Survey Research Methods (2016)
Vol. 10, No. 2, pp. 153-169
doi:10.18148/srm/2016.v10i2.6274

ISSN 1864-3361
http://www.surveymethods.org

Smartphones vs PCs: Does the Device Affect the Web Survey Experience and the Measurement Error for Sensitive Topics? A Replication of the Mavletova & Couper's 2013 Experiment

Daniele Toninelli
Department of Management, Economics and Quantitative Methods
University of Bergamo
Bergamo, Italy

Melanie Revilla
RECSM
Universitat Pompeu Fabra
Barcelona, Spain

More and more respondents use mobile devices to complete web surveys. These devices have different characteristics, if compared to PCs (e.g. smaller screen sizes and higher portability). These characteristics can affect the survey responses, mostly when a questionnaire includes sensitive questions. This topic was already studied by Mavletova and Couper (2013), through a two-wave experiment comparing PCs and mobile devices results for the same respondents in a Russian opt-in panel. We replicated this cross-over design, focusing on an opt-in panel for Spain, involving 1,800 panellists and comparing PCs and smartphones. Our results support most of Mavletova and Couper's (2013) findings (e.g. generally the used device does not significantly affect the reporting of sensitive information), confirming their robustness over the two studied countries. For other results (e.g. trust in data confidentiality), we found differences that can be justified by the diverse context/culture or by the quick changes that are still characterizing the mobile web survey participation.

Keywords: Web surveys; mobile participation; smartphones; survey optimization; sensitive questions; measurement error

1 Introduction

In the last years, a lot of effort has been made to study the use of mobile devices as tools for data collection (e.g., see Buskirk & C, 2012; de Bruijne & Wijnant, 2013, 2014; Toninelli, Pinter, & de Pedraza, 2015). This interest was stimulated by recent studies demonstrating that mobile devices are quickly spreading among the population of most of the countries. According to StatCounter StatCounter Global Stats (2015), the mobile Internet usage grew from 8.5% (September 2012) to 41.0% (September 2015). During the same period, the PCs (desktops and laptops, from here on) web accesses dropped from 91.5% to 59.0%. The speed of this phenomenon varies a lot by country, and in some countries the mobile devices owners are overtaking the PCs owners. Thus, recent results (e.g., Revilla & Ochoa, 2015) highlight that it became no-more-negligible. These facts caused the birth and the spread of the "unintended mobile participation" (Peterson, 2012; Wells, Bailey, & Link, 2013). This means that respondents attempt to participate in web surveys by means of mobile devices, even if surveys are not adapted for this kind of access. To better understand the phenomenon, Revilla, Toninelli, Ochoa, and Loewe (2016) studied the preferences of web panellists for different devices in seven countries. The authors found that 85.5% of the panellists usually prefer to participate in surveys by PCs. Nevertheless, if surveys are adapted for the mobile participation, the preference for smartphones considerably increases (from 5.5% to 31.0%) and the preference for PCs drops (from 85.5% to 38.8%). Moreover, the adaptation is a way to encourage the web survey participation itself.

These trends about the owned devices, the mobile web access, the survey participation and the preferences of respondents, all converge in underlining the significant and quick increase of the involvement in web surveys of the unintended mobile respondents, confirmed by several studies (in particular, de Bruijne & Wijnant, 2014).

Mobile devices have very different characteristics if compared to PCs (e.g. the reduced size of the screens, or the fact they allow the participation from any places or make more probable the presence of bystanders/strangers). The papers of Chae and Kim (2004) and of Sweeney and Crestani (2006) are some of the first works focused on these differences. Starting from these early years, several papers studied

Contact information: Daniele Toninelli, University of Bergamo, Department of Management, Economics and Quantitative Methods, Via dei Caniana 2, 24127 Bergamo, Italy (email: daniele.toninelli@unibg.it)

how the different characteristics of mobile devices can affect the survey responses and, more generally, the response process. Peytchev and Hill (2010), for example, analysed the effects of scrolling (caused by the reduced size of the screen of mobile devices) with different orientations of the scales, and in responding to open-ended questions.

This work was inspired by the paper of Mavletova and Couper (2013) which compares PC and mobile web participation when sensitive topics (e.g. alcohol consumption or deviant behaviours) are asked to respondents. In particular, the authors studied how social desirability affects the willingness to report sensitive information when a mobile device (feature phone or smartphone) is used rather than a PC. This was done considering several background and context variables: gender, presence of bystanders during the survey, place of participation, etc. This study "represents a first examination of possible social desirability biases of mobile web versus PC web surveys" (Mavletova & Couper, 2013, p. 202), but it also suffers from some limitations. The main one is that "the study is restricted to members of a volunteer online opt-in panel in Russia" (Mavletova & Couper, 2013, p. 200). Thus, the authors encourage further research in this specific area, highlighting the importance of replicating previous studies (Couper, 2014).

Therefore, in our paper we want to replicate the Mavletova and Couper (2013) research but in a different context, in order to test the robustness of their conclusions.

First, we study a different country: Spain, instead of Russia. The two countries are very dissimilar in terms of both Internet and smartphones penetration. According to Internet Live Stats (2015) the 2014 internet penetration (percentage of the total population with Internet) was 59.3% in Russia (+10% in comparison to 2013), whereas in Spain it was 74.4% (+3% in one year). The smartphones' usage is also different in the two countries: the percentage of people who use a smartphone for Russia was 45%, in 2014 and it is 61% in 2015 (+ 16 percentage points in one year), whereas for Spain it was 72% in 2014 and it is 80% in 2015 (+8 percentage points in one year). The average number of devices that one can use to access the web, per person, is similar: 2.4 in Russia, 2.9 in Spain (Consumer Barometer, 2014). Nevertheless, the percentage of people who access the internet at least as often via smartphones as via PCs is quite different: 32% Russia versus 64% Spain. In addition, the two countries are also very different if we consider the social desirability tendencies. For example, Steenkamp, de Jong, and Baumgartner (2009) analyze social desirability in both countries. They show that for moralistic response tendencies, Russia has a higher score than Spain, whereas for egoistic response tendencies, Spain has the highest observed score and Russia one of the lowest. The different cultural background of the two countries and the wider spread of smartphones observed in Spain could probably cause different tendencies in reporting sensitive information. Previous findings (e.g., the already cited Steenkamp et al., 2009) confirmed that there could be a different device effect on the willingness of reporting sensitive information, in these countries. We expect to observe a higher social desirability effect in Russia, where there could be a higher resistance in disclosing personal and sensitive information (e.g. deviant behaviors) and where people show a less frequent use of smartphones (and, thus, a lower familiarity with such devices, that could increase even more the social desirability bias). Thus, our study provides findings that can be complementary to the ones of Mavletova and Couper (2013) by completing the knowledge about the phenomenon of mobile web survey participation (i.e., taking into consideration a country quite different from Russia, Spain).

Second, our data were collected in 2015, so they are more recent and we can expect different results, because this phenomenon changed so quickly in the last years.

Third, our research has a slightly different focus. On the one hand, we compare smartphones and PCs results, instead of mobile devices (a more general category including smartphones and feature phones) and PCs; we are particularly interested in smartphone users because this device is the most common, within the panel studied in this paper (Revilla & Ochoa, 2015). On the other hand, we study if the used device can cause a bias in reporting sensitive information focusing on measurement error only (see Sect. 3). Finally, besides replicating what was done by Mavletova and Couper (2013), we also want to study the effect of the adaptation of the survey for the mobile participation (smartphone-optimized version of the questionnaire). This questionnaire's optimization aims at easing the survey participation through mobile devices by automatically adjusting the layout to the screen size. This topic was studied, for example, by McClain, Crawford, and Dugan (2012).

