Journal of Official Statistics, Vol. 38, No. 4, 2022, pp. 987–1017, http://dx.doi.org/10.2478/JOS-2022-0043

Response Burden and Dropout in a Probability-Based Online Panel Study – A Comparison between an App and Browser-Based Design

Caroline Roberts[1], *Jessica M.E. Herzing*[2], *Marc Asensio Manjon*[1], *Philip Abbet*[3], *and Daniel Gatica-Perez*[4]

Survey respondents can complete web surveys using different Internet-enabled devices (PCs versus mobile phones and tablets) and using different software (web browser versus a mobile software application, "app"). Previous research has found that completing questionnaires via a browser on mobile devices can lead to higher breakoff rates and reduced measurement quality compared to using PCs, especially where questionnaires have not been adapted for mobile administration. A key explanation is that using a mobile browser is more burdensome and less enjoyable for respondents. There are reasons to assume apps should perform better than browsers, but so far, there have been few attempts to assess this empirically. In this study, we investigate variation in experienced burden across device and software in wave 1 of a three-wave panel study, comparing an app with a browser-based survey, in which sample members were encouraged to use a mobile device. We also assess device/software effects on participation at wave 2. We find that compared to mobile browser respondents, app respondents were less likely to drop out of the study after the first wave and the effect of the device used was mediated by subjective burden experienced during wave 1.

Key words: App-based survey; attrition; mobile survey; usability; user satisfaction.

1. Introduction

Mobile Internet technologies presents numerous opportunities for survey research, as well as some important challenges (Link et al. 2014). Respondents can now access online surveys via web browsers on a number of different Internet-enabled devices (notably, desktop PCs and laptops, tablets and smartphones (Callegaro 2010; De Bruijne and Wijnant 2014a; Lugtig and Toepoel 2016; Peytchev and Hill 2010; Struminskaya et al.

[1] Institute of Social Sciences, Faculty of Social and Political Sciences University of Lausanne, Bâtiment Géopolis, Quartier Mouline, 1015 Lausanne, Switzerland. Emails: Caroline.Roberts@unil.ch and marc.asensiomanjon@unil.ch
[2] Interfaculty Centre for Educational Research (ICER), University of Bern, Fabrikstrasse 8, 3012 Bern, Switzerland. Email: jessica.herzing@unibe.ch
[3] Idiap Research Institute, Rue Marconi 19, 1920 Martigny, Switzerland. Emails: philip.abbet@idiap.ch and gatica@idiap.ch.
[4] EPFL (Ecole polytechnique fédérale de Lausanne), School of Engineering and College of Humanities, Inn building, Station 14, 1015 Lausanne, Switzerland. Email: daniel.gatica-perez@epfl.ch
Acknowledgments: This work was supported by the EPFL and University of Lausanne "Collaborative Research on Science and Society" (CROSS) 2019 Program, as part of the project "Leveraging on-device smartphone inference to address resistance to participate in social surveys (LOIS)"; by FORS, the Swiss Centre of Expertise in the Social Sciences; and by the Faculty of Social and Political Sciences at the University of Lausanne. We would like to thank Oliver Bornet (Idiap) for technical discussions about the mobile platform. We also thank all participants in the study.

2015)). While this range of access options means that long-standing coverage problems associated with web surveys are diminishing (Couper et al. 2017; Lee et al. 2019), research has highlighted difficulties around ensuring data quality where multiple response devices are used in the same survey (Antoun et al. 2017; Callegaro 2013; Lee et al. 2019; Revilla et al. 2016). There is evidence that compared to on PCs, answering questionnaires on mobile devices can take respondents longer (Couper et al. 2017; Couper and Peterson 2017), increase breakoff risk (Buskirk and Andrus 2012; Callegaro 2010; Couper et al. 2017; Mavletova and Couper 2015; Peytchev 2009), and affect measurement quality (Antoun 2015; Mavletova 2013; Mavletova and Couper 2013, 2015). Optimising online surveys for mobile devices in ways that motivate participation, engage respondents, and promote conscientious questionnaire completion has, therefore, become a key priority for survey practitioners (Antoun et al. 2018; De Bruijne and Wijnant 2014b; Mavletova et al. 2018; Peytchev and Hill 2010). This is especially important given the growing use of web-based data collection in longitudinal surveys, where mitigating panel attrition is a central, ongoing challenge (De Leeuw and Lugtig 2014).

Understanding the reasons why response behaviour and engagement vary across different access options is key to improving the design of future online surveys. While these reasons are manifold and intrinsically linked to the characteristics of respondents using them (Lugtig and Toepoel 2016), device and software usability and experienced burden have been identified as parts of the Equation (Callegaro et al. 2015; Couper et al. 2017). Web questionnaires designed for PCs are not always well-adapted to browsers on smaller touch-screen devices, and mobile internet connections (and being physically mobile) are less conducive to sustained concentration on questionnaires over long periods (Antoun et al. 2018; Callegaro et al. 2015; Couper et al. 2017) making the response task more demanding, and time-consuming (Antoun et al. 2018). Mobile software applications (apps) designed for hosting survey questionnaires offer ways to address some of these constraints, as well as other ways to make mobile-web surveys less burdensome and more engaging (Link et al. 2014). Most research investigating the utility of apps for surveys to date, however, has focused on the new measurement tools they offer and respondent willingness to use them to complete alternative data collection tasks (e.g., Jäckle et al. 2019; Keusch et al. 2019; Revilla et al. 2017; Wenz et al. 2019). Few studies have explicitly investigated whether perceptions and experiences of response burden among app respondents are more positive compared to respondents using mobile browsers, and if so, whether this, in turn, affects response behaviour and participation decisions.

In this study, we address this knowledge gap by investigating variation in experienced burden in the first wave of a three-wave panel study, across chosen response devices (PCs (desktop and laptops) verses mobiles (smartphones and tablets) and software (web browser versus mobile app), and the extent to which burden mediates device and software effects on willingness to participate in the subsequent panel wave. Specifically, we address the following research questions:

RQ1: To what extent does experienced response burden vary as a function of response device and software?

RQ2: Does willingness to participate at wave 2 of an online panel study vary as a function of the response device/ software used at wave 1?

RQ3: Does experienced response burden at wave 1 mediate device/software effects on willingness to participate at wave 2?

Before describing the research design and analytic approach in detail, we first review relevant literature relating to design challenges in mixed device web surveys and the problem of response burden – particularly in a panel setting. We also describe some of the opportunities apps offer for optimising web survey design for participants on mobile devices.

2.1. Design Challenges in Web Surveys and the Problem of Response Burden

While the proliferation of access options available for completing web surveys represents good news on the one hand for survey practitioners, on the other, research into its implications for data quality has highlighted areas for concern (De Bruijne and Wijnant 2014b; Maslovskaya et al. 2019; Toepoel and Lugtig 2015). Like modes of data collection, different devices have their own error properties (Couper et al. 2017), which can affect data comparability in mixed device settings (Toepoel and Lugtig 2015). While potentially problematic, measurement differences between devices are generally viewed as a lesser cause for concern than differences resulting from device-related selection errors (Couper et al. 2017; Keusch and Yan 2017; Toepoel and Lugtig 2018; Antoun et al. 2019; Struminskaya et al. 2015), that is, errors resulting from non-coverage and nonresponse which affect who is selected into the response sample (Klausch et al. 2015), and hence, its representativeness.

There are multiple explanations for differences in selection error in estimates based on data gathered from respondents on different devices. Firstly, different devices (and different brands of device and operating systems) tend to be used by different socio-demographic groups, who may, in turn, be more or less likely to use those devices to participate in web surveys. Secondly, in *mobile* web surveys, in which participation is only allowed on a mobile device, unit nonresponse rates are generally higher than for PC web surveys, suggesting an overall negative impact on response propensity of mobile modes. This has been attributed to both respondent and device characteristics (e.g., respondents' level of familiarity with the device and how they habitually use it versus its technological features, like the speed and reliability of internet connection (Couper et al. 2017). Thirdly, user characteristics, device features and environmental influences can all affect respondents' motivation to finish answering the questionnaire once started, as any difficulties experienced (e.g., technical problems, distractions) may cause them to want or have to stop (Link et al. 2014, 19). Indeed, there is consistent evidence that respondents on mobile devices have a higher breakoff rate than those responding on desktop and laptop computers (Buskirk and Andrus 2014; Couper and Peterson 2017; Couper et al. 2017; Guidry 2012; Wells et al. 2013) and that those on smartphones are more likely to quit the survey compared to those on tablets (where breakoff rates for the latter are closer to those on PCs (Guidry 2012; Wells et al. 2013). Less is known about whether, for the same reasons, mobile respondents in web-based panel studies are more likely to drop out of the study in subsequent waves of data collection, but there are reasons (reviewed below) to assume this may well be the case.

To mitigate non-comparability in data quality across response devices in web surveys it is essential to address the underlying mechanisms responsible for negative outcomes for mobile respondents. Explanations for differences in response behaviour by device (besides user characteristics) frequently invoke concepts relating to device usability and response burden. For example, Mavletova and Couper's (2015, 93) meta-analysis found breakoff rates for mobile respondents to be highest for longer questionnaires with complex design elements (e.g., grids, sliders and images) and for questionnaires designed for completion on a PC (see also Callegaro 2010). The latter require respondents to make additional effort on smaller screens (e.g., to scroll down the screen to read long questions or long lists of response options) and data input using a touch screen may be less comfortable and more error prone than on larger devices (Link et al. 2014).

Studies have also consistently found differences in completion times for respondents answering questionnaires on different devices, with mobile respondents taking longer, on average, than PC respondents (Antoun and Cernat 2020; Couper et al. 2017; Couper and Peterson 2017; Keusch and Yan 2017; Mavletova and Couper 2015). Like breakoffs, longer completion times have been attributed to device-related factors such as the greater need to scroll due to smaller screen sizes, the demands of text input without a keyboard, and increased transmission times due to connection speed (Antoun and Cernat 2020; Couper et al. 2017; Couper and Peterson 2017; Keusch and Yan 2017; Mavletova and Couper, 2013; 2015).

Contextual factors have also been alluded to, however, including mobile respondents being on the move when completing questionnaires, being in the presence of other people or exposed to other distractions, and multi-tasking (Couper et al. 2017; Wenz 2021). Longer completion times should indicate increased objective burden to the extent that they imply greater time sacrifice to complete the survey task and more effort required to finish (Office of Management and Budget 2006). However, the relationship with the respondent's subjective experience of burden is less clear cut, as taking longer to complete a questionnaire may also imply more task engagement (Lynn 2014; Read 2019). For this reason, it is important to take account of the interplay between subjective perceptions of burden and objective hindrances to participation that may influence willingness to continue participating in a survey (Couper et al. 2017), particularly in a longitudinal research setting.

