

Do Question Topic and Placement Shape Survey Breakoff Rates?

Survey Methods: Insights from the Field, Volume 2, issue 1 (2024)

Luke Plutowski, Vanderbilt University, USA

Elizabeth J. Zechmeister, Vanderbilt University, USA

How to cite this article : Plutowski, L. & Zechmeister, E. J. (2024). Do Question Topic and Placement Shape Survey Breakoff Rates? *Survey Methods: Insights from the Field*. Retrieved from <https://surveyinsights.org/?p=19071> The data used in this article is available for reuse from <http://data.aussda.at/dataverse/smif> at AUSSDA - The Austrian Social Science Data Archive. The data is published under the Creative Commons Attribution-ShareAlike 4.0 International License and can be cited as: Wilson, Carole; Plutowski, Luke; Zechmeister, Elisabeth, J. 2024, "Replication Data for: Do Question Topic and Placement Shape Survey Breakoff Rates? (OA edition)", <https://doi.org/10.11587/MMOPTD>, AUSSDA.

DOI : 10.13094/SMIF-2024-00005

Copyright : © the authors 2024. This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0)

Abstract : Survey breakoffs, or deliberate early terminations, are quite common and have the potential to reduce data collection efficiency and compromise data quality. Prior research has established that survey topics can influence participation rates, and conventional wisdom among practitioners is that more interesting questions should be placed at the beginning of questionnaires in order to increase retention. Building on these ideas, we test the extent to which the content and placement of questions in a survey jointly influence breakoff behavior through an experiment embedded in an April 2020 telephone survey in Haiti. Respondents were randomly assigned to answer questions about the COVID-19 pandemic at the beginning or near the end of the questionnaire. Those who heard the questions had only a slightly lower breakoff rate, but the treatment effect widened for the group that is comparatively more concerned about the coronavirus. Our findings show that - at least in times of crisis - adjusting survey content and structure to place timely topics upfront can modestly increase completion rates.

Introduction

Surveys have become a dominant method of research in many social scientific fields (Rossi et al. 2013) and are a key mechanism for inserting public opinion into democratic governance (Berinsky 2017). Unfortunately, many surveys suffer from high rates of breakoff - that is, instances in which respondents prematurely and deliberately stop their participation. Breakoffs can be permanent (i.e., the respondent never returns to the interview; Peytchev 2009) or temporary (i.e., the respondent leaves but returns to complete the questionnaire or reenters the panel at a later time; McGonagle 2013). Regardless of type, breakoff rates are often non-trivial in magnitude. A meta-analysis of web surveys found breakoff rates ranging from 0.4% to 30.9% (Mavletova and Couper 2015). A major telephone survey (the Panel Study of Income Dynamics) reported a temporary breakoff rate of 23% (McGonagle 2013). Others report a terminal breakoff rate for web surveys at 34% (Manfreda, Batagelj, Vehovar 2002).

We focus on terminal breakoffs in which individuals abandon the survey after it begins. Survey researchers aim to minimize these types of breakoffs for two main reasons: efficiency and quality. With respect to efficiency, canceling and replacing incomplete surveys incurs project costs (Keeter et al. 2016). With respect to quality, these breakoffs can produce biases if they do not occur randomly in the sample (Roßmann, Steinbrecher, and Blumenstiel 2015). While methodological studies of unit nonresponse abound, scholarship has been less interested in the issue of breakoff, so much so that breakoff rates are not usually reported (Peytchev 2009; Schaeffer and Dykema 2011). Further, the limited literature on breakoffs focuses primarily on web surveys rather than telephone or face-to-face interviewing (but see McGonagle 2013).

Some research has shown that survey topic and respondent interest influence response rates, breakoff rates, and other types of engagement (Galesic 2006; Groves, Presser, and Dipko 2004; Krosnick and Presser 2010; McGonagle 2013; Shropshire, Hawdon, and Witte 2009). Extending work in this domain, we test the effect of question topic and placement in generating engagement with the survey. Survey professionals often advise to place the most interesting and relevant questions at the beginning of a survey (see recommendations on question order from Pew Research 2023 and Qualtrics 2023). Yet to our knowledge there has been little to no experimental research that has systematically considered how question topic and question location in the instrument jointly shape breakoff behavior. We theorize that that topic-induced motivation could stem from any of three non-rival mechanisms. First, interesting questions pique attentiveness to the survey. Second, questions about salient issues make the survey seem worthwhile. Third, questions on “important” topics foster bonding between interviewer and interviewee, boosting the latter’s cooperation with the survey.

A challenge to assessing how breakoff rates are shaped by question topic and placement is finding topics that are relevant to broad swathes of the population. The onset of the COVID-19 pandemic in 2020 provided one such opportunity. We included a small set of questions about a highly salient and nationally important issue – the COVID-19 pandemic – on a phone survey conducted April-June 2020 about democratic governance in Haiti. Our core test consists of an experiment embedded within the survey: individuals were randomly assigned to answer questions related to the pandemic at the start or toward the end of the survey. We then measure the breakoff rate.