In order to investigate if there is a device/optimization effect, we implemented a two-wave survey experiment. For the first wave, we randomly assigned the selected panellists to a specific survey group, within the following: PC (*PC* group, from here on); smartphone-not-optimized (*SNO* group, from here on); smartphone-optimized (*SO* group, from here on). In the smartphone-optimized questionnaire, there is an automatic adaptation of the showed page to device's screen size. Moreover the size of buttons is enlarged and unnecessary elements are avoided. Thus, respondents do not need to zoom-in within the page, nor to scroll horizontally. This all allows to obtain an enhanced readability of the questionnaire and makes the answering process more friendly. For the second wave, we asked the same group of respondents to participate again in the same survey, but switching to another setting, in some cases. 1,608 panellists completed the survey in both waves.

The next section introduces our research background and our hypotheses. Section 3 presents our methodology and questionnaire. Section 4 explains the sensitive indices used in our analysis. The main findings are shown in Section 5. The last section summarises and discusses these main findings.

2 Research background and hypotheses

The first part of this section introduces the main previous literature findings about the comparison of PC and mobile web survey participation. This helps us in defining our objectives and hypotheses, which are introduced in the last part of the section.

2.1 Previous research comparing PCs and mobile devices

The comparison between a PC and a mobile participation has already been object of previous studies. Several works focused on understanding how mobile web users or mobile web survey respondents differ from a more general population (e.g., see Antoun, 2015b; Revilla & Ochoa, 2015). Others estimated the coverage error in surveys designed for computers (e.g., Mohorko, de Leeuw, & Hox, 2013) and for mobile devices (e.g., Fuchs & Busse, 2009).

Taking into consideration PCs and mobile devices survey experience, de Bruijne and Wijnant (2013) showed no significant differences in terms of the evaluation of questionnaire difficulty and of the interest and the enjoyment of respondents. Other authors (e.g., Andreadis, 2015; Mavletova & Couper, 2013) found longer response and questionnaire completion times for mobile devices, in comparison to PCs. Nevertheless, Toepoel and Lugtig (2014) demonstrated that the total response times with mobile devices are almost the same than with PCs, if one provides respondents with a mobile-friendly version of the questionnaire. Baker-Prewitt (2013), Bosnjak et al. (2013), Buskirk and Andrus (2014), and Villar, Callegaro, and Yang (2013) studied the break-off rates, finding higher rates for mobile web than for PC web surveys. Nevertheless, a version optimized for the mobile participation seems to help in lowering the break-off rates (Baker-Prewitt, 2013; Peterson, Mechling, LaFrance, Swinehart, & Ham, 2013; Stapleton, 2013).

The quality of data collected by means of mobile devices was also the object of several studies (Antoun, 2015a; de Bruijne & Wijnant, 2013; Mavletova & Couper, 2013; Revilla & Ochoa, 2015; Stapleton, 2013; Struminskaya, Weyandt, & Bosnjak, 2015; Wells et al., 2013, 2014). In some cases, the authors proposed best practices in order to enhance the results that can be obtained implementing mobile web surveys (e.g., Mavletova & Couper, 2015). Within this field, one of the main methodological questions pertains to the effect of the mobile web participation on the responses (Antoun, 2015a; Buskirk & Andrus, 2012; Guidry, 2012; McClain et al., 2012; Newell, Logan, Guo, Marks, & Shepperd, 2015; Peytchev & Hill, 2010).

2.2 What happens when sensitive questions are proposed?

One of the most interesting components of the device-effect is observed when reporting sensitive information. According to Tourangeau and Yan (2007, p. 859) “sensitive questions is a broad category that encompasses not only questions that trigger social desirability concerns but also those that are seen as intrusive by the respondents.” Moreover, Tourangeau, Groves, C., and Yan (2009) define a question “sensitive” if it implies highly undesirable or desirable answers, or if respondents have concerns about disclosing such kind of information to third parties. In answering about sensitive topics, respondents are subject to social desirability bias: they are likely to provide answers that present them in a better light, underreporting socially undesirable behaviours or over-reporting socially desirable ones (see Schaeffer, 2000; Tourangeau & Yan, 2007). Studying the use of sensitive questions in surveys, Tourangeau and Yan (2007) confirmed that the survey design and aspects like the self-administration or a private setting affect how respondents deal with sensitive questions, causing the misreporting of sensitive topics (Tourangeau, Groves, & Redline, 2010). Another study about these topics is the one of Sakshaug, Yan, and Tourangeau (2010). From this point of view, web and mobile web surveys maintain the same advantages of self-administered surveys: according to Kreuter, Presser, and Tourangeau (2008), the self-administered nature of web surveys helps in obtaining more honest answers to sensitive questions than interviewer-administered surveys. Studying the impact of the used device (PCs vs mobile devices) on the reporting of sensitive information, Mavletova and Couper (2013) found significant differences in two out of five sensitive indices: higher levels of alcohol consumption and income are reported by PC respondents. Thus, most of the time, it seems that the use of a particular device does not affect the reporting of sensitive information. Moreover, the authors found that the type of device affects the perceived privacy and that the bystanders' effect varied according to the content of the questions. Contrarily to this, the used device does not significantly affect other survey experience aspects: for example, there was “no evidence of differential satisficing by device, or by the context in which the survey was completed” (Mavletova & Couper, 2013, p. 200). Focusing on the same topic (the effect of the device on reporting sensitive information), Antoun (2015b) analysed data from the Longitudinal Internet Studies for the Social Sciences (LISS) probability panel, comparing the mobile-web and the PC-web participation. The author found no significant impact of completion device on response quality (in terms of both conscientious responding and disclosure of sensitive information), despite that “respondents in the mobile Web survey really were more mobile and more engaged with the other people and things around them compared to PC-Web respondents”.

Thus “smartphones appear to have no negative effects on response quality” if there is a careful design of mobile web surveys (i.e., avoiding certain question formats). Nevertheless, the study of Antoun (2015b) did not consider the effect of the optimization of the questionnaire for a mobile participation and did not include a mobile-optimized experimental condition. Moreover this study included people not owning mobile devices, providing them to these respondents.

2.3 Our paper’s goal and hypotheses

Our aim is to replicate the Mavletova and Couper (2013) experiment in a very different context (i.e. in a different country, using more recent data, and testing more experimental conditions), as highlighted in the last part of Section 1. We want to test hypotheses comparable to the ones of Mavletova and Couper (2013). In order to enhance the comparability between the two studies, the questions and the structure of the original experiment was kept unchanged, whenever possible. Moreover, our work is also based on the same methodology and we use data collected by a volunteer opt-in online panel as well; in addition, we are not providing our respondents with a smartphone, if they do not have one, contrarily to what done in Antoun (2015b). We aim at studying, following the criterion proposed by Mavletova and Couper (2013), a particular kind of measurement error, that we can define “relative measurement error”. This is understood as the average of the differences observed for all respondents between the two waves, rather than as the accuracy of responses.

Following Mavletova and Couper (2013), we expect the mobile participation to potentially take place in different contexts: public rather than private places, higher probability of having bystanders during the survey (in particular strangers), and so on. These differences in the survey context could lead to a higher social desirability bias, causing the misreporting of sensitive information and, consequently, higher measurement errors.

Our hypotheses can be classified into two different sets. The first set (H1 to H3) focuses specifically on the mobile web survey context. The goal is to check if the context is actually different, when participating in a survey by means of a smartphone, and if the main factors that can cause the social desirability bias are observed.

H1: using smartphones, it is *more common to participate in surveys in places different than home*;

H2: if a smartphone is used, there is a *higher probability of the presence of bystanders* (i.e. known or unknown third parties) during the participation;

H3: participating to web surveys through smartphones, it is *more common to fill the questionnaire in presence of strangers* (i.e. unknown third parties).

Once tested these hypotheses about the survey context, we can study the second set of hypotheses (H4 to H6). These are focused on the survey experience (H4) and, more specifically, on the consequences of the potential social desirability bias (i.e., on the measurement error, H5 and H6).

H4: the smartphone respondents *feel less comfortable* than the PC respondents, due to a perceived lack of privacy/trust in confidentiality;

H5: the higher probability of having bystanders for mobile web participants causes the *underreporting of sensitive information*.