In longitudinal surveys, response burden has also been cited as a common reason for attrition (Hoogendoorn and Sikkel 1998; Laurie 2008; De Leeuw and Lugtig 2014; Kleinert et al. 2019). Attrition refers to respondents dropping out of a panel study, either temporarily or permanently, and can be attributed to a variety of causes, including variation in intrinsic motivations to participate in and commitment to a study, as well as extrinsic factors such as incentives (Lynn 2008; Lugtig 2014). However, experienced burden plays a role in so-called 'panel fatigue' (Laurie et al. 1999; Lemay 2010) and is assumed to accumulate over the course of panel participation (especially with frequent data collection), leading for some to the decision to drop out (Lipps 2009; Watson and Wooden 2009; Lemay 2010). Burden may also play a key part in experienced 'shocks' that can lead to dropout (Lemay 2010; Lugtig 2014; Kleinert et al. 2019), for example, when objective burden in a particular survey wave is greater due to design features such as questionnaire content, length, difficulty, or due to stress or frustration provoked by response tasks (see e.g., Dillman et al. 1993; Galesic and Bosnjak 2009; Marcus et al. 2007; Lynn 2014; Lugtig 2014; and Kleinert et al. 2019). Perceptions (and recall) of such features can vary by response mode (Couper et al. 2017; Gummer and Daikeler 2020), so it is likely that device-related differences may arise in Internet panel studies also.

Further evidence that burden plays a role in panel attrition comes from the fact that attrition is often highest among people for whom the level of burden is objectively greater (Lugtig 2014). This includes, for example, people with lower levels of education or cognitive skills who may find it harder to complete questionnaires (Loosveldt and Carton 2001; Freese and Branigan 2012) and people from ethnic minorities, who may experience increased burden due to language difficulties (Lipps 2009). However, it may also be due to divergent subjective experiences of burden. For example, people with extravert personalities appear to be more susceptible to boredom and panel fatigue (Lugtig 2014). More generally, respondents who gain less enjoyment from completing surveys or respondents who experienced difficulties at the previous survey wave are also more likely to drop out of panels (e.g., Hill and Willis 2001; Kalton et al. 1990; Laurie et al. 1999; Lepkrowski and Couper 2002; Loosveldt and Carton 2001; Olsen 2005; Lugtig 2014). Finally, longer interview times (Hill and Willis 2001) and poor response quality attributed to fatigue (Loosveldt et al. 2002) have also been found to be predictive of attrition. These findings reinforce the hypothesis that the risk of attrition in an online panel study will be greater for respondents completing questionnaires on mobile device browsers if participation is experienced as especially burdensome.

2.2. *Opportunities Offered by Survey Apps*

To mitigate response burden for mobile-web respondents, it is recommended to adapt web questionnaires designed for PCs (or other modes) to accommodate constraints imposed by mobile devices (Buskirk and Andrus 2014; Peytchev and Hill 2010; Antoun et al. 2017; Couper et al. 2017; Herzing 2019). Mobile-optimised survey design has been shown to reduce required effort for mobile device respondents, leading to lower breakoff rates (Mavletova and Couper 2015; Stapleton 2013). Optimised designs can also help to improve respondents' interest and enjoyment, found to be important in motivating (ongoing) participation in mobile web surveys (Bosnjak et al. 2010; Galesic 2006).

Guidelines for adapting web questionnaires for mobile devices apply equally across browser and app-based platforms. However, apps offer additional ways to optimise web surveys to mobile devices that can potentially reduce burden and enhance enjoyment for mobile respondents (Jacobsen and Kühne 2021), which should, in principle, make them even more appealing than conventional browser-based surveys. In an app, the questionnaire resides locally on the phone, so it imposes fewer demands for a persistent internet connection, which may reduce breakoff risk (Link et al. 2014). Questionnaires can more easily be administered to respondents in shorter modules (simultaneously or over time), potentially offering them greater control over the duration of participation in any given response session, which may also mitigate the likelihood of breakoff and perceptions of burden (Johnson et al. 2012; Toepoel and Lugtig 2018). Apps also offer the possibility to contextualise the timing of survey requests (e.g., triggered by location or event) so they are received when participation is more relevant or convenient (Jäckle et al. 2018; Kreuter et al. 2018).

Apps also offer opportunities to potentially reduce response burden by replacing self-report measures with alternative types of data capture, including visual data and passively collected data (e.g., GPS, browser or app logs, or other sensor data (Keusch et al. 2019; Revilla et al. 2019). For example, in a study by Jäckle et al. (2019), respondents were asked to download an app in order to scan receipts for purchases, allowing them to save time and effort compared to entering details manually. Incorporating multimodal data collection (e.g., capturing photos, videos and audio) alongside conventional measures may also help to increase respondent engagement by making participation in surveys more enjoyable and varied (Link et al. 2014). Apps should, therefore, improve mobile respondents' participation experience, engagement, and motivation to continue participating and to optimise the response process. Assessing these possible benefits is key to finding ways to better optimise web-based surveys for existing and future mobile respondents. If response burden is indeed lower, then app-based data collection may be especially suitable for mobile respondents in online panel studies; just as the benefits of using apps are maximised in longitudinal research designs (Lugtig 2021).

It is important to note that despite their potential advantages, studies to date have generally found low levels of stated and actual willingness to participate in surveys via mobile apps. Here again, factors relating to burden appear among hypothesised explanatory variables (Keusch et al. 2019; Wenz et al. 2019). For example, Wenz et al. (2019, 4) identify potentially burdensome task characteristics (see also Keusch et al. 2019; Revilla et al. 2019), including having to download and install the app to begin with, and whether the survey involves active or passive data collection. Consistent with this, willingness to respond via an app varies both as a function of survey design features as well as respondent characteristics, with those for whom burden to complete tasks is likely lowest responding at higher rates. For example, studies have found younger respondents, and more generally, those who are more familiar with and experienced using mobile devices and other apps, are more willing to respond via a survey app (Pinter 2015; Scherpenzeel 2017; Elevelt et al. 2019; Jäckle et al. 2019; Keusch et al. 2019; Mulder and de Bruijne 2019; Wenz et al. 2019; Lawes et al. 2021; Struminskaya et al. 2021).

This complicates the task of comparing experiences of burden across response devices and software, as selection effects influence who responds via different access options and can bias estimates of burden derived from those samples. It also implies a need to acknowledge that burden may be a more complex phenomenon in the context of mixed device web surveys than originally conceptualised (e.g., by Bradburn 1978; Haraldsen 2004) and operationalised by many researchers (see Yan et al. 2019; Galesic and Bosnjak 2009; Link et al. 2014). This means that efforts to optimise surveys for mobile devices that focus solely on improving questionnaire design features may be insufficient for guaranteeing positive response experiences, as other factors beyond researchers' control (e.g., respondent experience, device usability and environmental factors), may mean that device or software-related differences persist. Nevertheless, to the extent that burden for mobile respondents can be alleviated through optimal design, apps may well be effective at reducing some of the negative effects of mobile devices on survey participation, and by extension, help to improve data quality.

3. Methods

3.1. Data

The data come from a three-wave online panel survey called ‘Selects-Civique', conducted in Switzerland during the 2019 Swiss Federal Election campaign, alongside the Swiss

Election Studies ('Selects') (Tresch et al. 2020). The survey included an experimental design (illustrated in Figure 1) to compare a (mobile-optimised) browser-based design with an app-based design. As for the 'Selects' survey, the target population for Selects-Civique was Swiss adults (18 and over) with the right to vote in federal elections, but the sample was restricted to people resident in French-speaking cantons only (due to resource constraints). A sample of 2,175 individuals was drawn from the Federal Statistical Office's sampling frame based on population registers maintained by municipalities and randomly assigned to one of two treatment groups. Group 1 (n = 1,088) received a mailed invitation to participate in the panel via a web browser, while Group 2 (n = 1,087) received a mailed invitation to participate in the panel via a mobile application called 'Civique.org', a pre-existing data collection platform intended as a citizen science initiative to gather multimodal data relevant to local civic causes (www.civique.org – first developed in 2015 and updated over time at Idiap Research Institute, Switzerland by D. Gatica-Perez, J.-I. Biel, O. Bornet, P. Abbet,, and D. Santani, at Idiap Research Institute, Switzerland).

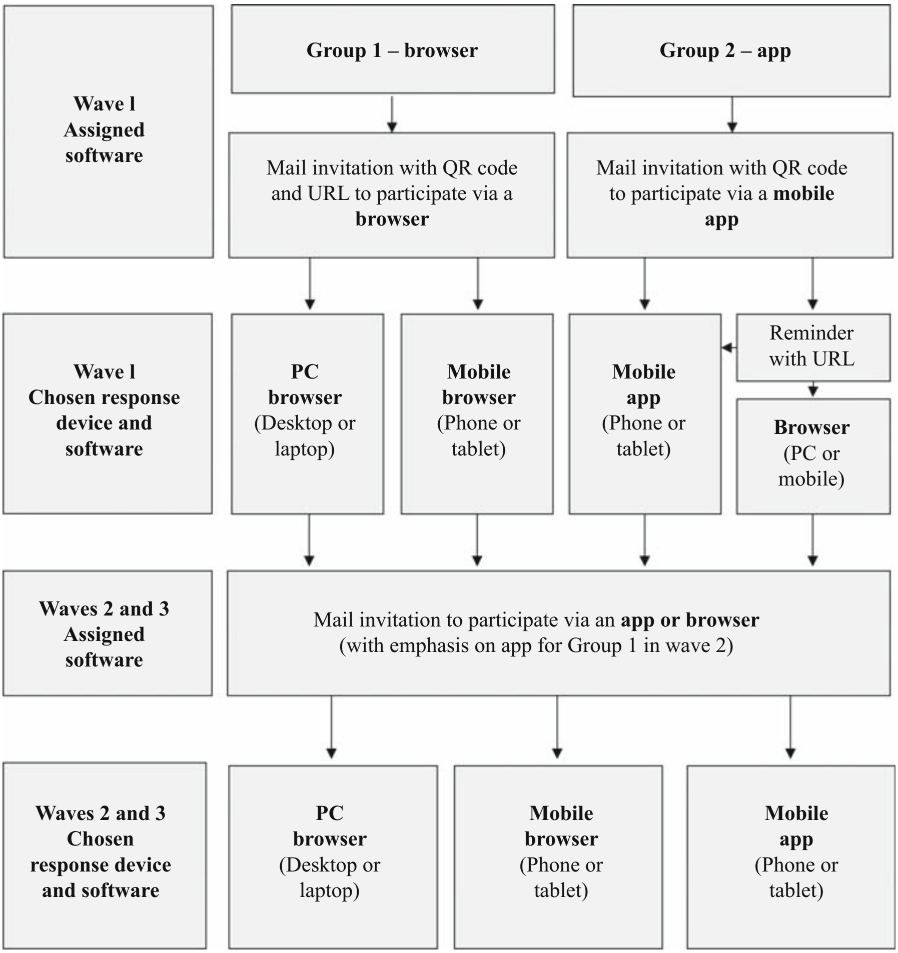

Fig. 1. Research design – Assigned software and chosen response device.

At wave 1 (fielded in May 2019), sample members received an unconditional incentive of USD 10. In both groups, the possibility to use a mobile device was emphasized in the invitation letter by displaying a Quick Response (QR) code, which directly forwarded respondents to the survey landing page (group 1) or the Civique app in either the Google Play or Mac App store depending on which operating system was detected (the app is available for Android and iOS mobile operating systems (but not for PCs)). To enable sample members without mobile devices to participate, the invitation also included a Uniform Resource Locator (URL) link to the survey landing page. Nonrespondents in group 2 were not informed about the browser version until the first reminder of wave 1. At wave 2 (fielded in October 2019), all wave 1 respondents from both groups 1 and 2 (irrespective of response device/ software used) were sent a postal invitation to participate via the app but were simultaneously offered the option of responding via a web browser. Participation via the app was emphasized visually and incentivised with a potentially higher conditional incentive (up to USD 20 for the app participation versus USD 10 for the browser).

