.uu.nl/10.1002/agr.20252_

Couper, M. P., & Miller, P. V. (2008). Web survey methods. *Public Opinion Quarterly, 72*(5), 831-835. https://doi.org/10.1093/poq/nfn066

Couper, M. P., & Peterson, G. J. (2017). Why do web surveys take longer on smartphones? Social Science Computer Review, 35(3), 357-377. https://doi.org/10.1177/0894439316629932

Cunningham, J. A., Neighbors, C., Bertholet, N., & Hendershot, C. S. (2013). Use of mobile devices to answer online surveys: Implications for research. *BMC Research Notes, 6*, 258. https://doi.org/10.1186/1756-0500-6-258

De Bruijne, M., & 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

De Bruijne, M., & Wijnant, A. (2014). Improving response rates and questionnaire design for mobile web surveys. Public Opinion Quarterly, 78(4), 951-962. https://doi.org/10.1093/poq/nfu046

De Leeuw, E. D. (2008). Self-administered questionnaires and standardized interviews. *Handbook of social research methods*, 313-327. DOI: -

Desai, S. C., & Reimers, S. (2019). Comparing the use of open and closed questions for Web-based measures of the continued-influence effect. *Behavior research methods*, *51*(3), 1426-1440. https://doi.org/10.3758/s13428-018-1066-z

Dolnicar, S., Grün, B., & Yanamandram, V. (2013). Dynamic, interactive survey questions can increase survey data quality. *Journal of Travel & Tourism Marketing, 30*(7), 690-699. https://doi.org/10.1080/10548408.2013.827546

Harb, E., Kapellari, P., Luong, S., & Spot, N. (2011). Responsive web design. *Version of*, *6*. Retrieved from http://courses.iicm.tugraz.at/iaweb/surveys/ws2011/g3-survey-resp-web-design.pdf

Harms, J., Seitz, D., Wimmer, C., Kappel, K., & Grechenig, T. (2015). Low-cost gamification of online surveys: Improving the user experience through achievement badges. *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in* Play*, 15*, 109-113. http://dx.doi.org/10.1145/2793107.2793146

Harris, L. R., & Brown, G. T. L. (2010). Mixing interview and questionnaire methods: Practical problems in aligning data. *Practical Assesment, Research, and Evaluation, 15*(1), 1-14. https://doi.org/10.7275/959j-ky83

Heiervang, E., & Goodman, R. (2011). Advantages and limitations of web-based surveys: Evidence from a child mental health survey. *Social Psychiat Epidemiol, 46*, 69-76. https://doi.org/10.1007/s00127-009-0171-9

Hussain, A., & Mkpojiogu, E. O. (2015). The effect of responsive web design on the user experience with laptop and smartphone devices. *Jurnal Teknologi*, *77*(4), 41-47. DOI:-

IBM SPSS Statistics (26.0.0) [Computer software]. United States: IBM Company.

Kim, S., Lee, J., & Gweon, G. (2019). Comparing data from chatbot and web surveys: Effects of platform and conversational style on survey response quality. *Proceedings of the 2019 CHI conference on human factors in computing systems, 86,* 1-12. DOI:-

Maslovskaya, O., Smith, P., & Durrant, G. (2020). Do respondents using smartphones produce lower quality data? Evidence from the UK Understanding Society mixed-device survey. *National Centre of for Research Methods Working Paper, 1* (20), 2-32. DOI:-

Mavletova, A. (2013). Data quality in PC and mobile web surveys. *Social Science Computer Review, 31*(6), 725-743. https://doi.org/10.1177/0894439313485201

McIntosh, P. (2011). Bowling: Entertainment for All Ages. *English Teaching Forum, 49*(4), 37-45. DOI:-

Millar, M., & Dillman, D. A. (2012). Encouraging survey response via smartphones. *Survey Practice*, *5*(3), 3095. https://doi.org/10.29115/SP-2012-0018

O'Hara, K., Massimi, M., Harper, R., Rubens, S., & Morris, J. (2014). Everyday dwelling with WhatsApp. *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing, 14*, 1131-1143. https://doi.org/10.1145/2531602.2531679

Peytchev, A. (2009). Survey breakoff. *Public Opinion Quarterly*, *73*(1), 74-97.