We find, in this case, some limited evidence that placing salient topics at the start of the survey can minimize breakoff behavior. As expected, the effect size is correlated with levels of interest in the topic: those who were more concerned about COVID were less likely to break off after hearing questions about it. This study contributes to literature on question topic and respondent engagement by offering an experimental test and a quantifiable measurement of the efficiency gained by starting the survey on a salient issue of the day. Adding to scholarship about topic-induced motivation, the results suggest that topic may be influential in the decision to discontinue the survey just as it is for the decision to participate in the first place (e.g., Groves, Presser, and Dipko 2004), though perhaps to a lesser degree. Based on the modest yet noticeable findings, we argue that, *ceteris paribus*, survey researchers should begin questionnaires with more salient topics in order to reduce breakoffs; yet, given only marginal gains, it is not essential to do so if it comes at the cost of interrupting the flow of the questionnaire or if no issue is particularly and broadly salient.

Topic-Induced Motivation and Survey Breakoff

Worldwide fixation on the COVID-19 pandemic accelerated in early 2020, reaching a first peak around April and remaining elevated for the duration of the year (Alshaabi et al. 2020). As the virus spread

worldwide, some survey practitioners reported that individuals were more willing to participate in interviews when the survey topic was related to the pandemic (e.g., Ambitho et al. 2020).

These anecdotes align with conventional wisdom among practitioners that questionnaires ought to begin with interesting questions so as to maintain respondent attention. In their guide to writing survey questions, Pew Research states that “it is often helpful to begin the survey with simple questions that respondents will find interesting and engaging” (Pew Research Center 2023). The textbook *Marketing Research and Information Systems*, prepared by the Food and Agriculture Organization of the United Nations, mentions that “Opening questions that are easy to answer and which are not perceived as being ‘threatening’, and/or are perceived as being interesting, can greatly assist in gaining the respondent’s involvement in the survey and help to establish a rapport” (Crawford 1997). Some even suggest adding “ringer” or “throwaway” questions about the hot topics of the day to boost interest in the survey (Qualtrics 2023).

Existing scholarship offers theoretical backing for the idea that the salience of the pandemic would shape willingness to engage in surveys. Previous work has found that the extent to which the topic of the survey is personally relevant and interesting influences response rates (Holland and Christian 2009; Keusch 2013; Krosnick and Presser 2010; Martin 1994; Van Kenhove, Wijnen and De Wulf 2002). The leverage-saliency theory (Groves, Singer, and Corning 2000; Groves, Presser, and Dipko 2004) posits that the outcome of each survey request is influenced by multiple attributes, including, among other factors, the survey topic.

We extend this scholarship by examining the role that topic plays in *breakoffs*, not unit response rates. The decision to participate in a survey and the decision to terminate it early are theoretically and analytically distinct. The latter decision is conditional on the former, so the populations of interest are different. Furthermore, breakoffs may be influenced by a whole host of variables that are unobserved prior to the beginning of the questionnaire, such as question wordings, cognitive load, order effects, interviewer dynamics, and so on. Finally, much of the literature linking survey topic and response rates assumes that the survey has a singular, unified topic that can be succinctly described in the introductory text; for many surveys (e.g., omnibus questionnaires), the themes are often not revealed until after the respondent agrees to participate.

We add to research on topic interest by considering how question topic and placement (location within the survey), jointly, can shape dynamics related to survey breakoff rates. To the degree that placing interesting and relevant questions at the start of the survey generates greater engagement, we posit this may occur through any of three non-rival mechanisms. First, respondents may experience a bump in interest when they engage with a personally relevant topic. Relatable questions should create an initial spike in interest, boosting their willingness to continue. Second, beginning with questions that are highly relevant could convince respondents that the broader research effort is worthwhile. Other researchers have found that belief in the importance of scientific studies predicts a lower breakoff rate (Roßmann, Blumenstiel, and Steinbrecher 2015). Third, beginning a questionnaire by acknowledging a highly salient issue may help establish a rapport between interviewer and respondent. Past research has linked rapport to survey engagement, non-response, and willingness to disclose sensitive attitudes (Garbarski, Schaeffer, and Dykema 2016; Sun, Conrad, and Kreuter 2020; Tu and Liao 2007). Questions that acknowledge important local issues could foster interviewer-interviewee bonding, which ought to boost respondent cooperation. Conversely, it may come across as insensitive or untimely to discuss seemingly irrelevant matters during times of crisis.