Thus, within groups, respondents could self-select their preferred response device and software. At wave 1, group 1 respondents could *only* participate via a web browser but had the choice whether to respond on a PC (desktop or laptop) or on a mobile device (smartphone or tablet) – using the QR code or by typing in the URL. Group 2 respondents could participate via the app (on a smartphone or tablet, but not a PC) or via a web browser on a PC or mobile device (see Figure 1). At wave 2, all respondents could complete the survey using their preferred access option. Our analysis compares mobile respondents using a browser at wave 1 with (a) mobile respondents using the app, and (b) PC respondents using a browser and we address the self-selection into the device response groups using weighting (described further below).

All app users were required to sign a consent form within the app detailing the data privacy and data protection policy (also available to browser respondents on the study website).

3.2. Questionnaires

The questionnaires mainly included questions about political attitudes, voting behaviour, media consumption, social network usage, and socio-demographics. In both the browser and app groups we used an optimal design strategy (Hox et al. 2017) aimed at maximizing data quality in each of the access options. The browser version (programmed in Qualtrics) used a mobile-optimised design, adjusting to screen size, with only one item displayed per screen. The intended completion time was around 20 minutes, the content dictated in part by the design of the parallel “Selects” study.

The app version followed a modular design process that used the available technical features of the platform to improve user experience. This involved splitting the browser version of the questionnaire into thematic sections, each taking one to five minutes to complete. At wave 1, there were nine modules, which were presented in the same order as in the browser questionnaire but were all made available at once, so respondents were free to choose the order of completion. It was not possible to skip questions within modules in the app, but modules could be left out or abandoned. In the browser, respondents were required to respond to all applicable questions (in the intended, fixed order), but where appropriate, options were provided to allow respondents to withhold their answer.

3.3. Indicators of Response Burden

As response burden can be evaluated both objectively and subjectively, we draw on indicators of both dimensions (Read 2019) – see the Online Supplemental Material (Table S1) for an overview. To measure *subjective response burden*, we analyse respondents' evaluations of experienced burden in the wave 1 questionnaire, based on four items that were asked at the end of the questionnaire (in the last module in the app). The items included statements presented with a five-point fully labelled Agree/Disagree response scale, where 1 meant 'completely agree' and 5 meant 'completely disagree'. The statements were: (1) 'The questionnaire was interesting', (2) 'The length of the questionnaire was adequate', (3) 'The questions were comprehensible', and (4) 'Filling in the questionnaire presented no difficulty'. We assumed that if respondents agreed with the statements, their overall experience at wave 1 was positive and their experienced burden low. Item 1 is relevant to respondent motivation to participate (Groves et al. 2004) and continue participating in the panel (Galesic 2006). Items 2, 3 and 4 are more direct indicators of perceived response burden, relating to length, effort and respondent ability/competence (Read (2019), originally proposed by Bradburn (1978) and Haraldsen (2004)). The items are also akin to common usability metrics used in user experience research, relating to satisfaction, enjoyment, engagement, experienced burden and task success (Geisen and Bergstrom 2017).

After confirming the inter-item correlations were positive and significant, we created a composite measure of subjective burden based on respondents' mean scores for all four items for use in preliminary pairwise comparisons between the groups of interest. Cronbach's alpha for the scale was .71 (a principal components analysis extracting one factor with an eigenvalue greater than one supported this decision). However, two of the items (comprehensibility of the questions and difficulty filling in the questionnaire) had lower correlations with the other two and did not differentiate respondents or device groups well (and indeed, were negatively correlated with the second factor, which had an eigenvalue just below one – output available in the See Online Supplementary Material). For this reason, we also computed a mean score based only on the other two items (questionnaire length and interest) for use in the subsequent (mediation) analyses.

To measure *objective response burden*, we calculated wave 1 completion times based on module completion timings from the app (summing the time taken to complete all the modules), and screen-by-screen timing variables from Qualtrics, which were summed across the items according to how they were grouped in each app module, and then across all the modules. We report mean completion times in minutes, normalised based on +-2 standard deviations (SD) from the mean. Completion times for a total of 16 cases (nine for the app, seven for the browser) were two SD above the mean, and we substituted these with the average time taken by the remaining respondents (following Revilla and Ochoa 2015), as these were likely indicative of interruptions rather than necessarily due to slower pace.

The number of items that were applicable to all respondents was slightly lower in the app version (91 questions applicable to all respondents) than in the browser version for respondents with a smartphone (96 questions applicable to all), and slightly higher than for browser respondents without a smartphone (83 questions applicable to all). The differences were in a module of questions on smartphone use, in which certain items were deemed not to be relevant to those already responding via the app (a question about the smartphone's operating system; a question on smartphone skills; and questions on willingness to complete a questionnaire on a mobile phone; to, download a survey app and to share GPS location) or to those without a smartphone, who answered a subset of questions about activities completed on the Internet instead. In addition, the *actual* number of questions answered varied by respondent as a function of whether follow-up questions to filters were asked. For this reason, we also compare groups on the number of answers given in wave 1 as an indicator of objective burden. However, as this variable was not normally distributed, highly associated with the response device and correlated with the response duration variable, we excluded it from the subsequent multivariate analyses.

3.4. Analytic Approach

Based on the reviewed literature, we hypothesise that both response burden and dropout will be greatest for mobile-browser respondents, compared with mobile-app respondents and PC-browser respondents. Underpinning this hypothesis is the assumption that variation in wave 2 drop-out by device and software is mediated by variation in experienced response burden at wave 1, which we test through a mediation analysis. On this basis, the analysis had two aims. The first was to compare subjective and objective experienced response burden across respondents completing wave 1 on different devices and using different software (RQ1). The second was to assess whether response device/software used at wave 1 predicts non-participation at wave 2 (RQ2), and if so, whether and to what extent the effect of device on non-participation at wave 2 is mediated by response burden experienced at wave 1 (RQ3).

To address RQ1, we initially conduct pairwise comparisons using t-tests to test differences in means for the subjective and objective measures of response burden described above. Subsequently, we use a regression-based approach as part of the mediation analysis conducted to address the other research questions (described in the next section). We focus on comparing those using a mobile browser with those using a PC browser to provide some control for the software type (notwithstanding possible differences between browser providers), and those using a mobile browser and those using the mobile app to provide some control for the device type. However, for interested readers, we also present the comparison between PC and App respondents (though this confounds device and software). At wave 1, there were only 20 app respondents who used a tablet and 32 mobile browser respondents who used a tablet, and due to small achieved sample sizes overall, we decided to pool tablet and smartphone respondents in all our analyses. We recognise that, for some users, tablets may be used in similar ways to laptop PCs. Our approach emphasises the greater portability of tablets and their typically smaller screen sizes, as well as the fact that the app used was available for tablet, but not PC. Future studies with larger samples responding on tablets may reconsider this classification.

Because our comparisons of interest are confounded by selection effects on the samples responding using different devices/software, we used a propensity score weighting approach to try to balance the samples using auxiliary data available for all sample members, following general recommendations for addressing questions of causal inference in social research (Harder et al. 2010; Rosenbaum and Rubin 1983). We first computed an inverse probability weight to address selectivity due to nonresponse at the first wave. We then computed separate weights for the pairwise comparisons across devices to control for the differential probability of responding using one software/ device type compared to the other and combined these with the general nonresponse weight for each of the pairwise comparisons. For the remaining analyses, in which we used regression-based methods to test our mediation hypothesis, we use the general nonresponse weight on its own. Details of how the weights were computed are available in the Online Supplemental Material (Computation of Weights).

3.4.1. Testing for Mediation

To address RQ2 and RQ3 and test the hypothesis that the effect of the response device on wave 2 dropout is mediated by experienced response burden at wave 1, we followed procedures for mediation analysis (Hayes 2017) based on those proposed by Baron and Kenny (1986), which are used widely in the social sciences (VanderWeele 2016). Mediation analysis allows the researcher to establish the extent to which an independent variable (e.g., response device) influences a dependent variable (e.g., dropout at wave 2 of a panel) “through one or more *mediator* variables” (e.g., experienced response burden) (Hayes 2017). In other words, it helps to shed light on *how* one variable influences another variable and explicitly tests hypotheses relating to the possible mechanisms involved (Hayes 2017).

At step 1, we regressed the indicator for non-participation in wave 2 (coded 1 if the person did not participate in wave 2 (dropped out) and coded 0 if they did participate) on the wave 1 device indicators (dummy variables indicating those who responded on a PC and those who responded via the app versus those who responded on a mobile browser (reference category). At step 2, we regressed the mediator variables (the two-item subjective burden indicator (questionnaire length and interest) and the measure of completion time) on the device indicators, to assess the relation between device and response burden. As we used multiple indicators of response burden, we first tested the mediation hypothesis separately for the subjective and objective indicators (together with the control variables). Although using the app significantly predicted reduced completion times, in the presence of the other variables in the model, this variable was not significantly associated with non-participation at wave 2. We focused, therefore, on the two-item measure of subjective burden, keeping completion time as a control variable. Both the completion time (in minutes) and subjective burden variables (the mean of two five-point scales) were recoded to range from 0 to 1, where 0 represented the minimum score, and 1 represented the maximum level of experienced burden – that is, the longest completion time and strong disagreement that the questionnaire was interesting and the length was adequate. At step 3, we added the mediator to the model predicting non-participation at wave 2, to assess whether subjective burden significantly predicts dropout, and whether the relation between device and dropout remains significant when controlling for response burden.

The mediation analysis was conducted using the SPSS macro ‘PROCESS’ version 3.5 (see Hayes 2017). As the main outcome variable was binary, we used logistic regression analysis at steps 1 and 3, and Ordinary Least Squares (OLS) regression at step 2. The procedure produces estimates of the direct and indirect effects of the dependent variable (device) via the mediator on the logged odds scale. To test whether the indirect effect is statistically significant, the macro uses non-parametric bootstrapping to estimate standard errors and 95% confidence intervals (based on 5,000 bootstrapped samples). To address potential confounding of the assumptions underpinning the mediation hypothesis (see VanderWeele 2016, 19–21), we included a number of control covariates in the logistic regression analyses (shown in Table S2 and discussed in detail in part C of the Online Supplemental Material). The same set of covariates were included in all models (and as independent variables in the OLS regression predicting the mediator). As PROCESS cannot handle weighted data, we present the results of the mediation for the unweighted data only.

4. Results

Before presenting the results of the analyses addressing the three research questions, we first present details of participation rates and breakoffs by response device and software, which provide insight into the extent of self-selection into the response samples and some preliminary differences of interest between groups. These are shown in Table 1. At wave 1, a total of 687 (31.6% – AAPOR RR2) sample members participated in the survey, 366 (33.6%) in the group assigned to the browser condition (group 1) and 321(29.5%) in the group assigned to the app (group 2). The difference in response rates between the treatment groups was statistically significant ($\chi^2(1) = 4.25$; $p < .05$). Of those participating, 298 (43.4%) participated via a PC browser, 152 (22.1%) via a mobile browser, and 237 (34.5%) via the app. Only 358 respondents participated in wave 2 (52.1% of those participating at wave 1), of which 139 (38.8%) used a PC browser, 23 (6.4%) used a mobile browser and 196 (54.7%) used the app. Out of the 184 wave 1 browser respondents from group 1, 89 (48.4%) switched to the app in wave 2, whereas 81 (44.0%) responded via a browser again. Only 14 respondents from group 1 (7.6%) responded on a mobile browser at wave 2. From group 2, 174 (54.2%) respondents participated in wave 2, of which 109 (62.6%) participated via the app and the remainder via a browser (only ten (5.8%) on a mobile). The dropout rate for the 32 tablet users among mobile browser respondents was 56.3% compared with 54.8% for smartphone users; for the 20 tablet users in the app group, the dropout rate was 20% compared with 42.6% for smartphone users.