Preston, C. C., & Colman, A. M. (2000). Optimal number of response categories in rating scales: reliability, validity, discriminating power, and respondent preferences. *Acta psychologica*, *104*(1), 1-15. https://doi.org/10.1016/S0001-6918(99)00050-5

Reja, U., Manfreda, K. L., Hlebec, V., & Vehovar, V. (2003). Open-ended vs. Close-ended Questions in Web Questionnaires. *Developments in applied statistics*, *19*(1), 159-177. DOI: -

Schlosser, S., & Mays, A. (2018). Mobile and dirty: Does using mobile devices affect the data quality and the response process of online surveys? *Social Science Computer Review, 36*(2), 212-230. https://doi.org/10.1177/0894439317698437

Seaborn, K., & Fels, D. I. (2014). Gamification in theory and action: A survey. *International Journal of human-computer studies*, *74*, 14-31. https://doi.org/10.1016/j.ijhcs.2014.09.006

Smith, G. (2008). Does gender influence online survey participation?: A record-linkage analysis of university faculty online survey response behavior. *ERIC Document Reproduction Service No. ED 501717,* 2-21.

Solomon, D. J. (2000). Conducting web-based surveys. *Practical Assessment, Research, and Evaluation, 7(*19), 1-4. https://doi.org/10.7275/404h-z428

Stapleton, C. E. (2013). The smartphone way to collect survey data. *Survey Practice, 6*(2), 1-7. DOI:-

Steinbrecher, M., Roßmann, J., & Blumenstiel, J. E. (2015). Why do respondents break off web surveys and does it matter? Results from four follow-up surveys. *International Journal of Public Opinion Research*, *27*(2), 289-302.

Symon, G., Cassell, C., & Dickson, R. (2000). Expanding our research and practice through innovative research methods. *European Journal of Work and Organizational Psychology, 9*(4), 457-462. https://doi.org/10.1080/13594320050203076

Toepoel, V. & Lugtig, P. J. (2015). Online surveys are mixed-device surveys. Issues associated with the use of different (mobile) devices in web surveys. methods, data, analysis, 9(2), 155-162. https://doi.org/10.12758/mda.2015.009

Tourangeau, R., Maitland, A., Rivero, G., Sun, H., Williams, D., & Yan, T. (2017). Web surveys by smartphone and tablets: Effects on survey responses. *Public Opinion Quarterly*, *81*(4), 896-929. https://doi.org/10.1093/poq/nfx035

Wright, K. B. (2005). Researching Internet-based populations: Advantages and disadvantages of online survey research, online questionnaire authoring software packages, and web survey services. *Journal of computer-mediated communication*, *10*(3), JCMC1034.

Yan, T., & Tourangeau, R. (2008). Fast times and easy questions: The effects of age, experience and question complexity on web survey response times. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *22*(1), 51-68.

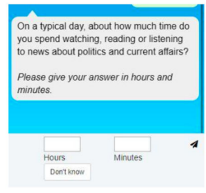
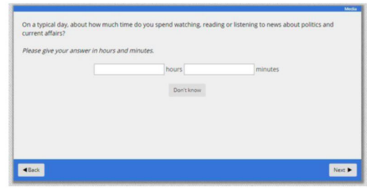
Zhang, X., Kuchinke, L., Woud, M. L., Velten, J., & Margraf, J. (2017). Survey method matters: Online/offline questionnaires and face-to-face or telephone interviews differ. *Computers in Human Behavior*, *71*, 172-180. https://doi.org/10.1016/j.chb.2017.02.006

Appendix 1

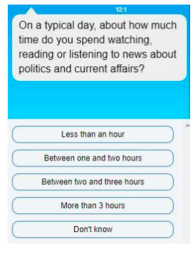
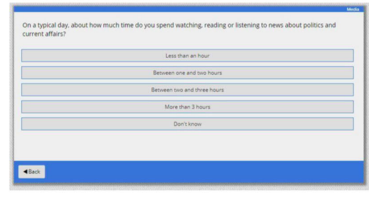
Screenshot of the answer scales and survey methods of the questions used. Video's are available upon request

Question about media use.

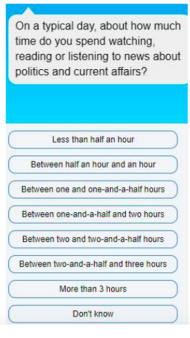
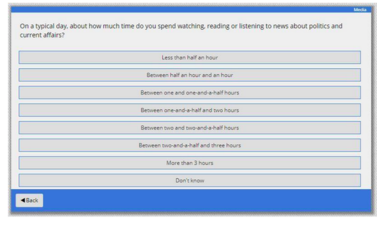
Section A: Media use: A1A

Questionnaire	RM version	Traditional version
[1A]. On a typical day, about how much time do you spend watching, reading or listening to news about politics and current affairs? Please give your answer in hours and minutes. **(source: European Social Survey (round 8))** ___ (open answer) Don't know		