The spread of the unintended mobile participation pushed the online fieldwork companies to develop questionnaires optimized for mobile devices, in order to ease their readability and the completion process. This adaptation and the obtained different layouts could also be a factor that potentially affects the comparability of collected data. Nevertheless, we expect that this factor does not significantly affect social desirability, which is our focus here.

H6: completing a *questionnaire optimized for smartphones rather than a not-optimized one does not significantly affect the reporting* of sensitive information.

3 Experiment Methodology

In order to test the previously listed hypotheses, we implemented a two-wave experiment. The web-based surveys were implemented in an opt-in online panel, Netquest (www.netquest.com). Netquest is active since 2001, covering some European (Spain and Portugal) and Latin American countries (e.g., Colombia, Mexico, Brazil). Our study is based on the panel for Spain. Further details about our experiment are provided in this section.

3.1 The experimental design

In order to maximize the comparability with the Mavletova and Couper paper, we planned our experiment using similar settings. The experiment was organized as a two-wave cross-over design. In the two waves the same respondents were involved. Following the suggestion of Mavletova and Couper (2013), we decided to keep as short as possible (one week) the lag between the two waves (in their study it was one month). This way, at the individual level, the potential differences over the two waves in providing the answers about sensitive behaviours should be minimized. This design allowed studying the relative bias in reporting sensitive information. We were able to evaluate the device effect, comparing answers provided by the same respondents but using two different kinds of devices: smartphones and PCs. Moreover, questions about the surveys experience allowed studying how much respondents felt comfortable during the participation, depending on the used device.

3.2 Data collection

The data were collected in 2015, from the 23rd of February to the 2nd of March (first wave) and from the 9th to the 18th

of March (second wave). In order to select the respondents within our target population (the panellists of Netquest), first, we screened out respondents with an age lower than 18. Then, we selected only panellists who have access to both PC and smartphone and who have used both of them to access to the Internet at least once, during the 30 days before the survey. Moreover, we selected only those who committed themselves to participate to both waves of the survey and to use the device we asked for. In order to encourage a second participation, a bigger incentive was guaranteed by Netquest for completing the second wave. Cross quotas for age and gender were used in order to guarantee in our sample a distribution similar to the one of the panel.

For the first wave we contacted 3,317 panellists: 2,720 of them (82.0%) opened the survey's introduction page. 581 respondents (21.4%) were screened out by the filter questions, because they had an age under 18, because they did not have Internet access by means of both devices, because they did not access the Internet in the previous 30 days or because they did not accept to commit themselves to participate twice to the same survey. After the preliminary filter questions, the remaining panellists were randomly assigned to a specific survey group: *PC*, *SNO* or *SO*. Moreover the devices used were automatically detected: if a respondent attempted to start the survey with a device different from the one requested and did not switch to the requested device, the respondent was screened out (296 cases; 10.9%). 1,843 respondents (55.6% of the contacted panellists) reached the first question and 1,800 (54.9%) completed the first wave survey. The 1,800 individuals that completed the survey were randomly assigned to three groups as follows: 554 units (34.5%) were assigned to the *PC* group, 536 (33.3%) to the *SNO* group and 518 (32.2%) to the *SO* group.

One week after the end of the first wave, the same 1,800 panellists were invited to participate in the second wave. They were randomly re-assigned again to one of the same categories (PC, SNO, SO). In this way, we obtained nine experimental groups: for three of them (control groups) the condition did not change, between the two waves, whereas for the other six (treatment groups) the respondents changed device or optimization settings. The control groups help in understanding if the detected differences are due to the device or to other disturbing factors (e.g. answering to the survey for the second time), whereas using data from the treatment groups we can detect if there is a device/optimization effect. A summary of the respondents' distribution over the nine experimental groups is shown in Table 1. A total of 1,608 panellists completed the second wave questionnaires (89.3% of participants from the first wave). This is the group of panellists that is analysed in the following.

3.3 The questionnaire

The questionnaire is very similar to the one used by Mavletova and Couper (2013) and was the same for both waves. It includes sets of questions about sensitive topics. Following the example of these authors, we used measures similar to the Marlowe-Crowne scale (Crowne & Marlowe, 1964) and the balanced inventory of desirable responding (Paulhus, 1984). The three versions of the online questionnaire (one for each survey group) are available at the following links:

1. *PC*: http://goo.gl/g9gAE4;
2. *SNO*: http://goo.gl/4c9d1C;
3. *SO*: http://goo.gl/5jF2vr.

4 Sensitive indices

We construct sensitive indices following Mavletova and Couper (2013), where possible; next, we explain how we computed each of these five sensitive indices.

1. The *attitude towards deviant practice* set proposes a list of 15 deviant behaviours (e.g. "lying in one's own interest"; see Appendix 1/Set #1 for details). The respondents have to express how much these behaviours can be justified, using a four-point scale. We used a similar scale as Mavletova and Couper (2013), but we changed the labels of some categories in order to obtain a more progressive evaluation of the attitude. Our scale is unipolar, with four increasing levels of acceptance of the behaviour, from "never can be justified" to "always can be justified". These options were recoded using a 0 to 3 scale (where: "never can be justified" = 0, "sometimes can be justified" = 1, "often can be justified" = 2, "always can be justified" = 3). Summing up the values of each respondent, we computed a score of socially undesirable answers. Then, we re-scaled this score into a 0 to 100 index, where 0 means that all behaviours can never be justified and 100 means that all behaviours can always be justified.

2. The *rate of deviant behaviour* set is made of 15 items (e.g. "Have you ever stolen something in a shop?"; see Appendix 1/Set #2) providing respondents with a "yes/no" answer option (1 = "yes"). The score for each respondent is the sum of positive answers. Dividing this score by 15, we obtain the respondent's rate of positive answers.

3. The *alcohol consumption* is measured by the question "How often did you drink alcohol in the last 30 days?". The answering options were (recoded values in parentheses): "Every day" (30); "4–6 times a week" (20); "2–3 times a week" (10); "Once a week" (4); "2–3 times in the last 30 days" (2,5); "Once in the last 30 days" (1); "I did not drink alcohol in the last 30 days" (0). The value observed for each respondent is divided by 30 and multiplied by 100, obtaining a 0–100 rate.

4. The frequency of undesirable *alcohol-related behaviour* set includes nine statements (e.g. "Have you ever

Table 1
Samples by group and wave

	1st wave			2nd wave		
Setting (1st/2nd wave)	Freq.	%	Valid %	Freq.	%	Valid %
Control groups						
PC / PC	200	11.1	11.1	188	10.4	11.7
SNO / SNO	200	11.1	11.1	187	10.4	11.6
SO / SO	200	11.1	11.1	179	9.9	11.1
Treatment groups						
PC / SNO	202	11.2	11.2	170	9.4	10.6
PC / SO	200	11.1	11.1	165	9.2	10.3
SNO / PC	200	11.1	11.1	182	10.1	11.3
SO / PC	200	11.1	11.1	184	10.2	11.4
SNO / SO	198	11.0	11.0	174	9.7	10.8
SO / SNO	200	11.1	11.1	179	9.9	11.1
Total	1,800	100.0	100.0	1,608	89.3	100.0
Nonresponse				192	10.7	
Total				1,800	100.0	

had sex under the influence of alcohol?"; the full list is available in Appendix 1/Set #4). The dummy variable ("yes/no"; 1 = "yes") is summed up in order to compute the score of each respondent. Dividing this score by 9 and multiplying the result by 100, we obtained the 0–100 respondent's rate.

5. The *monthly household income* (one question: "Approximately, what is the total monthly net income in your household?") is measured using the following categories (in brackets the recoded values, i.e. the intermediate levels of each class): 1) "Less than €1,000" (500); 2) "Between 1,001 and €1,500" (1,250); 3) "Between 1,501 and €2,500" (2,000); 4) "Between 2,501 and €3,500" (3,000); 5) "Between 3,501 and €4,500" (4,000); 6) "More than €4,500" (5,000). The recoded variable can be treated as continuous, because the distribution can be considered normal (see Appendix 2).