Prior studies on breakoffs and respondent interest are instructive though incomplete. For example, researchers have shown that some features of survey instrument design that affect respondent interest are associated with lower likelihood of breakoff (Galesic 2006; McGonagle 2013), though these studies focus not on question content or saliency but rather structural factors like types of questions, questionnaire length, and module introductions. Peytchev (2009) suggests that older respondents are less likely to break off in surveys that are mainly about health-related issues due to the topic's greater relevance for that age cohort, though this mechanism is not explored empirically. Two other studies suggest that interest in the survey topic (e.g., politics, ecological conservation) predicts lower breakoff tendencies in web surveys about those issues (Roßmann, Blumenstiel, and Steinbrecher 2015; Shropshire, Hawdon, and Witte 2009). This is consistent with the proposed theory, but we extend their work by examining multi-topic surveys and by experimentally testing whether question topic and placement combine to elicit lower breakoff rates.

Our study design leveraged the salience of the COVID-19 pandemic, a situation that was of interest, relevance, and importance to individuals around the world in 2020. Based upon observations from survey practice and research, we hypothesized the following about the relationship between questions about the pandemic and respondent engagement:

- **H1:** *Respondents who receive questions about the coronavirus first are less likely to break off than those who first receive questions about other topics.*

We also hypothesize that concern about the pandemic issue moderates the treatment effect. Those who are unconcerned about COVID-19 should have no reaction to questions about it; or perhaps may even be turned off by the mention of the issue, making them more likely to break off. To more precisely test the core of the argument about topic-induced motivation, we assess H1 not only for the full sample, but also for a subset of the sample that excludes those who express little concern about the coronavirus problem. We consider the more definitive test of the theoretical framework to be captured by this hypothesis:

- **H2:** *Respondents who believe the pandemic is a serious problem are less likely to breakoff when asked about the coronavirus first (versus asked about other questions first).*

Data and Methods

Survey Information and Questionnaire Design

Our core test is based on a national cellphone survey of adults (ages 18+) that was fielded in Haiti from April to June of 2020. The probabilistic sample was drawn from random digit dialing of active cell phone numbers, supplemented by frequency matching to realize census-derived targets on region, gender, and age cohorts. The survey touched on a variety of issues related to democratic governance (though the study information script only mentioned “the situation in Haiti”).

The selection of Haiti as a case was determined by survey objectives unrelated to this study. However, we consider the use of Haiti as a laboratory to be an additional novelty of this study, as methodological research rarely gathers data from developing countries, where best practices for survey research may differ from those in the United States and Western Europe due to differing cultural norms, languages, political institutions, and/or experiences with survey research. What is more relevant is the timing of the survey vis-à-vis the pandemic. When the survey began, in April, the COVID-19 pandemic was just

beginning to take root in Haiti, and it spread quickly in May and June (Rouzier, Liautaud, and Deschamps 2020). This unfortunate situation allows for a useful test of the theory as the virus became an overarching political and personal concern for Haitians over time, though it was not universally seen as the most important issue (thus allowing for variation in level of concern, necessary for our test of H2).

Our experiment design consisted of random assignment to one of two conditions. In the first, the COVID-First Condition, respondents were asked a set of 10 questions related to the pandemic at the start of the questionnaire. In the second, the COVID-Late Condition, respondents were asked those 10 questions toward the end of the survey. Lupu and Zechmeister (2021) has shown that this experiment has substantive effects (priming individuals to think about the pandemic influences certain democratic attitudes), which affirms individuals have the potential to be affected by the module. The present study aims to detect the effect of the treatment on respondent behavior, namely, whether the COVID-First design elevates interest and motivation relative to other topics (the COVID-Late group began with questions about the economy).

Figure 1 displays the structure of the questionnaire. After answering seven eligibility questions (e.g., age, citizenship) and agreeing to participate, respondents receive either the COVID module (10 questions) and then about 35 substantive survey items (exact number depends on branching) which we call the “core”, or vice versa. The core consists of modules on the following topics: economic situation, government services, interpersonal trust, democracy, crime, television, trust in institutions, voting, corruption, health / medical services, political interest and knowledge, and welfare. All respondents then answer an end block of questions concerning demographic characteristics (e.g., level of education), sampling information (e.g., number of cell phones used by respondent), and a battery of items about water access and related issues. Our analyses consider only data within the COVID module, the core, and the end block. An alternative approach would be to consider only data gathered prior to the end block because after that point, both groups have been treated; we also tried this approach and found no meaningful differences, so we do not show these results.