Section A: Media use: A1B

Questionnaire	RM version	Traditional version
[1b]. On a typical day, about how much time do you spend watching, reading or listening to news about politics and current affairs2? **(source: European Social Survey (round 8))** Less than an hour Between one and two hours Between two and three hours More than 3 hours Don't know		

Section A: Media use: A1C

Questionnaire	RM version	Traditional version
[1c]. On a typical day, about how much time do you spend watching, reading or listening to news about politics and current affairs2? **(source: European Social Survey (round 8))** Less than half an hour Between half an hour and an hours Between one and one-and-a-half hours Between one-and-a half and two hours Between two and two and a half hours Between two and-a half and three hours More than 3 hours Don't know		

Question about most important issue.

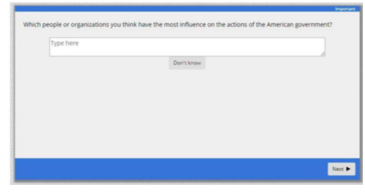
Section B: Most important problem:B_A4

Questionnaire

4. **(GSS)** Which people or organizations you think have the most influence on the actions of the American government?
___ **(open answer)**

RM version

Traditional version



Section B: Most important problem:B_B7

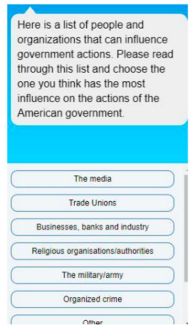
Questionnaire

7. **(GSS**)Here is a list of people and organizations that can influence government actions.
Please read through this list and choose the one you think has the most influence on the actions of the American government

The media
Trade Unions
Businesses, banks and industry
Religious organisations/authorities
The military/army
Organized crime
Other, ___
I can't choose
Don't know

[if "can't choose or "don't know" -> go to module c]

RM version



Traditional version

Section B: Most important problem:B_C3

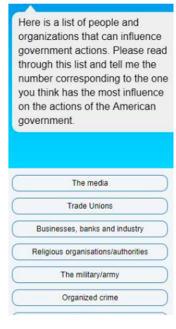
Questionnaire

3. (GSS) Here is a list of people and organizations that can influence government actions.
Please read through this list and tell me the number corresponding to the one you think has the most influence on the actions of the American government

The media
Trade Unions
Businesses, banks and industry
Religious organisations/authorities
The military/army
Organized crime
Schools/education
Ordinary People
Rich people
Other, ___
I can't choose
Don't know

[if "can't choose or "i don't know" -> go to module c]

RM version



Traditional version

Question about sports.

Section D: Leisure Time and Sports: D_1A

Questionnaire

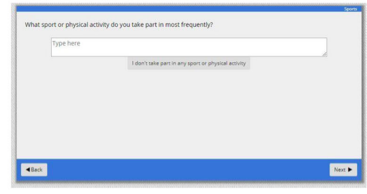
What sport or physical activity do you take part in most frequently? ((If you do not take part in any sport or physical activity, please tick the box provided below.)) *<OPEN-ENDED>*

Most frequent sport or physical activity
_____ (open answer)

I do not take part in any sport or physical activity

RM version

Traditional version



Section D: Leisure Time and Sports: D_1B

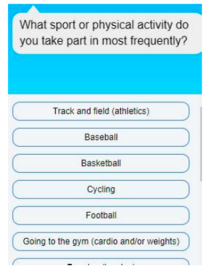
Questionnaire

[1b] What sport or physical activity do you take part in most frequently?

Track and field (athletics)
Baseball
Basketball
Cycling
Football
Going to the gym (cardio and/or weights)
Running (jogging)
Swimming for fitness
Treadmill or Gym equipment at home
Walking for fitness
Other, please write ___ **(open answer)**
Not sure

Do not know

RM version



Traditional version

Section D: Leisure Time and Sports: D_1C

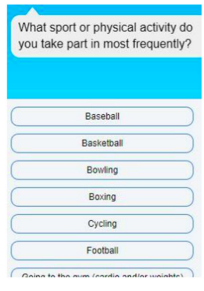
Questionnaire

[1c] What sport or physical activity do you take part in most frequently?

Baseball
Basketball
Bowling
Boxing
Cycling
Football
Going to the gym (cardio and/or weights)
Golf
Horse riding
Ice Hockey
Running (jogging)
Soccer
Swimming for fitness
Tennis
Track and field (athletics)
Treadmill or Gym equipment at home
Walking for fitness
Yoga/Pilates
Other, please write
Not sure

Do not know

RM version



Traditional version