5 Findings

Table 2 presents preliminary results about the averages and the standard deviations of the five indices introduced in the previous section.

Table 2 shows that there are no systematic differences between the three groups. For two out of four indices, the higher values are observed for the *PC* group, whereas for the other two the higher averages correspond to the *SNO* group. Nevertheless, the differences between the three groups are very small for all indices: the absolute difference varies from 0.1% to 5.6%. If we consider the "Monthly household income", there are no differences at all both considering the median and the mode.

The five sensitive indices introduced above and analysed in Table 2 are further discussed in the following of this section: but the simple analysis shown in Table 2 shows that there seem to be no effect of the device used and of the optimization of the questionnaires in terms of reporting sensitive information.

In the next part of this section, we first examine our preliminary hypotheses (i.e. the first set), mainly linked to the survey context. Then, we analyse our main hypotheses (i.e. the second set), focusing on the respondents' survey experience. In the last part, we also test the effect of using a smartphone and of a mobile-optimized version of the questionnaire on reporting sensitive information.

5.1 The survey context

In order to evaluate the first set of our hypotheses (H1 to H3), we used a series of chi-square tests crossing some context variables with the used device (Table 3). Our results support H2 and, partially, H3, as explained in the following.

Concerning the *place to participate in surveys (H1)*, surprisingly the percentage of questionnaires filled at home is even higher for smartphones than for PCs (77.1% vs 73.3% in wave 1 and 81.7% vs 78.9% in wave 2). This unexpectedly higher percentage of respondents through smartphones from "home" can, at least in part, be attributed to our experiment design. Let's consider a respondent that usually participates in surveys using the PC, at home. Even if this respondent reads the invitation email, say, on public transports (on a smartphone), he/she may wait to start with the survey when he/she is at home. Once at home, only when starting the survey using the PC, he/she discovers that he/she is requested to switch to the smartphone (if he/she is part of the SO or of the

Table 2
Sensitive indices: mean and standard deviations by survey group (average of wave 1 and wave 2)

	Survey group					
	PC		SO		SNO	
Sensitive Indices	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Pos. attitude tow. deviant practices	22.3	14.3	21.1	13.7	21.4	13.5
Rate of deviant behaviour	20.1	13.0	19.8	12.9	20.5	13.4
Monthly alcohol consump. (times)	17.7	25.1	17.4	25.5	17.3	26.1
Rate of alcoholic behaviour	22.4	22.9	22.4	23.4	22.6	23.6
	Median	Mode	Median	Mode	Median	Mode
Monthly household income	€ 1501–2500	€ 1501–2500	€ 1501–2500	€ 1501–2500	€ 1501–2500	€ 1501–2500

Table 3
Context variables (by wave and mode)

		Device (%)[a]				
Variable	Categories	PC	S[c]	Chi-Square	p-value	N[b]
Wave 1						
Place of participation	Outside home	26.7	22.9	3.119	0.077	1,793
	At home	73.3	77.1			
Presence of third parties	Not present	80.2	73.0	11.083	0.001	1,795
	Present	19.8	27.0			
Presence of strangers	Not present	98.7	97.5	2.693	0.101	1,795
	Present	1.3	2.5			
Wave 2						
Place of participation	Outside home	21.1	18.3	1.748	0.186	1,602
	At home	78.9	81.7			
Presence of third parties	Not present	83.2	70.6	30.435	0.000	1,605
	Present	16.8	29.4			
Presence of strangers	Not present	98.7	97.2	3.674	0.055	1,605
	Present	1.3	2.8			

[a] = percentages by column and by variable.
[b] = valid cases (i.e., with values available for both crossed variables) [c] S = smartphones (SNO + SO).

SNO group). Thus, following the instructions, he/she probably stays at home. This situation could involuntarily cause an higher observed percentage of respondents participating from home.

Anyway, despite both studies have this same potential bias, our results differ from the ones of Mavletova and Couper (2013, p. 197). They observed a higher survey completion outside home for mobile devices (45%) than for PCs (29%). This discrepancy of our results might highlight how the smartphone participation in surveys is quickly evolving (since our data were collected few years later) and how common it is nowadays to use these devices even in a domestic environment. Our findings may also be related to the different context of our study: Spain is characterized by a different cultural background, by different attitudes in using mobile devices, and, thus, by a different approach in the smartphones' usage. Moreover, our results are consistent to what already found in other works (e.g., de Bruijne & Wijnant, 2013; Revilla et al., 2016): the preferred place to participate in surveys is still "home", even using a smartphone that provides respondents with a higher portability.

Table 3 shows the chi-square test applied to the variables "place of participation" (recoded into a dummy: "home/outside home") and "used device" ("PC/smartphone"). There is no significant link between place and used device for both waves (wave 1: $p = 0.077$; wave 2: $p = 0.186$). This confirms that the higher preference for filling the questionnaire at home for both the modes (>73%) is independent from the used device. Nevertheless, if for the variable "place of participation" we take into account all the original categories, the results change: the chi-square test suggests a significant link between the two variables for both the waves (df = 5; p = 0.000). This is due to the effect of the category "work". Considering together the places of participation "home" and "work", the preference for PCs is higher than for smartphones (96.6% vs 91.3% in wave 1 and 95.4% vs 92.5% in wave 2). If we also add the category "place of study", we observe a further increase of the difference in favour of PCs (wave 1: 97.8% vs 92.7% for smartphones; wave 2: 97.0% vs 93.3%). Maybe potential restrictions for the use of mobile devices at work or in places of study or the availability of a fixed-PC, or other unknown factors still make PC the favourite option. Overall, our findings do not confirm H1.

If we take into consideration H2 (regarding the *presence of third parties*), the percentage for the presence of bystanders is higher in places of study or, even more, in public transports (respectively, 60.9% and 64.3% in wave 1, and 70.6% and 76.9% in wave 2). The chi-square tests applied to the variables "presence of third parties" ("yes/no") and "place of participation" confirm a significant link (p < 0.0001) considering both the original categories of the variable "place of participation" and the recoded dummy "at home/outside home". Applying the chi-squared test to the variables "presence of third parties" ("yes/no") and "used device" ("PC/smartphone"), we detect a significant link in both wave 1 (p = 0.001) and wave 2 (p < 0.001). In particular, in wave 1, bystanders are present in 19.8% of the cases for PCs respondents and 27.0% for smartphones respondents. In wave 2, the difference is even bigger: 16.8% for PCs, 29.4% for smartphones. Thus, we found significantly higher percentages for the presence of third parties, when a smartphone is used. The second assumption is supported by our findings in both waves. Moreover, the percentages regarding the smartphones are in line with the findings of Mavletova and Couper (2013), mostly if we consider the wave 2 data: in this paper they found that 16% of surveys were completed in the presence of third parties when using a PC and 29% when using a mobile device. The panellists of Spain and of Russia are not different, from this point of view.

About H3, in both waves more than 94% of the questionnaires filled in *presence of strangers* were filled outside home. The percentage of survey in presence of strangers is 1.3% for PCs, and about 2.5% when a smartphone is used. Nevertheless, there is no significant link between the device used and the percentage of strangers (p = 0.101 for wave 1; p = 0.055 for wave 2). Overall, our findings do not completely support H3.

The evaluation of the first set of hypotheses allows us to define the survey context. Now we want to evaluate if having significantly higher percentages for the presence of third parties when smartphones are used influences respondents' behaviour and survey experience. In the next two subsections, we test the second set of hypotheses from two points of view. First, we study the perceived confidentiality, the sensitivity of the questions and how much the respondent feels (un)comfortable. Then, we take into account the willingness of respondents in reporting sensitive information.

5.2 The survey experience

First, we want to test if the higher probability of being surrounded by bystanders, with smartphones, affects how the respondents experience the survey (H4). In particular, we measured the concepts introduced hereunder by means of the following three questions[1]:

1. *No trust in the confidentiality*: the item asks if the respondent trusts the confidentiality of the survey; unipolar four-point scale recoded into a dummy: 0 = "I trust in the confidentiality" ("completely" + "partially" + "a little"); 1 = "I do not trust in the confidentiality".