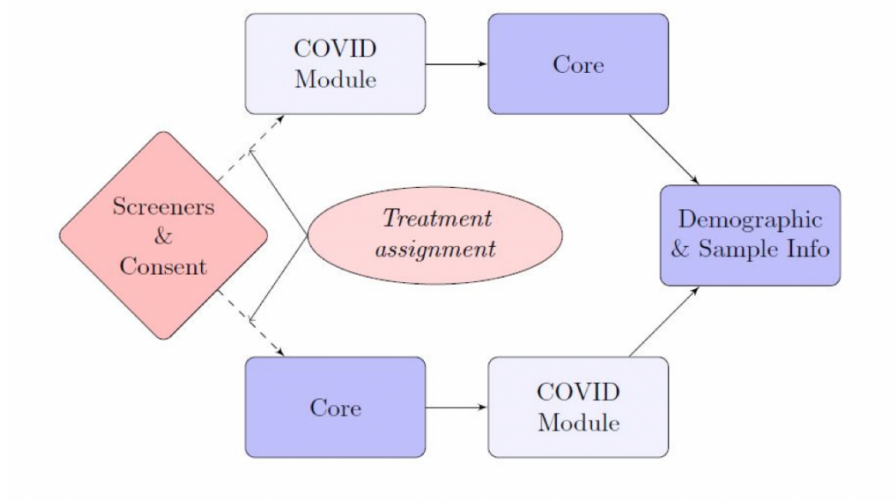


Figure 1: Questionnaire Structure

Measurement

We do not use survey weights in the analyses, for four reasons. First, as a practical matter, weighting is difficult for a study of breakoffs precisely because respondents may drop out before they provide demographic information at the end of the survey. Second, as a study of breakoffs, our main population of

interest is survey takers rather than the population of Haiti. Third, treatments are assigned at random, and we have no a priori theoretical reason to expect that the proposed treatment effects would be heterogenous according to any particular sampling or demographic variable(s). Finally, research finds that, for survey experiments, the benefits of using weights (decreased bias) are relatively small while the costs (loss of statistical power) are substantial (Miratrix et al. 2018).

Per our theoretical framework, concern about COVID should act as a moderator for the relationship between treatment and respondent engagement. For this test, we remove those who responded to a 5-category question about the coronavirus outbreak in their country by reporting that the pandemic is “not so serious” or “not serious at all” or they “have not thought much about” the issue (other options: “very serious” and “somewhat serious”). A total of 439 individuals said that COVID was less than serious, compared to 1,407 who said the outbreak was “very” or “somewhat” serious. There were a total of 28 “don’t know” responses, 246 non-response/refusals, and 109 N/As (i.e., respondents who broke off before being asked the question).

We include the nonresponses in the analyses because eliminating them would unfairly bias the results toward our hypotheses. The N/As disproportionately come from the COVID-Late group (since the people who break off are most likely to do so near the beginning of the survey), so eliminating them would show an artificially low breakoff rate for the COVID-Late group. For this same reason, common methods for testing moderator effects, like an interaction or two-way ANOVA, are inappropriate when using breakoffs as a dependent variable. The N/As cannot be dropped from the analysis, and we can neither group them into with the “COVID serious” or the “COVID unserious” conditions as it would arbitrarily place nearly all breakoff cases into one group or another. For a similar reason, we do not drop the “don’t knows” and refusals. The rate of item nonresponse is much higher for questions at the beginning of the survey (since disinterested respondents give nonresponses and then drop out). Thus, eliminating the “don’t knows” and refusals would eliminate respondents who are both a) more likely to come from the COVID-First group and b) more likely to break off.

Data Analysis

Since this design is experimental and treatment assignment is random, our primary analysis consists of a two-sample z-test to detect significant differences between the COVID-First and COVID-Last conditions on proportion of breakoffs (H1). We then repeat the analysis after eliminating the “COVID not serious” group (H2).

As a further test to eliminate possible confounding variables, we run multivariate logistic regressions for both the full and limited samples. The models include fixed effects for interviewer and also control for respondent age, gender, region, and urban/rural residence. Other sociodemographic variables were asked only at the end of the survey and are thus not available for analysis (many had broken off before answering these questions). We considered a multilevel model as well, with respondents nested within interviewers. However, a Hausman test revealed that a fixed effect model is preferred (i.e., we reject the null hypothesis that individual-specific effects are not correlated with the covariates). Further, we prefer a fixed effects model because of the relatively small number of interviewers (11, compared to the rule of thumb of 30; Kreft and Bokhee 1994) and because we are not theoretically interested in making inferences about the interviewers, merely controlling for their effect. Regardless, the treatment effect changes very little whether we use a random or fixed effect model.

Before proceeding, we acknowledge the limitations of this analysis. Though we believe that the COVID-19

pandemic offered a treatment in the form a broadly salient issue in which most people would be interested, the assumption that respondents are eager to discuss the topic may not hold true for all (even if they think it is an important or serious issue). The pandemic was also a one-of-a-kind situation, and it is uncertain the extent to which it is applicable to other newsworthy events or topics that social scientific researchers may ask about in public opinion surveys (e.g., an election, a war, or the economy). Further, the present data analysis assesses only an average treatment effect and does not probe further into mechanisms that may underlie any potential treatment effects. We recommend that researchers and practitioners take the results and adapt them to the specific circumstances of their own surveys and domains. Context-specific knowledge will inform the applicability of the results to other studies.