We define breakoffs based on the modular questionnaire design used in the app in two ways: (1) starting the survey and not completing the 9th and final module; and (2) starting the survey and not completing the last three modules. The final module included the questionnaire evaluation measures used as indicators of subjective burden, so failure to complete this module meant that the respondent was excluded from our analyses. Out of the 687 respondents, 621 had complete data for this module, and form the analytic sample for the subsequent analyses. Not completing the final three modules meant that the participant failed to complete the module of socio-demographic questions (module 7), rendering their preceding answers less usable (module 8 was about motivations for participating, so may have been considered less pertinent to the main survey topic and

Table 1. Participation and break-off rates by treatment group, device and software type.

	Total		PC Browser		Mobile Browser		App	
	n	%	*n*	%	*n*	%	*n*	%
Wave 1								
Group 1 (browser) (N^4=1,088)	366	33.6	232	63.4	134	36.6	-	-
Group 2 (app) (N=1,087)	321	29.5	66	20.6	18	5.6	237	73.8
Total (N=2,175)	687	31.6	298	43.4	152	22.1	237	34.5
Break-off[1]	65	3.0	22	7.4	16	10.5	27	11.4
Break-off[2]	58	2.7	19	6.4	16	10.5	23	9.7
Prefer not to do W2	153	22.3	68	22.8	42	27.6	43	18.1
Analytic sample[3]	621	28.6	275	44.3	136	21.9	210	33.8
Wave 2								
Group 1 (N=366)	184	50.3	81	44.0	14	7.6	89	48.4
Group 2 (N=321)	174	54.2	55	31.6	10	5.8	109	62.6
W1 PC (N=298)	164	55.0	113	68.9	3	1.8	48	29.3
W1 Mobile (N=152)	62	40.8	6	9.7	14	9.7	42	67.7
W1 App (N=237)	132	55.7	20	8.4	6	4.5	106	80.3
Total (N=687)	358	52.1	139	38.8	23	6.4	196	54.7
Break-off[1]	17	4.7	1	0.7	5	21.7	11	5.6
Break-off[2]	12	3.4	1	0.7	4	17.4	7	3.6

Notes: [1]Break-off = started survey but did not complete final module (questionnaire evaluation). [2]Break-off = started survey but did not complete the last three modules. [3]Analytic sample includes all respondents responding to questionnaire evaluation measures. [4]N= Total sample contacted at waves 1 and 2; n = sample responding.

hence, provoked participants to quit). According to these definitions, breakoff rates were higher for respondents on mobile devices (whether using a browser or app) than for PC respondents. Breakoff rate (1) at wave 1 was 7.4% for participants using PCs, 10.5% for participants using a mobile and 11.4% for participants using the app; and rate (2) was 6.4% PC participants, 10.5% for mobile participants and 9.7% for participants using the app. Differences between samples responding on different devices were not significant, however. Respondents who did complete the second-to-last module were asked how they would like to be contacted at wave 2, with the option to say they would prefer not to participate again. A higher proportion of respondents using the mobile browser (27.6%) selected this option than of respondents on a PC (22.8%; difference not statistically significant) or using the app (18.1%; $\chi^2(1) = 4.88$; $p < .05$).

4.1. Device/Software Effects on Experienced Response Burden (RQ1)

Table 2 shows the adjusted means (weighting for nonresponse at wave 1) for the subjective burden indicators by device/ software and the results of independent samples t-tests for each of the pairwise comparisons. Participants' evaluations of the wave 1 questionnaire were generally positive across all response device groups. However, statistical differences were observed between the main comparison groups of interest (mobile browser versus pc and mobile browser versus app) on some measures. App respondents were significantly more likely than mobile respondents to agree the questionnaire was interesting (mean [M] = 2.12, SD = .80 vs. M = 2.36, SD = .89; t-value[t] = 2.70 (Levene's Test for Equality of Variances was statistically significant so equal variances are not assumed), degrees of freedom [d.f.] = 265.1, p-value [p] <.001), and that the length of the questionnaire was adequate (M = 1.76, SD = .82 vs. M = 2.55, SD = 1.10; $t = 7.32$ (equal variances not assumed), d.f. = 230.9, p < .001). App respondents also had significantly lower mean scores on the composite measures, indicating overall less experienced burden (see column 5, Table 2). Respondents who used a PC browser at wave 1 were more likely to agree that the length of the questionnaire was adequate compared with respondents who used a mobile browser (M = 2.28, SD = 1.02 vs. M = 2.55, SD = 1.10; $t = 2.61$ (equal variances not assumed), d.f. = 249.6, p < .05). These samples also differed significantly on the composite score combining the questionnaire interest and length measures. There were no differences on the other subjective burden measures (see column 4, Table 2). Subjective burden measures were also significantly more positive for app users than for PC users, except for the indicator for difficulty (see column 6, Table 2).

Compared with mobile browser respondents, app respondents had significantly shorter completion times (M = 21.43, SD = 8.75 vs. M = 23.61, SD = 10.95; $t = 2.32$ (equal variances not assumed), d.f. = 238.36, p < .05), despite answering more questions in total (M = 106.01, SD = 2.90 vs. M = 98.52, SD = 2.51; t = -25.68 (equal variances not assumed), d.f. = 310.93, p < .001). There was no significant difference between the two browser groups in terms of completion times. However, the mean number of questions answered varied significantly between respondents using a PC browser and mobile respondents (M = 100.59, SD = 4.45 vs. M = 98.52, SD = 2.51; t = -4.93 (equal variances not assumed), d.f. = 390.40., p < .001). App respondents also differed from PC

Table 2. Subjective and objective response burden measures – adjusted means by wave 1 device type and pair-wise comparisons.

	Means[1]			Pairwise contrasts[2]		
	(1) PC Browser (n = 275)	(2) Mobile Browser (n = 136)	(3) App (n = 210)	(4) Mobile-PC Browser	(5) Mobile Browser -App	(6) PC Browser - App
	Mean (SD)	Mean (SD)	Mean (SD)	(2)-(1) *p*	(2)-(3) *p*	(3)-(1) *p*
Subjective burden indicators:						
The questionnaire was interesting (1-5[3])	2.27 (0.96)	2.36 (0.89)	2.12 (0.80)	0.09	0.24**	0.15†
The length of the questionnaire was adequate (1-5)	2.28 (1.02)	2.55 (1.10)	1.76 (0.82)	0.27*	0.79***	0.52***
The questions were comprehensible (1-5)	1.43 (0.65)	1.38 (0.53)	1.31 (0.53)	-0.05	0.07	0.12*
Filling in the questionnaire presented no difficulty (1-5)	1.49 (0.76)	1.42 (0.68)	1.45 (0.81)	-0.07	-0.03	0.04
Mean score on all items	1.87 (0.66)	1.93 (0.54)	1.66 (0.55)	0.06	0.27***	0.21***
Mean of interesting and length	2.28 (0.88)	2.46 (0.87)	1.94 (0.71)	0.18*	0.52***	0.34***
Objective burden indicators:						
No. of questions answered	100.59 (4.45)	98.52 (2.51)	106.01 (2.90)	-2.07***	-7.49***	-5.42***
Completion duration in minutes	25.03 (10.36)	23.61 (10.95)	21.43 (8.75)	-1.42	2.18*	3.60**

Notes: [1]Means adjusted for wave 1 nonresponse. [2]Pairwise contrasts based on Independent Samples T-Tests using propensity-score weights to adjust for observed differences in sample composition for each pairwise comparison combined with wave 1 nonresponse weights; [3]Agree-disagree response scales where 1 = completely agree and 5 = do not agree at all. SD = standard deviation; p = p-value: † $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

respondents on the objective burden indicators, with faster completion times, despite answering more questions.

To summarise, we find significant variation in experienced response burden as a function of response device and software (RQ1), even when adjusting for observed differences in the device/software comparison groups due to differential selection error at wave 1.

4.2. *Effects of Wave 1 Response Device on Wave 2 Participation (RQ2)*

Parameter coefficients for the logistic regression model predicting drop-out at wave 2, estimated at step 1 of the mediation analysis are shown in Table 3 (Model 1). The overall fit of the model was significant ($\chi^2(21) = 52.836$, $p < .001$; Hosmer and Lemeshow's test was non-significant, also indicating good model fit), and, based on Nagelkerke's pseudo R^2, the model accounted for around 11% of the variation in the probability of wave 2 nonresponse. Controlling for the socio-demographic and the other control variables, both app and PC respondents were significantly less likely to drop out of the study after wave 1 compared to respondents who used a mobile browser (see column 1, Table 3). The odds ratio for responding on an app (exponentiated beta coefficient [Exp(*B*)] = .516, confidence interval $[CI]_{.95} = [.267, .997]$) indicates that the odds of dropping out of the study after wave 1 were around 48% lower for app respondents than they were for mobile respondents. The odds ratio for responding on a PC (Exp(*B*) = .630, $CI_{.95} = [.394, 1.01]$) indicates that the odds of dropping out for PC respondents were around 37% lower than they were for mobile respondents. Thus, we indeed find differences in willingness to participate at wave 2 of an online panel study as a function of the response device/software used at wave 1, with mobile browser respondents dropping out at a higher rate than PC or app respondents.

4.3. *Mediating Effects of Experienced Response Burden on Non-Participation at Wave 2 (RQ3)*

To assess whether experienced response burden mediates effects of the response device on the decision not to participate in wave 2, we first estimated the parameter coefficients of the Ordinary Least Squares (OLS) regression equation predicting subjective burden with the device indicators and the control variables from model 1 (see Table 3, Model 2). The overall fit of the model was good ($R^2 = .328$; $F_{(21,599)} = 13.95$; $p < .001$). Responding using the app (compared to a mobile browser) significantly reduced experienced subjective burden ($B = -.101$, $SE = .028$, $p < .001$), while responding on a PC (compared to a mobile browser) reduced burden, but not significantly ($B = -.035$, $SE = .020$, $p = .078$). These results support the findings of the t-tests that experienced response burden varies as a function of response device and software (RQ1), even when controlling for other respondent characteristics.

Independent of response device, respondents who were interested in politics also reported significantly less burden ($B = -.056$, $SE = .016$, $p < .001$), as did respondents who were motivated to participate in wave 1 by the possibility to contribute to science ($B = -.361$, $SE = .030$, $p < .001$). Respondents who reported using the Internet less than once a day reported higher levels of burden ($B = .040$, $SE = .019$, $p < .05$), as did

Table 3. Coefficients from logistic regression models predicting non-participation at wave 2(1 and 3) and OLS regression predicting subjective burden (2).