2. *Sensitive questions*: the item asks if the respondent feels the questions as sensitive; four-point scale recoded into a dummy: 1 = "Sensitive questions" ("completely" + "a lot" + "a little"); 0 = "Not sensitive questions".

3. *Felt uneasy*: the question detects if the respondent felt uneasy during the survey; four-point scale, recoded into a dummy: 1 = "I felt uneasy answering to some questions"; 0 = "I did not feel uneasy answering to any questions".

In addition to this, by means of the following variable we wanted to evaluate the involvement of the respondent in potentially distracting activities, during the survey participation:

1. *Multi-tasking*: the original variable (asking which kind of other activities were performed during the survey) was recoded into a dummy: 1 = "At least one activity was performed"; 0 = "No other activity was performed".

Table 4 summarizes the chi-square tests results applied to each of these four variables and the used device (PC vs smartphone).

The variable "trust in the confidentiality of the survey", shows a slightly higher percentage of "do not trust" for smartphones (1.3% in wave 1 and 2.0% in wave 2) than for PCs (0.5% in wave 1 and 1.4% in wave 2). Nevertheless, there is no significant relation between this variable and the device used (p = 0.129). Similarly, if we consider the "perceived questions sensitivity", the data show not link with the used device (p = 0.368 in wave 1; p = 0.426 in wave 2). Nevertheless, the percentages observed for smartphones are slightly higher than the ones observed for PCs (94.7% vs 93.7% in wave 1; 94.0% vs 93.0% in wave 2). The chi-square test confirms the independence also between the variables "felt uneasy during the survey" and "used device" (p = 0.625 in wave 1; p = 0.079 in wave 2). Moreover, the percentage is not consistent over the two waves: more respondents felt uneasy during wave 1, when using smartphones (27.9% vs

[1] Note that the recode into dummy is used in order to ease the comparison with the 2013's Mavletova and Couper paper.

Table 4

Respondents' perception and behavior during the survey (by wave and device)

		Device (%)[a]				
Variable	Categories	PC	S[c]	Chi-Square	Sig.	N[b]
Wave 1						
No trust in confidentiality	Trust	99.5	98.7	2.299	0.129	1,797
	Do not trust	0.5	1.3			
Sensitive questions	Not sensitive	6.3	5.3	0.809	0.368	1,793
	Sensitive	93.7	94.7			
Felt uneasy	Did not feel uneasy	73.2	72.1	0.239	0.625	1,796
	Felt uneasy	26.8	27.9			
Multi-tasking	No other activities	29.1	24.8	3.791	0.052	1,800
	Other activities	70.9	75.2			
Wave 2						
No trust in confidentiality	Trust	98.6	98.0	0.628	0.428	1,599
	Do not trust	1.4	2.0			
Sensitive questions	Not sensitive	7.0	6.0	0.635	0.426	1,601
	Sensitive	93.0	94.0			
Felt uneasy	Did not feel uneasy	68.6	72.8	3.083	0.079	1,604
	Felt uneasy	31.4	27.2			
Multi-tasking	No other activities	28.5	25.2	2.015	0.156	1,608
	Other activities	71.5	74.8			

[a] percentages by column and by variable. [b] valid cases (i.e., values available for both the crossed variables)
[c] S = smartphones (SNO + SO).

26.8% of PCs), whereas for wave 2 we observe the opposite (27.2% for smartphones vs 31.4% for PCs). Finally, the "multi-tasking" (i.e. performing other potentially distracting activities during the survey participation) is more frequent, with smartphones (75.2% vs 70.9% observed for PC in wave 1; 74.8% vs 71.5% in wave 2). Nevertheless, the chi-square test suggests no significant link (wave 1: $p = 0.052$; wave 2: $p = 0.156$).

The same analysis was carried out testing the independence between the two smartphones groups (i.e., *SO* and *SNO*) and each of the four previously studied variables. No significant linkage is detected for all of these variables, and for both waves: "no trust in the confidentiality" (wave 1: $p = 0.795$; wave 2: $p = 0.546$); "perceived sensitivity of the questions" (wave 1: $p = 0.517$; wave 2: $p = 0.300$); "felt uneasy" (wave 1: $p = 0.742$; wave 2: $p = 0.338$). Moreover, the optimization has no significant link with the performance of other activities (wave 1: $p = 0.216$; wave 2: $p = .416$).

Summarizing our findings, smartphones make the participation from several different locations possible, and the mobile participation significantly increases the probability of the presence of bystanders during the survey. However, this all does not significantly affect the perceived privacy, nor the trust in the confidentiality of the survey, neither the feeling comfortable of respondents. There is independence between each of these aspects and the device used. The optimization of the survey also does not play a role from these perspectives. Moreover, even if, during the survey participation, the smartphones respondents could be more distracted (e.g. by reading emails, or talking with other persons), there is no significant link between the performance of these activities and the used device or the optimization of the survey. Thus, H4 is not supported by our findings.

5.3 Relative bias in reporting sensitive information

In this subsection, we study the potential differences in reporting sensitive information due to the use of smartphones rather than PCs (hypothesis H5) and due to an optimized rather than a not-optimized version of the survey (H6). We focus on the five sensitive indices introduced in Section 4.

Analogously to what done in Mavletova and Couper (2013), we performed a *Linear Mixed Models* analysis (*LMM*; see West, Welch, & Gałecki, 2007). These models are a specific case of the general linear models taking into consideration the linear relationship between factors/covariates[2] and a dependent variable; they also allow

[2]We define factors as categorical predictors. In LMM each level of a factor can have a different linear effect on the value of the dependent variable. We define covariates as continuous/scale predictors. The values of the covariates are assumed to have a linear correlation with the values of the dependent variable within combinations

modelling the covariance structure of the error terms. This is useful because we can take into account the random effect linked to each respondent and the within-subject correlation (and non-constant variability) between the two waves (errors between the two waves can be considered correlated for the same respondent).

The estimated model includes, at level 1, the fixed effects of the wave ("Wave" variable, also inserted as a repeated effect component), and, at level 2, the subject ("Panellist ID"). In addition to "Wave", as fixed effects we included: the survey condition/group ("Cond": *PC*, *SNO* or *SO*), the gender ("Gender"), the presence of bystanders ("Bystd") and the place of participation ("Home"). This model is very similar to the Mavletova and Couper (2013) one, except that we inserted "Age" as a covariate, instead of as a fixed factor (considering the dummy based on the age group 18-34), and except the inclusion of the variables "Bystd" and "Home"[3]. The model can be formalized as follows:

$$Y_{it} = \beta_{00} + \beta_C \text{Cond} + \beta_G \text{Gender} + \beta_W \text{Wave} + \beta_B \text{Bystd} + \beta_H \text{Home} + \beta_A \text{Age} + u_{0i} + \epsilon_{ti} \quad (1)$$

where Y_{it} is the value of a certain sensitive index (Y) computed for the i-th respondent for time t (wave 1 or 2), β_C is the fixed effect of the survey condition/group (PC, SNO, SO), β_G is the fixed effect of "Gender" ("male" = 1, "female" = 2), β_W is the fixed effect of "Wave" ("wave 1" = 1, "wave 2" = 2), ; β_B is the fixed effect of the presence of bystanders, "Bystd" ("yes" = 1, "no" = 0), β_H is the fixed effect of the place of participation, "Home" ("at home" = 1, "outside home" = 0), β_A is the effect of the covariate "Age" (numeric variable), and u_{0i} and ϵ_{ti} are the random effects at the two levels.