Descriptive Statistics and Sample Characteristics

The target number of complete interviews on the survey was 2,000. In our data, a total of 2,390 interviews passed the “Screeners & Consent” block and thus were assigned to a treatment group by the survey software. For this survey experiment, we do not believe that response rates are necessary or even meaningful. Measures like AAPOR’s codes RR1-4 assess the percentage of completed (and sometimes partial) interviews out of all attempts. This percentage has no real relevance to the present study because we are not interested only in the percentage of complete (or partial) interviews, but rather we are focused on breakoffs among those who participate. Further, as a study of cooperation among survey participants, we are not interested in generalizing to the overall Haitian population but rather only to the portion of the population that is willing and able to take surveys. Therefore, any imbalances between the sample and the broader population resulting from noncontact, ineligibility, unknown eligibility, or refusals are irrelevant to the present study because the excluded individuals are not part of the survey-taking population.

We operationalize breakoffs as deliberate hangups or, in other words, interviews in which the respondent says they do not wish to continue the survey or they ask to be called back at another time but never answer the callback. For the breakoff rate analysis, then, we drop 61 cases in which an interview ended early for inadvertent reasons (e.g., dropped call, poor connection). Though this strategy may exclude some real breakoff cases (if, for example, someone abruptly hangs up without saying anything), it is impossible to distinguish these from cases in which there was a real technical malfunction and the respondent would have otherwise completed the interview. Thus, we use the more narrow definition to test the first two hypotheses. In analysis not shown, we did assess the data using a more inclusive definition of breakoff, and the results do not meaningfully change. We also omit 60 cases that were terminated for “other” reasons (usually, because a quota is overfilled) and 40 interviews that were completed but rejected by quality control supervisors. This leaves 2229 observations.

In total, there are 211 breakoffs and 2,018 complete interviews, meaning the overall breakoff rate, i.e., breakoffs divided by [breakoffs plus completes], is 9.5%. The interviewer team consisted of 11 staff, each conducting between 183 and 256 interviews. Breakoff rates between interviewers ranged from 2.7% to 17.2%. Because of the wide range of breakoff rates, we add fixed effects for interviewer in the logistic regression to control for interviewer effects.

The data show that the two treatment groups are balanced on observable demographic characteristics including gender, age, region, urban/rural residence, education, and wealth (table in appendix). In the multivariate analysis, we control for these variables save education and wealth, which are asked at the end of the questionnaire (at which point many have broken off). A simple analysis of the limited set of demographic variables asked in the eligibility block suggests that women were about 30% (3.9

percentage points) more likely to drop out ($p < 0.01$) than men. Age, urban/rural residence, region, and time of interview (day or night, weekend or weekday) are not significantly associated with breakoff rate in bivariate analyses.

Results

Table 1 shows the breakoff rate in each experiment condition. Among the full sample, the breakoff rate for the COVID-Late group is 1.7 percentage points higher than it is for the COVID-First group. This means that for a 2000-person survey, 34 fewer interview attempts need to be made to reach the target sample size when topic-induced motivation is elevated. This difference, however, is not statistically significant (one-tailed p -value = 0.09). To test H2, we use the concern about the seriousness of the pandemic measure to filter out respondents who believe the COVID outbreak is less than “somewhat serious.” As the second data column of Table 1 (“Limited Sample”) shows, the treatment effect nearly doubles to 3.1 percentage points ($p < 0.025$), in line with our expectation from H2.

Table 1. Breakoff Rate, by Question Ordering (%) in Haiti		
	Full Sample	Limited Sample
COVID Questions First	8.6 (0.8)	8.7
COVID Questions Late	10.3 (0.9)	11.8
Net Difference	-1.7 (1.2)	-3.1 (1.4)
p-value (one-tailed)	0.09	0.02
<i>n</i>	2,229	1,790

Note: Standard errors in parentheses. Standard errors for “Net difference” are calculated from two-sample difference-in-proportion z-test.

This difference in the treatment effect between the two samples is consistent with the topic-induced motivation framework: those who care more about the coronavirus should be more affected by the placement of questions related to it. Examining Table 1 more closely reveals that the growth in the treatment effect after limiting the sample is attributable mostly to the COVID-Late group; the breakoff rate for the COVID-First group does not change much between the samples. This lends additional credence to the topic-induced motivation argument. According to the theory, those who are more preoccupied with COVID (the limited sample) should be particularly sensitive to not being asked about the issue. Similarly, we would not expect the difference between the full and limited sample to be large for the COVID-First respondents since they are asked about the topic right away.

To further isolate the effect of the treatment, we estimate logistic regressions using the full and limited sample, while controlling for respondent age, gender, region, and urban/rural residence, along with adding interviewer fixed effects. The results are shown in Table 2. The results show a similar pattern to Table 1. The treatment has an effect in the expected direction in both cases, but it is only statistically significant in the sample in which those who care little about the coronavirus are filtered out. Men are less likely to break off, and, as described in the Descriptive Statistics section, there is significant variation in breakoff rate by interviewer. There is no significant variation by region of residence.