	(1) Non-participation at Wave 2				(2) Subjective Burden			(3) Non-participation at Wave 2 (+ subjective burden)			
	B	SE	*p*	Exp(B)	*B*	SE	*p*	*B*	SE	*p*	Exp(B)
Device (ref. Mobile):											
PC browser	-.463	.239	.053†	.630	-.035	.020	.078†	-.415	.242	.086†	.661
App	-.662	.336	.049*	.516	-.101	.028	.000***	-.516	.342	.132	.597
Subjective burden	-	-		-	-	-		1.493	.500	.003**	4.449
Response duration	.250	.571	.661	1.284	.029	.048	.544	.208	.576	.718	1.231
Respondent sex: female	.126	.187	.502	1.134	.002	.016	.899	.129	.189	.497	1.137
Respondent age (ref. 18–30)											
Aged 31–55	.588	.293	.045*	1.800	.013	.024	.594	.576	.295	.051†	1.779
Aged 56 and over	.590	.378	.119	1.804	-.060	.031	.055†	.687	.383	.073†	1.988
Marital status (ref. never married):											
Married	-.323	.267	.226	.724	.056	.022	.014*	-.408	.271	.132	.665
Divorced, widowed	.005	.309	.987	1.005	.068	.026	.009**	-.096	.314	.761	.909
Household size (ref. 1 member):											
2 members	.501	.284	.078†	1.651	.001	.024	.966	.506	.286	.077†	1.659
3 or more members	.345	.290	.234	1.412	-.002	.024	.935	.355	.292	.224	1.427
Lives in a rural area	-.060	.190	.753	.942	.033	.016	.039*	-.109	.193	.572	.897
Main activity (ref. in F/T paid work):[1]											
In P/T paid work	-.428	.242	.076†	.652	-.023	.020	.245	-.403	.244	.098†	.668
Student / apprentice/ in training	.086	.355	.809	1.090	-.011	.030	.711	.104	.357	.771	1.109
Retired/ unemployed/ home-maker	-.440	.254	.083	.644	.002	.021	.927	-.455	.257	.077†	.634
Education (ref. Secondary level or lower):[2]											
Tertiary level	-.473	.179	.008**	.623	.006	.015	.668	-.494	.181	.006**	.610
Interested in politics	-.141	.184	.446	.869	-.056	.016	.000***	-.057	.188	.761	.944
Motivated to contribute to science	-1.372	.357	.000***	.254	-.361	.030	.000***	-.869	.394	.027*	.419
Uses Internet less than once a day	-.404	.227	.075†	.668	.040	.019	.034*	-.477	.232	.039*	.620

Table 3. Continued

	(1) Non-participation at Wave 2				(2) Subjective Burden			(3) Non-participation at Wave 2 (+ subjective burden)			
	B	SE	*p*	Exp(B)	*B*	SE	*p*	*B*	SE	*p*	Exp(B)
Doesn't use smartphone to access Internet	-.269	.292	.357	.764	-.029	.024	.220	-.224	.295	.447	.799
Does use tablet to access Internet	-.099	.189	.601	.906	.013	.016	.409	-.123	.192	.520	.884
Invited to app at wave 1 (Group 2)	.101	.265	.702	1.107	-.014	.022	.519	.122	.269	.648	1.130
Constant	.949	.473	.045*	2.582	.539	.039	.000***	.171	.540	.752	1.186
N	621				621				621		
Nagelkerke R^2/ R^2	.109				.328				.127		

Notes: [1] F/T = full-time; P/T = part-time. [2]19 cases selecting the 'other education' category were added to the reference category (primary). ref. = reference category; N = cases analysed; B = Unstandardised beta coefficients (on logged-odds scale for logistic regressions [1] and [3]); SE = standard error; Exp(B) = Exponentiated beta coefficient; *p* = p-value: † p <0.1, *p < 0.05, ** p < 0.01, *** p < 0.001. Data weighted to adjust for nonresponse at wave 1.

respondents living in a rural area (B = .033, SE = .016, $p < .05$). Finally, compared to respondents who were single (never married) married (B = .056, SE = .022, $p < .05$) and divorced (B = .068, SE = .026, $p < .05$) respondents reported slightly increased burden. None of the other independent variables were significant predictors of subjective burden.

Given that responding on a PC browser compared with a mobile browser was not significantly predictive of increased burden (only at the 10% level), we focused the mediation analysis on the effect of using the app compared to a mobile browser (keeping responding on a PC in the model as a control covariate). Table 3 (Model 3) shows the log-odds and odds ratios for the full model predicting non-participation at wave 2. Subjective burden has a statistically significant and positive direct effect on the outcome variable (Exp(B) = 4.449, $CI_{.95}$ = [1.671, 11.842]). For each unit increase in reported subjective burden, the odds of not responding at wave 2 of the study increase by around 345%. The direct effect of the app indicator is reduced in the presence of the subjective burden measure and is no longer statistically significant at the 5% level (adjusted Exp(B) = .597, $CI_{.95}$ = [.305, 1.167]; unadjusted Exp(B) = .556, $CI_{.95}$ = .278, 1.110]). The effect of using the app compared to a mobile browser on non-participation at wave 2 is, thus, at least partially mediated by subjective burden. The path model (with the (unadjusted) direct and indirect effects presented in a log-odds metric) is shown in Figure 2 (note that these estimates are based on the full model including all the control variables shown in Table 3). The (unadjusted) indirect effect of the app via burden on drop-out is negative and statistically significant (Exp(B) = .855, $CI_{.95}$ = [-.353, -.042]; effect tested using non-parametric bootstrapping – see Figure 2). In other words, the reduction in experienced subjective burden resulting from using the app instead of a mobile browser results in a reduction in the odds of dropping out of the study after wave one of around 15%. This is illustrated in Figure 3, which shows the probability of dropping out for app and mobile browser respondents across levels of response burden. These findings confirm the hypothesis that experienced response burden at wave 1 mediates device effects on willingness to participate at wave 2 of an online panel study (RQ3).

Among the control variables, it is noteworthy that the size of the coefficient for being motivated to participate by the possibility to contribute to science (highly significant in both models 1 and 2) is also reduced by the inclusion of the mediator, and its effect on the outcome is also partially mediated via subjective burden (Indirect Effect [IE] Exp(B) = .571, $CI_{.95}$ = [-1.02, -.187].

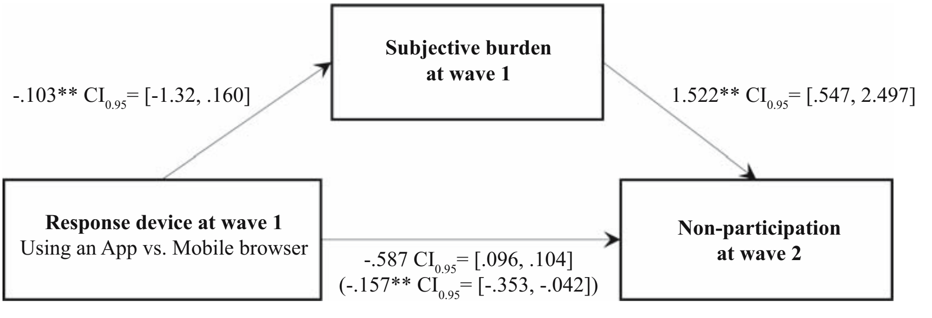

Fig. 2. Regression coefficients and 95% confidence intervals for the relationship between response device at wave 1 and non-participation at wave 2 (direct and indirect effects (in parentheses) shown in a log odds metric).

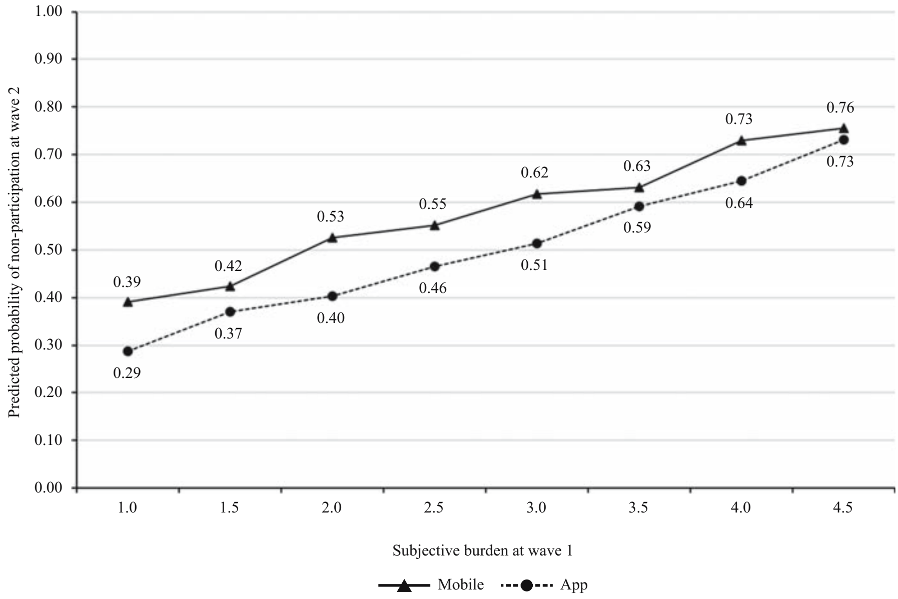

Fig. 3. Predicted probabilities of non-participation at wave 2 for respondents using a mobile browser or an app across levels of experienced subjective burden.

5. Discussion

In this article we reported the results of a study using data from an online panel survey with an embedded experiment comparing an app-based design and a mobile-adapted web browser design. Previous research has found that participants in online surveys using browsers on mobile devices are more likely to break off (see Mavletova and Couper 2015 for a review), due, in part, to greater experienced response burden (Allum et al. 2018) or reduced enjoyment (Bosnjak et al. 2010). In the context of panel surveys, survey experience and increased response burden influence the likelihood of attrition (Gummer and Daikeler 2020; Lugtig 2014), meaning that panel respondents using mobile devices may be at greater risk of dropping out. While studies investigating correlates of hypothetical willingness to participate in mobile data collection have also identified burden-related factors as sources of resistance (e.g., Read 2019; Wenz et al. 2019), few studies have explicitly tested the mediating effect of burden on *actual* participation – especially in the context of a probability-based general population survey. For these reasons, we assessed whether and how respondents' experienced burden using a given device (PC versus smartphone or tablet) or software (browser versus app) in the first wave of a panel affected their response propensity in the second panel wave.

Our first research question concerned the extent to which experienced response burden varied as a function of device and software. In the bivariate analyses, we observed a number of differences in subjective and objective burden between respondents using different devices, with higher levels of burden for mobile browser respondents compared with app and PC respondents. In the mediation analysis, responding using an app (compared to responding on a mobile browser) was associated with significantly reduced subjective burden, even when controlling for completion times, sociodemographic differences in the composition of the response samples, and other control variables known to correlate with participation in surveys/ attrition, that could plausibly also relate to experienced burden (such as topic interest). Responding on a PC, however, was not significantly associated with burden (or at least, was so to a lesser extent).

Our second research question concerned the extent to which willingness to participate at wave 2 varied as a function of the wave 1 response device. Respondents who used a mobile browser at wave 1 of the study were significantly more likely not to participate in wave 2 than respondents who used the app or a PC browser. Our third research question was whether experienced response burden at wave 1 mediates device effects on willingness to continue participating at wave 2. When the measure of subjective burden was added to the model predicting non-participation at wave 2, it was strongly and positively significant, while the effect of responding using the app was no longer so. The mediation analysis confirmed a significant, negative indirect effect of the app on drop-out in wave 2 via subjective burden. This implies that responding on an app can motivate ongoing panel participation due to the app's capacity to reduce response burden.