Minimizing the -2 Restricted Log Likelihood information criterion and other adjusted criteria (e.g., the Akaike's Information Criterion and the Schwartz's Bayesian Criterion), we found that models with repeated effects fit better than models without repeated effects, also leading to a reduction of the estimates' standard error. Moreover we studied the residual covariance (R) matrix: when the variances of the two waves were nearly equal, we opted for a stricter covariance structure, using a compound symmetry covariance structure; in the other cases we requested an unstructured covariance structure[4]. In order to check whether the compound symmetry variance structure was more appropriate, we computed a likelihood ratio test for the equal variances hypothesis (test based on the -2 Restricted Log Likelihoods statistics). The best choice is shown in the second row of Table 5. As method of estimation, we tested both the *Restricted Maximum Likelihood (REML)* and the *Maximum Likelihood (ML)*. The results were similar. Here we present the models obtained using the REML, because this method is more likely to produce unbiased estimates of the variance and of the covariance parameters.

Table 5 shows the estimated models for the different sensitive indices.

Our results do not support hypothesis *H5*. Within the five sensitive indices, the device effect is significant only for one ("Income"; $p = 0.039$). An underreporting of the household income is detected when a smartphone is used with an optimized version of the survey. Nevertheless, running the analysis again using a dummy variable "Settings" (0 = "PC"; 1 = "smartphone"), this variable is not anymore significant for the index "Income" ($p = 0.067$), neither for the other indices. However, these findings generally support hypothesis H6: a significant effect of the optimization of the survey for a mobile participation on reporting sensitive information is observed only for one index out of five.

Looking to the other variables of the model, for "Gender" results are consistent with the Mavletova and Couper (2013) paper: males usually report significantly higher levels for all the indices. "Wave" is not significant for any of the indexes, but one: there is a significantly higher reporting for "Alcohol behaviour" in the first wave ($p = 0.000$). The presence of bystanders is never significant. The place of participation is only significant for two out of five indices: for "Deviant behaviour" ($p = 0.041$) and for "Alcohol behaviour" ($p = 0.025$), with a higher reporting of these behaviours when the participation takes place outside home. Finally, the variable "Age" is significant for all the sensitive indexes (all p-values < 0.001). Older respondents report higher "Alcohol consumption" and "Income", and lower "Positive attitude towards deviant practices", "Deviant behaviour" and "Alcohol behaviour".[5]

6 Conclusions

This research was conceived as a replication of the experiment discussed in the Mavletova and Couper (2013) paper, with some adaptations. We implemented a two-wave web survey with a cross-over design. The same survey was proposed to the same panellists, randomly assigning them to a specific experimental group (*PC*, *SNO* or *SO*) in each wave.

of factor levels (or cells).

[3] Estimates of the LMMs excluding these last two variables showed very similar results.

[4] Choosing the compound symmetry structure, we require a constant variation and a constant covariation (i.e., we assume to have constant variance for the repeated measures). Using an unstructured repeated covariance, we allow the repeated measures to be correlated and to have unequal variances. For each sensitive indices, we actually apply both methods. Then, we tested whether restricting the variances to be equal across waves would produce better estimates.

[5] We also re-run the complete analysis with the variable "Age" recoded into the same classes used by Mavletova and Couper (2013): "18-34" and ">34". The results obtained did not change: the variable was still always significant (all p-values < 0.001).

Table 5
Linear mixed models coefficients by sensitive indices

	Positive attitude towards deviant practices		Deviant Behaviour		Alcohol Consumption		Alcohol Behaviour		Income	
Covar. struct.	Compound		Unstructured		Compound		Unstructured		Unstructured	
Parameter	Est.	Std. Err.	Est.	Std. Err.	Est.	Std. Err.	Est.	Std. Err.	Est.	Std. Err.
Intercept	33.10***	1.004	22.31***	1.069	−7.39***	2.035	36.90***	1.883	1340.87***	92.78
Group: SO	−0.19	0.344	−0.38	0.288	0.05	0.687	0.06	0.480	−51.21*	24.74
Group: PC	0.45	0.346	−0.31	0.290	0.40	0.692	0.12	0.484	12.56	24.81
Gender: M	1.96***	0.544	4.36***	0.593	7.24***	1.106	5.78***	1.051	105.95*	51.69
Wave: First	0.13	0.213	0.10	0.174	0.11	0.424	1.12***	0.287	5.00	14.78
No Bystand.	−0.49	0.366	−0.11	0.313	−0.25	0.733	−0.94	0.523	−2.00	27.02
Outs. Home	0.43	0.429	0.77*	0.375	0.98	0.857	1.41*	0.631	46.99	32.65
Age	−0.20***	0.025	−0.11***	0.027	0.58***	0.050	−0.48***	0.048	18.08***	2.35

A second version of the model with a different reference category (PC) for survey group is shown in Appendix C.
* $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$

We allowed respondents to participate only by means of the selected device. We involved 1,800 panellists of the opt-in online panel Netquest, focusing on the panel for Spain.

We studied the device effect, comparing PCs and smartphones. In addition, we studied if there was an effect due to the optimization of the survey for a mobile participation. For reaching these objectives, similarly to Mavletova and Couper (2013), first, we evaluated the respondents' survey experience. Then, we studied the relative bias in reporting sensitive information.

Our main findings are the following:

• *Place of participation*: even if smartphone respondents are freer to answer from any place, the most recurrent place of participation is "home" (we should remind that these high percentages could be explained, at least to a certain extent, by our experimental design, as explained in section 5.1).

• *Presence of third parties and of strangers*: it is significantly higher when a smartphone is used.

• *Survey experience*: there is no significant device effect on the perceived confidentiality of the survey, on the perceived sensitivity level of the questions and on the feeling uncomfortable during the survey.

• *Reporting of sensitive information*: generally there is no significant impact of the device (PC vs smartphone) and of the context of participation (in terms of presence of bystanders) on the studied sensitive indices. Other background variables seem to be more linked to the reporting of sensitive information, like "Gender" and "Age". For two out of five indices also the place of participation (at home/outside home) has a significant effect.

• *Optimization of the survey* for a mobile participation: it does not play a key role in affecting the reporting of sensitive information (its effects in the LMMs we estimated is significant only for one index out of five).

On the one hand, these results confirm the robustness of the majority of the Mavletova and Couper (2013) findings. Indeed, these authors also found higher proportions for the presence of third parties and strangers for smartphone respondents. Concerning the survey experience, they also did not find any significant effect of the device on the perceived privacy. About the reporting of sensitive information, they also found that the comparability of data collected by means of PCs rather than by means of smartphones is not generally affected by the used device, when sensitive information is asked. Thus, this suggests that several of the findings about the comparison of PCs and smartphones for web survey completion are robust to the survey context, since they are verified in two very different countries (in terms of Internet and mobile web coverage and usage, as well as in terms of culture) and at two different moments in time. Therefore, we can presume that our results can probably be extended to other countries.

On the other hand, some of our results differ from the ones of Mavletova and Couper (2013). In particular, they found that the link between the device and the place of participation is significant, whereas we found that it is not. Besides, they found that "respondents tended to trust in data confidentiality more when they completed the questionnaire on a PC" (Mavletova & Couper, 2013, p. 200), but we could not confirm this. All this suggests that some of the findings are influenced by the research context. First, the use of smartphones has become even much more common within the lag of time between the two experiments. For example, we found that it has been as much popular as the use of PCs (or even more than this) to participate in our survey, also in a domestic environment. This change can play a role to explain why our findings differ from the ones of Mavletova and Couper (2013). Second, the differences in culture or tendencies in the use of mobile devices due to the different national context (Russia and Spain) can also have an impact. It seems that these findings cannot be easily generalised to different survey contexts.

To summarize, some of our results confirm the robustness of Mavletova and Couper's (2013) findings, whereas others highlight that some aspects of the mobile survey participation are different in the two contexts (e.g. the perceived trust in data confidentiality).