Table 2. Logistic Models Predicting Breakoff Rate		
	Model 1 (Full Sample)	Model 2 (Limited sample)
COVID Qs First	-0.14 (0.15)	-0.31* (0.16)
Age	0.01 (0.01)	0.01 (0.01)
Male	-0.30* (0.15)	-0.49** (0.16)
Urban	-0.26 (0.15)	-0.15 (0.16)
Constant	-1.92 (0.33)	-2.06 (0.36)
Interviewer Fixed Effects	✓	✓
Region Fixed Effects	✓	✓
<i>n</i>	2198	1763

*Note: Standard errors in parentheses. We drop 31 cases which are missing “urban”. The baseline category for “male” (gender of respondent) include women and two cases in which “other” was chosen. p-values are one-tailed for the treatment and two-tailed for the other variables. * $p < 0.05$ ** $p < 0.025$.*

As an aside, the breakoff data seem to coincide with the timing of the pandemic. Data collection took place between April and June, when the pandemic was relatively new. Concern gradually grew over the data collection period (mean of 4.11 on the 5-point concern about the seriousness of the pandemic scale in April, compared to 4.43 in June). Notably, this coincided with a drop in overall breakoff rate (14.09% in April, 4.12% in June). Those who took the survey later on (late May and into June 2020) were significantly less likely to break off ($p < 0.01$). Though not a scientifically rigorous test, this result is in line with the motivating hypothesis that people became more interested in surveys about COVID as the issue became omnipresent. We also tested for heterogeneous treatment effects and found no significant interaction between treatment and date of interview, age, gender, or urban/rural residence. Those living in the North region seemed to be comparatively more affected by the treatment, though the differences between it and the other regions are only marginally significant, and likely due to chance.

Discussion

Existing literature has documented how the stated topic of a survey influences response rates. Prior research has also shown that many facets of questionnaire design, like cognitive load, mode, or structure (e.g., number of questions on a page), are linked to engagement with and motivation to complete surveys. Little research, however, has systematically and experimentally tested how the placement and topic of survey questions can affect respondent engagement once they begin the survey.

We investigate the relationship between question placement and topic, on the one hand, and breakoff behavior in surveys, on the other hand. In doing so, we posit and test the idea that initiating surveys with interesting and relevant questions increases participants’ engagement and, thus, reduce breakoffs. We theorize that two factors – topic (capacity of the survey to produce interest) and question placement (location of a module within the survey) – are jointly important in motivating engagement. We offer a theoretical framework that permits any one of three non-rival micro-mechanisms to undergird this

dynamic: salient questions pique respondent interest, relevant questions convince respondents the survey is worthwhile, and/or questions on important topics induce bonding between the interviewer and interviewee. We test the argument by leveraging one particularly salient topic in 2020: the COVID-19 pandemic. Whereas most research studies this topic with web surveys in developed countries, we focus on behavior during a phone study in a less developed democracy.

In an original experiment carried out in a national phone survey in Haiti, we find a pattern of results that are overall supportive of the notion that question placement and question topic jointly matter. Frontloading the survey with questions about the COVID-19 pandemic led to marginally fewer breakoffs, though the result is not statistically significant. We do find support for the topic-induced motivation argument in a more precise test: the treatment effect widens when those who are unconcerned with the pandemic are removed from the analysis.

We had the fortunate opportunity to repeat this experiment in another probabilistic national phone survey in Ecuador. Fieldwork for this survey took place between December 2020 to January 2021. The sample design, questionnaire, and training protocols were broadly the same as those in Haiti, but the target sample size in Ecuador was only 800 adults. Though breakoffs were much rarer in Ecuador (1.6%), we found that the breakoff rate was over three times higher for those assigned to the COVID-Late Condition compared to the COVID-First Condition, in line with our expectations. The estimated treatment effect was 1.7 percentage points (one-tailed p -value < 0.05), coincidentally the same effect size found in Haiti. However, in the case, the treatment effect barely moved when we remove 60 respondents who said the coronavirus issue was less than serious, unlike what we showed in Haiti, possibly because the number of breakoffs was already so low. Nevertheless, the conclusion was the same as in Haiti: frontloading the questionnaire with COVID items has a marginal but noticeable effect.

Though the effects we find are not immense in magnitude and are on the borderline of statistical significance, based on the consistency of the treatment effect between the two countries, we conclude that surveys will be slightly more efficient if participants are asked questions related to the top issue of the day at the beginning of the questionnaire. Even marginal gains can be beneficial for survey practitioners. Due to a lower breakoff rate, fewer respondents will need to be interviewed to reach the target sample size. In practical terms, for a 2000-person survey, our results would suggest that placing questions about the most salient issue first would turn about 34 breakoffs into complete interviews (62 if the issue were universally seen as a serious issue), which saves a day or two of fieldwork for a survey like ours. However, we consider the COVID outbreak to be a “most likely” case (Gerring 2007) for the topic-induced motivation theory, and, finding only mild effects, we would not necessarily recommend adding “throwaway” questions or adjusting the questionnaire in a way that comprises other survey objectives for this purpose alone. But all else equal, we consider it wise to begin surveys with the most important issues of the day.