Because responding on a PC was not a significant predictor of subjective burden, we did not test the mediation hypothesis for the comparison between PC and mobile browser respondents. However, the effect of responding on a PC on subjective burden was 'significant' at the 10% level. We noted effects of this size because small samples combined with low response rates in this study likely affected the possibility to detect statistical relationships of interest (with a larger sample, we might have observed a similar mitigating effect, as has been concluded by Mavletova and Couper 2015). The focus on the positive effects for respondents of completing the survey using the app, however, is of greater contemporary interest (both theoretical and practical) because of the many opportunities app-based surveys offer researchers in terms of new data collection capabilities (Jäckle et al. 2018; Link et al. 2014).

We addressed potential confounding of the assumptions underpinning the mediation model (VanderWeele 2016) by including several control covariates in the regression analyses. Of these potential confounds, two variables were found to be statistically significant *negative* predictors of non-participation at wave 2 of the study: having completed a tertiary-level qualification and being motivated to participate in the survey by the possibility to contribute to science. Level of education was not significantly associated with experienced burden. However, respondents who were motivated by the possibility to contribute to science reported less burden; they were more likely to evaluate the questionnaire length as adequate and the questionnaire content as interesting, and were subsequently, less likely to drop-out of the panel at wave 2 (in line with findings of Keusch et al. 2019). Another potential confounding variable was topic interest. However, although respondents who were more interested in politics were also more likely to agree the questionnaire was interesting, political interest was not associated with non-participation at wave 2. Sensitivity analyses can be used to assess in more detail whether confounding violates assumptions of mediation analysis (VanderWeele 2015, 2016), but given these findings, we did not investigate this further.

5.1. Limitations

The need to address potential confounds is particularly important given that we did not have strict control over exposure to the treatment of interest – that is, the response devices selected by respondents. While software type (whether the respondent used the app or a browser) was randomly assigned, the choice of device (PC or mobile) was based on respondent preferences. This means that differences in the characteristics of respondents using different devices at wave 1 could account for both experiences of burden and response propensity at the second wave of the survey. Besides the inclusion of control variables in the regression analyses already described, we also used propensity score weights to control for self-selection into the response samples. It should be noted that the effectiveness of this method depends on which variables are used to estimate response probabilities (Roberts et al. 2020). We were able to make use of administrative data from the sampling frame but found relatively few differences between response samples on the sociodemographic variables analysed (see Online Supplemental Material, Table S2). As a result, we cannot rule out the possibility that unobserved selection errors may be partially responsible for the findings reported here (Lugtig and Toepoel 2016).

Another feature of the study design that may have influenced our findings is the fact that respondents assigned to the browser-based design were invited to switch to the app at wave 2 (and encouraged to do so through the offer of a higher incentive), while the group 2 respondents who used the app at wave 1 could continue using the app. In other words, effects of the response device used at wave 1 are potentially confounded with the effect of the mode switch for group 1 respondents, which has been shown in other studies to lead to drop out (e.g., Sakshaug and Kreuter 2011; Sakshaug et al. 2010; Tourangeau et al. 2013). Nevertheless, respondents in both groups were informed that they could continue using the browser to participate if they wished to do so and presumably those with strong preferences for using a browser (or no alternative) would have continued to use it in subsequent panel waves if motivated to do so.

The difference found in experiences of using the app versus using a mobile browser implies that the underlying causes of burden for mobile respondents may be less about user, device or environmental influences, but rather, the freedom and flexibility (and novelty, perhaps) offered by the software. In the app, it was possible to use a modular questionnaire design where all modules were available at once, giving respondents complete control over when, where, and how to complete them. This had a positive influence on respondents' perceptions of the adequacy of the length of the questionnaire and of how interesting it was, and subsequently, on their probability of participation at wave 2. This is partially consistent with the findings of Johnson et al. (2012), who found positive benefits of a modular design in an app on breakoffs (though in our study, the wave 1 breakoff rate was still higher in the app than for browser respondents). Toepoel and Lugtig (2018) reported greater dropout *with* a modular design, but one in which modules were administered over the course of several days. These findings suggest there may be benefits of making all modules available simultaneously.

As we are not able to disentangle the 'pure' effect of using an app from the effect of the modular questionnaire, we recommend that future research investigate the advantages and potential disadvantages of alternative modular designs, both within and across different response devices, in both app- and browser-based surveys. This is also needed to inform understanding of how modular questionnaire designs are used by respondents and of any possible negative effects they may have (e.g., on measurement quality due to differential context effects linked to question order – see Dillman et al. (2014) for an overview).

6. Conclusion

This study lends further support to a well-established and fairly consistent literature about the negative effects of perceived and experienced burden on willingness to participate in surveys. It also confirms the assumptions of previous research into device differences in survey outcomes that they appear to be mediated by differences in experienced burden. Experienced burden in the study reported here was greatest for respondents who completed the first wave using their mobile browser, and this had a negative impact on willingness to participate at the subsequent panel wave. The novel contribution, however, is the finding that using a mobile app (with a modular questionnaire design) effectively reduces response burden compared to using a mobile browser (with a conventional questionnaire) and positively influences willingness to continue participating in the panel.

App-based data collection platforms offer a range of potential benefits to researchers, and combined with a panel design, a potentially cost-efficient one. The present study suggests there may be considerable advantages for respondents also. If it is indeed the case that apps reduce response burden then our findings would imply a broader utility for them as a data collection mode in general population studies, beyond the more 'niche' uses for which they have been deployed in surveys to date (Link et al. 2014). The use of apps in panel designs offers other advantages also, including the possibility to make use of the new measurement tools they offer and the opportunity to build study identity and loyalty through a purpose-designed interface and infrastructure for panel management (Link et al. 2014). More importantly, however, if apps offer more effective mobile optimisation solutions than browsers, they present a better alternative for reaching the growing number of mobile-dependent Internet users and survey respondents.

Benefits of responding via an app could be emphasised to potentially positive effect in efforts to motivate willingness to participate in app-based studies, which has been shown elsewhere to present an important barrier (so far) to the successful integration of apps in survey research (e.g., Keusch et al. 2019; Mulder and de Bruijne 2019; Revilla et al. 2019; Wenz et al. 2019). Similarly, the finding that being motivated to contribute to science improves both the survey experience and respondent engagement may offer further clues as to how to optimise the framing of invitations to participate in app-based surveys. Given the importance of motivation, combined with perceptions and prior experiences of burden in participation decisions, new research should test ways of making these benefits more salient in survey requests and monitoring and evaluating both in the future development of app-based survey methodology.

7. References

Allum, N., F. Conrad, and A. Wenz. 2018. "Consequences of Mid-Stream Mode-Switching in a Panel Survey." *Survey Research Methods* 12: 43–58. DOI: https://doi.org/10.18148/srm/2018.v12i1.6779.

Antoun, C. 2015. Who are the Internet Users, Mobile Internet Users, and Mobile-Mostly Internet Users?: Demographic Differences Across Internet-Use Subgroups in the U.S." In *Mobile Research Methods: Opportunities and Challenges of Mobile Research Methodologies* by D. Toninelli, R. Pinter, and P. Pedraza: 99–117. Ubiquity Press. DOI: http://dx.doi.org/10.5334/bar.g.

Antoun, C., and A. Cernat. 2020. “Factors Affecting Completion Times: A Comparative Analysis of Smartphone and PC Web Surveys.” *Social Science Computer Review* 38: 477–489. DOI: https://doi.org/10.1177/0894439318823703.

Antoun, C., F.G. Conrad, M.P. Couper, and T.B. West. 2019. “Simultaneous Estimation of Multiple Sources of Error in a Smartphone-Based Survey.” *Journal of Survey Statistics and Methodology* 7: 93–117. DOI: https://doi.org/10.1093/jssam/smy002.

Antoun, C., M.P. Couper, and F.G. Conrad. 2017. “Effects of Mobile versus PC Web on Survey Response Quality.” *Public Opinion Quarterly* 81: 280–306. DOI: https://doi.org/10.1093/poq/nfw088.

Antoun, C., J. Katz, J. Argueta, and L. Wang. 2018. “Design Heuristics for Effective Smartphone Questionnaires.” *Social Science Computer Review* 36: 557–574. DOI: https://doi.org/10.1177/0894439317727072.

Baron, R.M. and D.A. Kenny. 1986. “The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations.” *Journal of Personality and Social Psychology* 51: 1173–1182. DOI: https://doi.org/10.1037/0022-3514.51.6.1173.

Bosnjak, M., G. Metzger, and L. Gräf. 2010. “Understanding the Willingness to Participate in Mobile Surveys: Exploring the Role of Utilitarian, Affective, Hedonic, Social, Self-Expressive, and Trust-Related Factors.” *Social Science Computer Review* 28: 350–370. DOI: https://doi.org/10.1177/0894439309353395.

Bradburn, N. 1978. *Respondent burden.* Health Survey Research Methods, DHEW Publication No. 79-3207: 35–40. Washington, DC: U.S. Department of Health, Education, and Welfare. Available at: http://www.asasrms.org/Proceedings/papers/1978_007.pdf (accessed March 2022).

Buskirk, T.D., and C. Andrus. 2012. “Online Surveys Aren’t Just for Computers Anymore! Exploring Potential Mode Effects between Smartphone vs. Computer-based Online Surveys”. In JSM Proceedings, Survey Research Methods Section. Alexandria, VA: American Statistical Association: 5678–5691. Available at: http://www.asasrms.org/Proceedings/y2012f.html (accessed March 2022).

Buskirk, T.D. and C. Andrus. 2014. “Making Mobile Browser Surveys Smarter: Results from a Randomized Experiment Comparing Online Surveys Completed via Computer or Smartphone.” *Field Methods* 26: 322–342. DOI: https://doi.org/10.1177/1525822X14526146.

Callegaro, M. 2010. “Do You Know Which Device Your Respondent Has Used to Take Your Online Survey?” *Survey Practice* 3: 1–12. DOI: https://doi.org/10.29115/sp-2010-0028.

Callegaro, M. 2013. “From Mixed-Mode to Multiple Devices: Web Surveys, Smartphone Surveys and Apps: Has the Respondent Gone Ahead of Us in Answering Surveys?”, *International Journal of Market Research* 55: 317–320. DOI: https://doi.org/10.2501/IJMR-2013-026.

Callegaro, M., K. Manfreda, and V. Vehovar. 2015. *Web Survey Methodology*. London, UK: SAGE.

Couper, M.P., C. Antoun, and A. Mavletova. 2017. “Mobile Web Surveys.” In *Total Survey Error in Practice* edited by P.P. Biemer, E.D. de Leeuw, S. Eckman, B.

Edwards, F. Kreuter, L.E. Lyberg, N.C. Tucker, and B.T. West: 133–154. Hoboken, NJ: Wiley. DOI: https://doi.org/10.1002/9781119041702.ch7.

Couper, M.P. and G.J. Peterson. 2017. "Why Do Web Surveys Take Longer on Smartphones?" *Social Science Computer Review* 35: 357–377. DOI: https://doi.org/10.1177/0894439316629932.