Our study suffers from some limitations. Some of them are similar to the ones of Mavletova and Couper's (2013) paper. First, we involved panellists of an online opt-in panel that are not representative of the general population. Thus, further studies with probability samples could be developed. Second, studies involving a wider target population or a wider set of countries would be helpful. Moreover, our findings are referred to the smartphones usage only. Nevertheless, more research is needed for other kinds of mobile devices (e.g., tablets, for whose the bigger screens can emphasize the desirability bias phenomenon). Finally, as mentioned by Mavletova and Couper (2013, p. 202), "the differences between PC and mobile-based web surveys may depend on the types of questions being compared". This is also confirmed by our results (significant effect of the device found only for one out of five). Thus, further studies are needed in order to check the robustness of the results across a wider range of topics. Moreover, our study is mostly focused on the study of relative measurement error: further research can also take into account other aspects, such as coverage error, completion times, and break-off rates.

Acknowledgements

We would like to thank, first, Aigul Mavletova and Mick Couper for their precious help, for their suggestions and advices in setting up the experiment and for inspiring this work from the beginning. We are also very thankful to Netquest for providing us with the data used in this paper, and in particular to Carlos Ochoa for his advices during the whole process. Finally, we would like to thank the University of Bergamo: this research has been partially supported by the 60% University funds.

References

Andreadis, I. (2015). Comparison of response times between desktop and smartphone users. In D. Toninelli, R. Pinter, & P. de Pedraza (Eds.), *Mobile research methods: opportunities and challenges of mobile research methodologies* (pp. 63–79). London: Ubiquity Press. doi:10.5334/bar.e

Antoun, C. (2015a). *Effects of mobile versus pc web on survey response quality: a crossover experiment in a probability web panel*. Paper presented at the Annual Conference of the Americal Association for Public Opinion Research, Hollywood, FL.

Antoun, C. (2015b). Who are the Internet users, mobile Internet users, and mobile-mostly Internet users? Demographic differences across Internet-use subgroups in the U.S. In D. Toninelli, R. Pinter, & P. de Pedraza (Eds.), *Mobile research methods: opportunities and challenges of mobile research methodologies* (pp. 99–117). London: Ubiquity Press. doi:10.5334/bar.g

Baker-Prewitt, J. (2013). *Mobile research risk: what happens to data quality when respondents use a mobile device for a survey designed for a PC?* CASRO Online Research Conference 2013 Proceedings, San Francisco, US, 2013, Burke, Cincinnati, OH.

Bosnjak, M., Poggio, T., Becker, K., Funke, F., Wachenfeld, A., & Fischer, B. (2013). *Online survey participation via mobile devices*. Paper presented at the 68th Annual Conference of the American Association for Public Opinion Research (AAPOR), 16–19 May, Boston, Massachusetts, US.

Buskirk, T. & Andrus, C. (2012). *Online surveys aren't just for computers anymore! Exploring potential mode effects between smartphone vs. computer-based online surveys*. Paper presented at AAPOR Annual Conference. May 17–20, 2012. Orlando, USA.

Buskirk, T. & Andrus, C. (2014). Making mobile browser surveys smarter: results from a randomized experiment comparing online surveys completed via computer or smartphone. *Field Methods*, *Published online before print*. doi:10.1177/1525822X14526146

Buskirk, T. & C, A. (2012). Smart surveys for smart phones: exploring various approaches for conducting online mobile surveys via smartphones. *Survey Practice*, *5*(1), 1–11.

Chae, M. & Kim, J. (2004). Do size and structure matter to mobile users? An empirical study of the effects of screen size, information structure, and task complexity on user activities with standard web phones. *Behaviour & Information Technology*, *23*, 165–181.

Consumer Barometer. (2014). Retrieved from https://www.consumerbarometer.com/en/

Couper, M. (2014). *Internet survey methods: the past, the present and the future*. Keynote opening address of the ISM – Internet Survey Methodology workshop, December 2–4, 2014. Bozen, Italy. Retrieved from http://www.unibz.it/en/economics/research/ism2014/default.html

Crowne, D. & Marlowe, D. (1964). *The approval motive: studies in evaluative dependence*. New York: Wiley.

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.

de Bruijne, M. & Wijnant, A. (2014). Mobile response in web panels. *Social Science Computer Review*, *32*(6), 728–742.

Fuchs, M. & Busse, B. (2009). The coverage bias of mobile web surveys across european countries. *International Journal of Internet Science*, *4*, 21–33.

Guidry, K. (2012). *Response quality and demographic characteristics of respondents using a mobile device on a web based survey*. Paper presented at AAPOR Annual Conference. May 17–20, 2012. Orlando, USA.

Internet Live Stats. (2015). Retrieved from http://www.internetlivestats.com/internet-users-by-country/

Kreuter, F., Presser, S., & Tourangeau, R. (2008). Social desirability bias in CATI, IVR, and web surveys: the effects of mode and question sensitivity. *Public Opinion Quarterly*, *72*(5), 847–865.

Mavletova, A. & Couper, M. (2013). Sensitive topics in PC web and mobile web surveys: is there a difference? *Survey Research Methods*, *7*(3), 191–205.

Mavletova, A. & Couper, M. (2015). A meta-analysis of breakoff rates in mobile web surveys. In D. Toninelli, R. Pinter, & P. de Pedraza (Eds.), *Mobile research methods: opportunities and challenges of mobile research methodologies* (pp. 81–98). London: Ubiquity Press. doi:10.5334/bar.f

McClain, C., Crawford, S., & Dugan, J. (2012). *Use of mobile devices to access computer-optimized web instruments: implications for respondent behavior and data quality*. Paper presented at AAPOR Annual Conference. May 17–20, 2012. Orlando, USA.

Mohorko, A., de Leeuw, E., & Hox, J. (2013). Internet coverage and coverage bias in europe: developments across countries and over time. *Journal of Official Statistics*, *29*(4), 609–622. doi:10.2478/jos-2013-0042

Newell, S., Logan, H., Guo, Y., Marks, J., & Shepperd, J. (2015). Evaluating tablet computers as a survey tool in rural communities. *The Journal of Rural Health*, *31*(1), 108–117.

Paulhus, D. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology*, *46*(3), 598–609.

Peterson, G. (2012). *Unintended mobile respondents*. Paper presented at CASRO Technology Conference, 31 May, New York, NY. Retrieved from http://c.ymcdn.com/sites/www.casro.org/resource/collection/D0686718-163A-4AF4-A0BB-8F599F573714/Gregg%5C_Peterson%5C_-%5C_Market%5C_Strategies.pdf

Peterson, G., Mechling, J., LaFrance, J., Swinehart, J., & Ham, G. (2013). *Solving the unintentional mobile challenge*. Paper presented at the CASRO Online Research Conference, San Francisco, USA.

Peytchev, A. & Hill, C. (2010). Experiments in mobile web survey design: similarities to other modes and unique considerations. *Social Science Computer Review*, *28*(3), 319–335.

Revilla, M. & Ochoa, C. (2015). Open narrative questions in pc and smartphones: is the device playing a role? *Quality and Quantity*, *First published online*. doi:s11135-015-0273-2

Revilla, M., Toninelli, D., Ochoa, C., & Loewe, G. (2016). Do online access panels need to adapt surveys for mobile devices? *Internet Research*, *26*, accepted, forthcoming.

Sakshaug, J., Yan, T., & Tourangeau, R. (2010). Nonresponse error, measurement error, and mode of data collection: tradeoffs in a multi-mode survey of sensitive and nonsensitive items. *Public Opinion Quarterly*, *74*(5), 907–933.

Schaeffer, N. (2000). Asking questions about threatening topics: a selective overview. In A. Stone, C. Bachrach, J. Jobe, H. Kurtzman, & CainV.S. (Eds.), *The science of self-report: implications for research and practice* (pp. 105–121). Mahwah, NJ: Lawrence Erlbaum Associates.

Stapleton, C. (2013). The smart(phone) way to collect survey data. *Survey Practice*, *6*(2), 1–7.

StatCounter Global Stats. (2015). Retrieved from http://gs.statcounter.com/%5C#desktop+mobile-comparison-ww-monthly-201202-201502

Steenkamp, J.-B., de Jong, M., & Baumgartner, H. (2009). Socially desirable response tendencies in survey research. *Journal of Marketing Research*, *46*, 1–74.