Our study contributes to understanding respondent behavior during surveys. We show that the decision to complete a survey once it begins is influenced jointly by topic and question order. This has implications for assessing substantive survey results: because respondents may break off depending on question order (H1), and this behavior is tied to level of interest in the topic (H2), the distribution of valid responses to a question could change depending on where questions are placed in the instrument. For example, in this case, researchers who ask questions about COVID at the end of a questionnaire will show artificially lower levels of concern about the pandemic, because those who are highly preoccupied with the issue already broke off.

In future work, we recommend additional experiment-based studies consider the extent to which similar dynamics can be found across other types of surveys and in different contexts (including but not restricted to different types of crises). While we believe that we would observe similar results in a context in which there is another type of highly salient issue or event (say, the assassination of Haiti’s president), it is not clear if question topic and placement would induce engagement for a moderately important issue (say, an upcoming election). Furthermore, we note that our study does not permit us to assess the micro-mechanisms (piqued interest, belief it is time well-spent, or interviewer-interviewee bonding) that may produce topic-induced motivation. Therefore, we further recommend that future research work to assess these varying paths through which question placement and topic may affect engagement. Additional explorations could also more carefully measure how “interesting” respondents find different topics, since “seriousness” does not necessarily equate to eagerness to talk about an issue. Finally, in this study, we also found substantial differences in breakoff rates by country, interviewer, and gender. As scholarship on breakoff behaviors continues to expand, additional avenues of research ought to include work that considers what factors drive variation across these three variables, and how this relationship might be affected by questionnaire characteristics.

Appendix

Table A: Treatment Group Assignment by Demographic Variables

	COVID-First Group Mean (SE)	COVID-Late Group Mean (SE)	Difference (SE)	t-test p-value (two-sided)
Gender (proportion male)	0.501 (0.017)	0.482 (0.016)	0.019 (0.024)	0.426
Age (years)	37.815 (0.454)	37.849 (0.427)	-0.034 (-0.054)	0.957
Region: North (proportion)	0.194 (0.014)	0.202 (0.013)	-0.008 (0.019)	0.651
Region: Center (proportion)	0.209 (0.014)	0.220 (0.013)	-0.011 (0.019)	0.585
Region: West (proportion)	0.372 (0.017)	0.369 (0.016)	0.003 (0.023)	0.906
Urban (proportion urban)	0.623 (0.017)	0.619 (0.016)	0.004 (0.023)	0.862
Education (years of schooling)	12.665 (0.144)	12.465 (0.144)	0.200 (0.204)	0.328
Wealth (PCA score)	0.013 (0.079)	-0.012 (0.072)	0.025 (0.106)	0.814

Notes: Category “Region: South” omitted for redundancy. “Wealth” is based on principal components analysis scores, which are in turn based on reported possession of various household items. Gender, age, region, and urban were assessed near the beginning of the questionnaire, and education and wealth near the end (and thus the latter have much higher missingness/standard errors).

References

1. Ambitho, Angela, Yashwant Deshmukh, Johnny Heald, and Ignacio Zuasnabar. “Changes in Survey Research in Times of COVID-19.” Webinar from World Association for Public Opinion Research, Virtual, 21 May 2020. <https://wapor.org/resources/wapor-webinars/webinar-may-2020/>
2. Alshaabi, Thayer, Michael V. Arnold, Joshua R. Minot, Jane Lydia Adams, David Rushing Dewhurst, Andrew J. Reagan, Roby Muhamad, Christopher M. Danforth, and Peter Sheridan Dodds. “How the