De Bruijne, M. and A. Wijnant. 2014a. "Mobile Response in Web Panels." *Social Science Computer Review* 32: 728–742. DOI: https://doi.org/10.1177/0894439314525918.

De Bruijne, M. and A. Wijnant, A. 2014b. "Improving Response Rates and Questionnaire Design for Mobile Web Surveys." *Public Opinion Quarterly* 78: 951–962. DOI: https://doi.org/10.1093/poq/nfu046.

De Leeuw, E.D. and P. Lugtig. 2014. *Dropouts in Longitudinal Surveys.* Wiley StatsRef: Statistics Reference Online. John Wiley & Sons, Ltd. DOI: 10.1002/9781118445112.stat06661.pub2.

Dillman, D.A., M.D. Sinclair, and J.R. Clark. 1993. "Effects of Questionnaire Length, Respondent-friendly Design, and a Difficult Question on Response Rates for Occupant-addressed Census Mail Surveys." *Public Opinion Quarterly* 57: 289-304. DOI: 10.1086/269376.

Dillman, D.A., J.D. Smyth, and L.M. Christian. 2014. *Internet, Mail, and Mixed-mode Surveys: The Tailored Design Method*. Hoboken, NJ: John Wiley & Sons.

Elevelt, A., P. Lugtig, and V. Toepoel. 2019. "Doing a Time Use Survey on Smartphones Only: What Factors Predict Nonresponse at Different Stages of the Survey Process?" *Survey Research Methods* 13: 195–213. DOI: https://doi.org/10.18148/srm/2019.v13i2.7385.

Freese, J. and A. Branigan. 2012. "Cognitive Skills and Survey Nonresponse — Evidence from Two Longitudinal Studies in the United States." *EurAmerica* 42: 221–247. Available at: https://www.ea.sinica.edu.tw/eu_file/134009378914.pdf.

Galesic, M. 2006. "Dropouts on the Web: Effects of Interest and Burden Experienced during an Online Survey." *Journal of Official Statistics* 22: 313–328. Available at: https://www.scb.se/contentassets/ca21efb41fee47d293bbee5bf7be7fb3/dropouts-on-the-web-effects-of-interest-and-burden-experienced-during-an-online-survey.pdf

Galesic, M., and M. Bosnjak. 2009. "Effects of Questionnaire Length on Participation and Indicators of Response Quality in a Web Survey." *Public Opinion Quarterly* 73: 349–360. DOI: https://doi.org/10.1093/poq/nfp031.

Geisen, E., and J.R. Bergstrom. 2017. *Usability Testing for Survey Research*. Burlington, MA: Morgan Kaufmann.

Groves, R., S. Presser, and S. Dipko. 2004. "The Role of Topic Interest in Survey Participation Decisions." *Public Opinion Quarterly* 68: 2–31. DOI: https://doi.org/10.1093/poq/nfh002.

Guidry, K.R. 2012. "Response Quality and Demographic Characteristics of Respondents Using a Mobile Device on a Web-Based Survey?" Presentation at the American Association for Public Opinion Research's 67th Annual Conference. May 18, Orlando, FL, U.S.A. Available at: https://www.aapor.org/AAPOR_Main/media/AnnualMeeting-Proceedings/2012/01_KRGuidry_F1_Mobile-Device-Respondents.pdf (accessed March 2022).

Gummer, T., and J. Daikeler. 2020. "A Note on How Prior Survey Experience with Self-Administered Panel Surveys Affects Attrition in Different Modes." *Social Science Computer Review* 38: 490–498. DOI: https://doi.org/10.1177/0894439318816986.

Haraldsen, G. 2004. "Searching for Response Burdens in Focus Groups with Business Respondents." In Proceedings of the QUEST 2003: Proceedings of the 4th Conference on Questionnaire Evaluation Standards, 21-23 October 2003, edited by P. Prüfer, M. Rexroth, and F.J.J. Fowler: 13–123. Mannheim: Zentrum für Umfragen, Methoden und Analysen -ZUMA-. Available at: https://nbn-resolving.org/urn:nbn:de:0168-ssoar-49200-6 (accessed August 2021).

Harder, V.S., E.A. Stuart, and J.C. Anthony. 2010. "Propensity Score Techniques and the Assessment of Measured Covariate Balance to Test Causal Associations in Psychological Research." *Psychological Methods* 15: 234–249. DOI: https://doi.org/10.1037/a0019623.

Hayes, A.F. 2017. *Introduction to Mediation, Moderation, and Conditional Process Analysis.* New York, NY: The Guilford Press.

Herzing, J.M.E. 2019. *Mobile Web Surveys*. FORS Guide No. 01, Version 1.0. Lausanne: Swiss Centre of Expertise in the Social Sciences FORS. DOI: https://doi.org/10.24449/FG-2019-00001.

Hill, D.H. and R.J. Willis. 2001. "Reducing Panel Attrition. A Search for Effective Policy Instruments." *Journal of Human Resources* 26: 416–438. DOI: https://doi.org/10.2307/3069625.

Hoogendoorn, A.W., and D. Sikkel. 1998. "Response Burden and Panel Attrition." *Journal of Official Statistics* 14: 189–205. Available at: https://www.scb.se/contentassets/ca21efb41fee47d293bbee5bf7be7fb3/response-burden-and-panel-attrition.pdf

Hox, J., E.D. de Leeuw, and T. Klausch. 2017. "Mixed Mode Research: Issues in Design and Analysis." In *Total survey error in practice* edited by P.P. Biemer, E.D. de Leeuw, S. Eckman, B. Edwards, F. Kreuter, L.E. Lyberg, N.C. Tucker, and B.T. West: 511–530. Hoboken, NJ: Wiley. DOI: https://doi.org/10.1002/9781119041702.ch23.

Jäckle, A., J. Burton, M.P. Couper, and C. Lessof. 2019. "Participation in a Mobile App Survey to Collect Expenditure Data as Part of a Large-Scale Probability Household Panel: Coverage and Participation Rates and Biases." *Survey Research Methods* 13: 23–44. DOI: https://doi.org/10.18148/srm/2019.v1i1.7297

Jäckle, A., A. Gaia, and M. Benzeval. 2018. *The Use of New Technologies to Measure Socioeconomic and Environmental Concepts in Longitudinal Studies*. CLOSER Resource report. London, UK: UCL, Closer, Institute of Education. Available at: https://www.closer.ac.uk/wp-content/uploads/CLOSER-resource-The-use-of-new-technology-to-measure-socio-economic-and-environmental-concepts.pdf (accessed March 2022).

Jacobsen, J. and S. Kühne. 2021. Using a Mobile App When Surveying Highly Mobile Populations: Panel Attrition, Consent, and Interviewer Effects in a Survey of Refugees. *Social Science Computer Review* 39: 721–743. DOI: https://doi.org/10.1177/0894439320985250.

Johnson, A., F. Kelly, and S. Stevens. 2012. "Modular Survey Design for Mobile Devices." Presentation at the CASRO Online Conference 2012, March 7–8, Las Vegas, NE, U.S.A. Available at: https://c.ymcdn.com/sites/www.casro.org/resource/collection/E270CC91-6B72-4C37-BCC0-5503/CBB66C55/Paper_-_Frank_Kelly_and_Alex_-Johnson_-_Lightspeed_Research_and_Kantar_Operations.pdf.

Kalton, G., J. Lepkowski, G.E. Montanari, and D. Maligalig. 1990. "Characteristics of Second Wave Nonrespondents in a Panel Survey." In Proceedings of the Survey Research Methods Section: American Statistical Association: 462–467. In *JSM Proceedings,* Survery Research Methods Section. Alexandria, VA, U.S.A. Available at: http://www.asasrms.org/Proceedings/y1990f.html (accessed August 2021).

Keusch, F., B. Struminskaya, C. Antoun, M.P. Couper, and F. Kreuter. 2019. "Willingness to Participate in Passive Mobile Data Collection." *Public Opinion Quarterly* 83: 210–235. DOI: https://doi.org/10.1093/poq/nfz007.

Keusch, F. and T. Yan. 2017. "Web versus Mobile Web: An Experimental Study of Device Effects and Self-Selection Effects." *Social Science Computer Review* 35: 751–769. DOI: https://doi.org/10.1177/0894439316675566.

Klausch, T., J. Hox, and B. Schouten. 2015. "Selection Error in Single- and Mixed Mode Surveys of the Dutch General Population." *Journal of the Royal Statistical Society Series A* 178: 945–961. DOI: https://doi.org/10.1111/rssa.12102.

Kleinert, C., B. Christoph, and M. Ruland. 2019. "Experimental Evidence on Immediate and Long-term Consequences of Test-induced Respondent Burden for Panel Attrition." *Sociological Methods and Research*: 1–32. DOI: https://doi.org/10.1177/0049124119826145.

Kreuter, F., G.-C. Haas, F. Keusch, S. Bähr, and M. Trappmann. 2018. "Collecting Survey and Smartphone Sensor Data with an App: Opportunities and Challenges Around Privacy and Informed Consent." *Social Science Computer Review* 38: 533–549. DOI: https://doi.org/10.1177/0894439318816389.

Laurie, H. 2008. "Handbook of Longitudinal Research: Design, Measurement, and Analysis." In *Minimizing Panel Attrition* edited by S. Menard: 167–184. Elsevier.

Laurie, H., R. Smith, and L. Scott. 1999. ''Strategies for Reducing Nonresponse in a Longitudinal Panel Survey.'' *Journal of Official Statistics* 15: 269–282.

Lawes, M., C. Hetschko, J.W. Sakshaug, and S. Griessemer. 2021. "Contact Modes and Participation in App-based Smartphone Surveys: Evidence from a Large-scale Experiment." *Social Science Computer Review*. DOI: https://doi.org/10.1177/0894439321993832.

Lee, H., S. Kim, M. Couper, and Y. Woo. 2019. "Experimental Comparison of PC Web, Smartphone Web, and Telephone Surveys in the New Technology Era." *Social Science Computer Review* 37: 234–247. DOI: https://doi.org/10.1177/0894439318756867.

Lemay, M. 2010. ''Understanding the Mechanism of Panel Attrition.'' Unpublished Doctoral thesis, Doctor of Philosophy, University of Maryland, College Park, MD.

Lepkowski, J.M. and M.P. Couper. 2002. "Nonresponse in the Second Wave of Longitudinal Household Surveys." In *Survey Nonresponse* edited by R.M. Groves, D.A. Dillman, J.E. Eltinge, and R.J.A. Little: 259–272. New York: Wiley.

Link, M.W., J. Murphy, M.F. Schober, T.D. Buskirk, J. Hunter Childs, and C. Langer Tesfaye. 2014. "Mobile Technologies for Conducting, Augmenting and Potentially Replacing Surveys." *Public Opinion Quarterly* 78: 779–787. DOI: https://doi.org/10.1093/poq/nfu054.

Lipps, O. 2009. *Attrition of Households and Individuals in Panel Surveys.* SOEP- Papers 164. Berlin: DIW. Available at: http://www.diw.de/documents/publikationen/73/diw_01.c.96125.de/diw_sp0164.pdf (accessed March 2022).