Struminskaya, B., Weyandt, K., & Bosnjak, M. (2015). The effects of questionnaire completion using mobile devices on data quality. evidence from a probability-based general population panel. *Methods, data, analyses*, *9*(9), 261–292.

Sweeney, S. & Crestani, F. (2006). Effective search results summary size and device screen size: is there a relationship? *Information Processing & Management*, *42*, 1056–1074.

Toepoel, V. & Lugtig, P. (2014). What happens if you offer a mobile option to your web panel? Evidence from a probability-based panel of internet users. *Social Science Computer Review*, *32*(4), 1–17.

Toninelli, D., Pinter, R., & de Pedraza, P. (Eds.). (2015). *Mobile research methods: opportunities and challenges of mobile research methodologies*. London: Ubiquity Press. doi:10.5334/bar

Tourangeau, R., Groves, R., C., K., & Yan, T. (2009). The presentation of the survey, nonresponse, and measurement error. *Journal of Official Statistics*, *25*, 299–321.

Tourangeau, R., Groves, R., & Redline, C. (2010). Sensitive topics and reluctant respondents: demonstrating a link between nonresponse bias and measurement error. *Public Opinion Quarterly*, *74*(3), 413–432.

Tourangeau, R. & Yan, T. (2007). Sensitive questions in surveys. *Psychological Bulletin*, *133*(5), 859–883.

Villar, A., Callegaro, M., & Yang, Y. (2013). Where am I? A meta-analysis of experiments on the effects of progress indicators for web surveys. *Social Science Computer Review*, *31*(6), 744–762. doi:10 . 1177 / 0894439313497468

Wells, T., Bailey, J., & Link, M. (2013). Filling the void: gaining a better understanding of tablet-based surveys. *Survey Practice*, *6*(1), 1–9.

Wells, T., Bailey, J., & Link, M. (2014). Comparison of smartphone and online computer survey administration. *Social Science Computer Review*, *32*(2), 238–255.

West, B., Welch, K., & Gałecki, A. (2007). *Linear mixed models: a practical guide using statistical software*. Boca Raton: Chapman & Hall/CRC.

Appendix A
Set of questions

Set #1:

"For each of the following activities, please indicate at what extent you believe they can be justified". *Response scale:* 1 "Can always be justified"; 2 "Can often be justified"; 3 "Can sometimes be justified"; 4 "Can never be justified".

1. Lying in your own interest
2. Cheating on taxes
3. Accepting a bribe in the course of one's own duties
4. Paying cash for services to avoid taxes
5. Prostitution
6. Married people having an extra-marital affair
7. Married people lying to their partners
8. Having casual sex
9. Avoiding a fare on public transport
10. Stealing in a shop
11. Taking marijuana or hashish
12. Abortion
13. Euthanasia (i.e. terminating the life of a terminally ill)
14. Suicide
15. Scientific experiments on human embryos

Set #2

"Have you ever done the following actions?" *Response options:* "Yes" / "No".

1. Have you ever stolen something in a shop?
2. Have you ever taken marijuana, hashish or ecstasy?
3. Have you or your relatives ever given bribe for some services?
4. Have you ever failed to report a crime committed by other people to the authorities?
5. Have you ever found a wallet without returning it back?
6. Have you ever had casual sex?
7. Have you ever been unfaithful to one of your partners?
8. Have you ever simulated illness in order to receive sick leave?
9. Have you ever tried to commit suicide?
10. Have you ever taken money or things from other people without their permission and without confessing it?
11. Have you ever been treated because of a venereal disease?
12. Have you ever watched pornographic movies, or websites, or magazines?
13. Have you ever had a same sex experience?
14. Have you ever accepted a bribe?
15. Have you ever taken out a loan, knowing that you cannot pay the money back?

Set #4

"Please, answer the following questions". *Response options:* "Yes" / "No".

1. Have you ever been drunk during several days?
2. Have you ever had sex under the influence of alcohol?
3. Have you ever felt ashamed by what you had done the day before when you were drunk?
4. Have you ever drunk alcohol alone?
5. Have you ever felt you needed to drink alcohol early in the morning to steady your nerves or get rid of a hangover?

6. Have you ever forgotten something because of alcohol?
7. Have you ever missed a class or a day of work because of drinking?
8. Have you ever drunk so much alcohol that you could not control yourself?
9. Have you ever lost control while drinking alcohol?

Appendix B
Distribution of monthly household total income

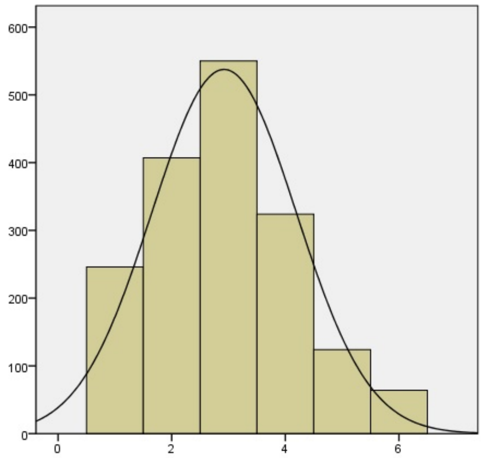

Figure B1. Wave 1, Original variable (distribution by classes)

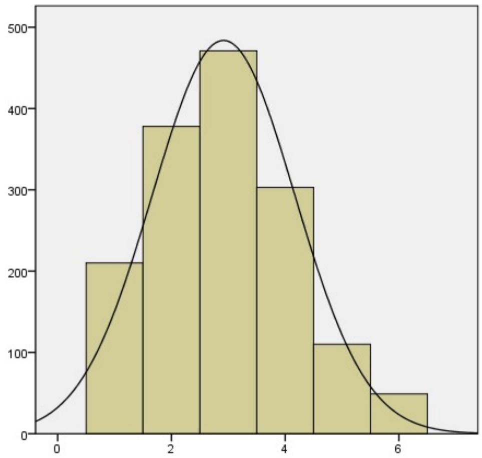

Figure B2. Wave 2, Original variable (distribution by classes)

Table B1
Recoded values (normality statistics):

Statistics	Wave 1	Wave 2
Valid cases	1,715	1,521
Missing	85	279
Median	2,000	2,000
Mode	2,000	2,000
Skewness	.695	.666
Std. Error of Skewness	.059	.063
Kurtosis	.080	.033
Std. Error of Kurtosis	.118	.125

Appendix C
LMM for sensitive indices (version #2 – reference category: PC)

Table C1
Linear mixed models coefficients by sensitive indices

	Positive attitude towards deviant practices		Deviant Behaviour		Alcohol Consumption		Alcohol Behaviour		Income	
Covar. struct.	Compound		Unstructured		Compound		Unstructured		Unstructured	
Parameter	Estimate	Std. Err.	Estimate	Std. Err.	Estimate	Std. Err.	Estimate	Std. Err.	Estimate	Std. Err.
Intercept	33.10***	1.004	22.31***	1.069	−7.39***	2.035	36.90***	1.883	1340.87***	92.78
Group: SO	−0.63	0.346	−0.07	0.291	−0.35	0.692	−0.06	0.484	−63.78*	24.83
Group: SNO	−0.45	0.346	−0.31	0.290	−0.40	0.692	−0.12	0.484	−12.56	24.81
Gender: M	1.96***	0.544	4.36***	0.593	7.24***	1.106	5.78***	1.051	105.95*	51.69
Wave: First	0.13	0.213	0.10	0.174	0.11	0.424	1.12***	0.287	5.00	14.78
No Bystand.	−0.49	0.366	−0.11	0.313	−0.25	0.733	−0.94	0.523	−2.00	27.02
Outs. Home	0.43	0.429	0.77*	0.375	0.98	0.857	1.41*	0.631	46.99	32.65
Age	−0.20***	0.025	−0.11***	0.027	0.58***	0.050	−0.48***	0.048	18.08***	2.35

* $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$