- World's Collective Attention is being Paid to a Pandemic: COVID-19 Related N-Gram Time Series for 24 Languages on Twitter." *Plos One* 16, no. 1 (2021).
3. Berinsky, Adam J. "Measuring Public Opinion with Surveys." *Annual Review of Political Science* 20 (2017): 309-329.
 4. Crawford, I.M. *Marketing Research and Information Systems*. Rome, Food and Agriculture Organization Of The United Nations, 1997.
 5. Galesic, Mirta. "Dropouts on the Web: Effects of Interest and Burden Experienced During an Online Survey." *Journal of Official Statistics* 22, no. 2 (2006): 313-328.
 6. Garbarski, Dana, Nora Cate Schaeffer, and Jennifer Dykema. "Interviewing Practices, Conversational Practices, and Rapport: Responsiveness and Engagement in the Standardized Survey Interview." *Sociological Methodology* 46, no. 1 (2016): 1-38.
 7. Gerring, John. "Is There a (Viable) Crucial-Case Method?" *Comparative Political Studies* 40, no. 3 (2007): 231-253.
 8. Groves, Robert M., Stanley Presser, and Sarah Dipko. "The Role of Topic Interest in Survey Participation Decisions." *Public Opinion Quarterly* 68, no. 1 (2004): 2-31.
 9. Groves, Robert M., Eleanor Singer, and Amy Corning. "Leverage-Saliency Theory of Survey Participation: Description and an Illustration." *Public Opinion Quarterly* 64, no. 3 (2000): 299-308.
 10. Holland, Jennifer L., and Leah Melani Christian. "The Influence of Topic Interest and Interactive Probing on Responses to Open-Ended Questions in Web Surveys." *Social Science Computer Review* 27, no. 2 (2009): 196-212.
 11. Keeter, Scott, Courtney Kennedy, Michael Dimock, Jonathan Best, and Peyton Craighill. "Gauging the Impact of Growing Nonresponse on Estimates from a National RDD Telephone Survey." *Public Opinion Quarterly* 70, no. 5 (2006): 759-779.
 12. Keusch, Florian. "The Role of Topic Interest and Topic Saliency in Online Panel Web Surveys." *International Journal of Market Research* 55, no. 1 (2013): 59-80.
 13. Kreft, Ita GG, and Bokhee Yoon. "Are Multilevel Techniques Necessary? An Attempt at Demystification." (1994).
 14. Krosnick, Jon A., and Stanley Presser. "Question and Questionnaire Design." In *Handbook of Survey Research 2nd Edition*, eds. James D. Wright and Peter V. Marsden, p. 263-313. Bingley: Emerald Group Publishing Ltd., 2010.
 15. Manfreda, Katja Lozar, Zenel Batagelj, and Vasja Vehovar. "Design of Web Survey Questionnaires: Three Basic Experiments." *Journal of Computer-Mediated Communication* 7, no. 3 (2002): JCMC731.
 16. Mavletova, Aigul, and Mick P. Couper. "A Meta-Analysis of Breakoff Rates in Mobile Web Surveys." In *Mobile Research Methods: Opportunities and Challenges of Mobile Research Methodologies*, eds. Daniele Toninelli, Robert Pinter, and Pablo de Pedraza, p. 81-98. London: Ubiquity Press, 2015.
 17. McGonagle, Katherine A. "Survey Breakoffs in a Computer-Assisted Telephone Interview." *Survey Research Methods* 7, no. 2 (2013): 79-90.
 18. Miratrix, Luke W., Jasjeet S. Sekhon, Alexander G. Theodoridis, and Luis F. Campos. "Worth Weighting? How to Think about and Use Weights in Survey Experiments." *Political Analysis* 26, no. 3 (2018): 275-291.
 - Musch, Jochen, and Ulf-Dietrich Reips. "A Brief History of Web Experimenting." In *Psychological Experiments on the Internet*, pp. 61-87. Academic Press, 2000.
 19. Peytchev, Andy. 2009. "Survey Breakoff." *Public Opinion Quarterly* 73 (1): 74-97.
 20. Pew Research Center. "Writing Survey Questions." *Pew Research Center*. Accessed August 10, 2023. <https://www.pewresearch.org/our-methods/u-s-surveys/writing-survey-questions/>
 21. Qualtrics. "Survey Question Sequence, Flow, & Style Tips." *Qualtrics*. Accessed August 10, 2023. <https://www.qualtrics.com/experience-management/research/question-sequence-flow-style/>
 22. Rossi, Peter H., James D. Wright, and Andy B. Anderson, eds. Preface to *Handbook of Survey Research*. P. xv. New York: Academic Press, 2013.
 23. Roßmann, Joss, Markus Steinbrecher, and Jan Eric Blumenstiel. "Why do Respondents Break Off Web

- Surveys and Does It Matter? Results from Four Follow-Up Surveys. *International Journal of Public Opinion Research* 27, no. 2 (2015): 289-302.
24. Rouzier, Vanessa, Bernard Liautaud, and Marie Marcelle Deschamps. "Facing the Monster in Haiti." *New England Journal of Medicine* 383, no. 1 (2020): e4.
 25. Schaeffer, Nora Cate, and Jennifer Dykema. "Questions for Surveys: Current Trends and Future Directions." *Public Opinion Quarterly* 75, no. 5 (2011): 909-961.
 26. Shropshire, Kevin, James Hawdon, and James Witte. "Web Survey Design: Balancing Measurement, Response, and Topical Interest." *Sociological Methods and Research* 37, no. 3 (2009): 344-370.
 27. Sun, Hanyu, Frederick G. Conrad, and Frauke Kreuter. "The Relationship between Interviewer-Respondent Rapport and Data Quality." *Journal of Survey Statistics and Methodology* 9, no. 3 (2021): 429-448.
 28. Tu, Su-Hao, and Pei-Shan Liao. "Social Distance, Respondent Cooperation and Item Non-response in Sex Survey." *Quality & Quantity* 41, no. 2 (2007): 177-199.
 29. Van Kenhove, Patrick, Katrien Wijnen, and Kristof De Wulf. "The Influence of Topic Involvement on Mail-Survey Response Behavior." *Psychology & Marketing* 19, no. 3 (2002): 293-301.