Loosveldt, G. and A. Carton. 2001. "An Empirical Test of a Limited Model for Panel Refusals." *International Journal of Public Opinion Research* 13:173–185. DOI: https://doi.org/10.1093/ijpor/13.2.173

Loosveldt, G., J. Pickery, and J. Billiet. 2002. "Item Nonresponse as a Predictor of Unit Nonresponse in a Panel Survey." *Journal of Official Statistics* 18: 545–557. Available at: https://www.scb.se/contentassets/ca21efb41fee47d293bbee5bf7be7fb3/item-nonresponse-as-a-predictor-of-unit-nonresponse-in-a-panel-survey.pdf.

Lugtig, P. 2014. "Panel Attrition: Separating Stayers, Fast Attriters, Gradual Attriters, and Lurkers." *Sociological Methods & Research* 43: 699–723. DOI: https://doi.org/10.1177/0049124113520305

Lugtig, P. 2021. "What Panel Surveys and Smartphone-App Studies can Learn from Each Other." Presentation at the 9th Conference of the European Survey Research Association, Online, July 9. Available at: https://www.europeansurveyresearch.org/conferences/program2021?sess = 29#218 (accessed March 2022).

Lugtig, P. and V. Toepoel. 2016. "The Use of PCs, Smartphones, and Tablets in a Probability-Based Panel Survey: Effects on Survey Measurement Error." *Social Science Computer Review*, 34: 78–94. DOI: https://doi.org/10.1177/0894439315574248.

Lynn, P. 2008. "The Problem of Nonresponse." In *International Handbook of Survey Methodology* edited by E.D. de Leeuw, J.J. Hox, and D.A. Dillman: 35–55. New York: Psychology Press.

Lynn, P. 2014. "Longer Interviews May Not Affect Subsequent Survey Participation Propensity." *Public Opinion Quarterly* 78: 500–509. DOI: http://dx.doi.org/10.1093/poq/nfu015.

Marcus, B., M. Bosnjak, S. Lindner, S. Pilischenko, and A. Schütz. 2007. "Compensating for Low Topic Interest and Long Surveys. A Field Experiment on Nonresponse in Web Surveys." *Social Science Computer Review* 25: 372–383. DOI: https://doi.org/10.1177/0894439307297606.

Maslovskaya, O., G.B. Durrant, P.W.F. Smith, T. Hanson, and A. Villar. 2019. "What are the Characteristics of Respondents Using Different Devices in Mixed-Device Online Surveys? Evidence from Six UK Surveys." *International Statistical Review* 87: 326–346. DOI: https://doi.org/10.1111/insr.12311.

Mavletova, A. 2013. "Data Quality in PC and Mobile Web Surveys." *Social Science Computer Review* 31: 725–743. DOI: https://doi.org/10.1177/0894439313485201.

Mavletova, A. and M.P. Couper. 2013. "Sensitive Topics in PC Web and Mobile Web Surveys: Is there a Difference?" *Survey Research Methods* 7: 191–205. DOI: https://doi.org/10.18148/srm/2013.v7i3.5458.

Mavletova, A. and M.P. Couper. 2015. "A Meta-Analysis of Breakoff Rates." In *Mobile Research Methods: Opportunities and Challenges of Mobile Research Methodologies* edited by D. Toninelli, R. Pinter, and P. Pedraza: 81–98. Ubiquity Press. DOI: https://doi.org/10.5334/bar.f.

Mavletova, A., M.P. Couper, and D. Lebedev. 2018. “Grid and Item-by-Item Formats in PC and Mobile Web Surveys.” *Social Science Computer Review* 36: 647–668. DOI: https://doi.org/10.1177/0894439317735307.

Mulder, J. and M. de Bruijne. 2019. “Willingness of Online Respondents to Participate in Alternative Modes of Data Collection.” *Survey Practice* 12: 1–11. DOI: https://doi.org/10.29115/SP-2019-0001.

Office of Management and Budget. 2006. *Standards and Guidelines for Statistical Surveys*. Executive Office of the President of the United States. Available at: https://unstats.un.org/unsd/dnss/docs-nqaf/USA_standards_stat_surveys.pdf (accessed August 2021).

Olsen, R.J. 2005. “The Problem of Respondent Attrition: Survey Methodology is the Key.” *Monthly Labor Review* 128: 63–70. Available at: http://www.jstor.org/stable/23804052 (accessed March 2022).

Peytchev, A. 2009. “Survey Breakoff.” *Public Opinion Quarterly* 73: 74–97. DOI: https://doi.org/10.1093/poq/nfp014.

Peytchev, A. and C.A. Hill. 2010. “Experiments in Mobile Web Survey Design: Similarities to Other Modes and Unique Considerations.” *Social Science Computer Review* 28: 319–335. DOI: https://doi.org/10.1177/0894439309353037.

Pinter, R. 2015. “Willingness of Online Access Panel Members to Participate in Smartphone Application-Based Research.” In *Mobile Research Methods*, edited by D. Toninelli, R. Pinter, and P. de Pedraza: 141–156. London: Ubiquity Press.

Read, B. 2019. “Respondent Burden in a Mobile App: Evidence from a Shopping Receipt Scanning Study.” *Survey Research Methods* 13: 45–71. DOI: https://doi.org/10.18148/srm/2019.v1i1.7379.

Revilla, M., M.P. Couper, and C. Ochoa. 2019. “Willingness of Online Panelists to Perform Additional Tasks.” *Methods, Data, Analyses* 13: 223–252. DOI: https://doi.org/10.12758/mda.2018.01.

Revilla, M. and C. Ochoa. 2015. “What are the Links in a Web Survey among Response Time, Quality, and Auto-Evaluation of the Efforts Done?” *Social Science Computer Review* 33: 97–114. DOI: https://doi.org/10.1177/0894439314531214.

Revilla, M., C. Ochoa, and G. Loewe. 2017. “Using Passive Data from a Meter to Complement Survey Data in Order to Study Online Behavior.” *Social Science Computer Review* 35: 521–536. DOI: https://doi.org/10.1177/0894439316638457.

Revilla, M., D. Toninelli, and C. Ochoa. 2016. “PCs versus Smartphones in Answering Web Surveys: Does the Device Make a Difference?” *Survey Practice* 9: 1–6. DOI: https://doi.org/10.29115/sp-2016-0021.

Roberts, C., Vandenplas, C., and J. Herzing. 2020. “A Validation of R-indicators as a Measure of the Risk of Bias using Data from a Non-response Follow-up Survey.” *Journal of Official Statistics* 36: 675-701. DOI: https://doi.org/10.2478/jos-2020-0034.

Rosenbaum, P.R., and D.B. Rubin. 1983. “The Central Role of the Propensity Score in Observational Studies for Causal Effects.” *Biometrika* 70: 41–55. DOI: https://doi.org/10.1093/biomet/70.1.41.

Sakshaug, J.W., and F. Kreuter. 2011. “Using Paradata and Other Auxiliary Data to Examine Mode Switch Nonresponse in a “Recruit-and-Switch” Telephone Survey.” *Journal of Official Statistics* 27: 339–357. Available at: https://www.scb.se/contentassets/ca21efb41fee47d293bbee5bf7be7fb3/using-paradata-and-other-auxiliary-data-to-examine-mode-switch-nonresponse-in-a-34recruit-and-switch34-telephone-survey.pdf.

Sakshaug, J.W., T. Yan, and R. Tourangeau. 2010. "Nonresponse Error, Measurement Error, and Mode of Data Collection: Tradeoffs in a Multi-Mode Survey of Sensitive and Non-Sensitive Items." *Public Opinion Quarterly* 74: 907–933. DOI: https://doi.org/10.1093/poq/nfq057.

Scherpenzeel, A. 2017. "Mixing Online Panel Data Collection with Innovative Methods." In *Methodische Probleme von Mixed-Mode-Ansätzen in der Umfrageforschung* [Methodological Problems of Mixed-Mode Approaches in Survey Research], edited by S. Eifler and F. Faulbaum: 27–49. Wiesbaden: Springer.

Stapleton, C.E. 2013. "The Smart(phone) Way to Collect Survey Data." *Survey Practice* 6: 1–7. DOI: https://doi.org/10.29115/sp-2013-0011.

Struminskaya, B., K. Weyandt, and M. Bosnjak. 2015. "The Effects of Questionnaire Completion using Mobile Devices on Data Quality. Evidence from a Probability-Based General Population Panel." *Methods, Data, Analyses* 9: 261–292. DOI: https://doi.org/10.12758/mda.2015.014.

Struminskaya, B., Toepoel, V., Lutgtig, P., Haan, M., Luiten, A., and B. Schouten. 2021. "Understanding willingness to share smartphone-sensor data." *Public Opinion Quarterly*. DOI: https://doi.org/10.1093/poq/nfaa044.

Toepoel, V. and P. Lugtig. 2015. "Online Surveys are Mixed-Device Surveys. Issues Associated with the Use of Different (Mobile) Devices in Web Surveys." *Methods, Data, Analyses* 9: 155–162. DOI: https://doi.org/10.12758/mda.2015.009.

Toepoel, V. and P. Lugtig. 2018. "Modularization in an Era of Mobile Web: Investigating the Effects of Cutting a Survey into Smaller Pieces on Data Quality." *Social Science Computer Review*. DOI: https://doi.org/10.1177/0894439318784882.

Tourangeau, R., F.G. Conrad, and M.P. Couper. 2013. *The Science of Web Surveys*. Oxford, GB: Oxford University Press.

Tresch, A., L. Lauener, L. Bernhard, and L. Scaperrotta. 2020. *Selects: Panel Survey* (waves 1–3) – 2019 [Dataset]. Distributed by FORS, Lausanne, 2020. www.selects.ch. DOI: https://doi.org/10.23662/FORS-DS-1184-1.

VanderWeele, T.J. 2015. *Explanation in Causal Inference: Methods for Mediation and Interaction*. Oxford, UK: Oxford University Press.

VanderWeele, T.J. 2016. "Mediation Analysis: A Practitioner's Guide." *Annual Review of Public Health* 37: 17–32. DOI: https://doi.org/0.1146/annurev-publhealth-032315-021402.

Watson, N. and M. Wooden. 2009. ''Identifying Factors Affecting Longitudinal Survey Response.'' In *Methodology of Longitudinal Surveys* edited by P. Lynn: 157–182. Chichester, England: John Wiley.

Wells, T., J.T. Bailey, and M.W Link. 2013. "Filling the Void: Gaining a Better Understanding of Tablet-Based Surveys." *Survey Practice* 6: 1–9. DOI: https://doi.org/10.29115/sp-2013-0002.

Wenz, A. 2021. "Do Distractions During Web Survey Completion Affect Data Quality? Findings from a Laboratory Experiment." *Social Science Computer Review* 3: 148–161. DOI: https://doi.org/10.1177/0894439319851503.

Wenz, A., A. Jäckle, and M.P. Couper. 2019. "Willingness to Use Mobile Technologies for Data Collection in a Probability Household Panel." *Survey Research Methods* 13: 1–22. DOI: https://doi.org/10.18148/srm/2019.v13i1.7298.

Yan, T., S. Fricker, S., and S. Tsai. 2019. "Response Burden: What Predicts It and Who is Burdened Out?" In *Advances in Questionnaire Design, Development, Evaluation and Testing* edited by P. Beatty, D. Collins, L. Kaye, J.L. Padilla, G. Willis, and A. Wilmot, 193-212. Hoboken, NJ: John Wiley & Sons. DOI: https://doi.org/10.1002/9781119263685.ch8.

Received December 2020
Revised August 2021
Accepted March 2022