

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. Our primary test is an experiment embedded in a 2020 survey in Haiti. Respondents were randomly assigned to a survey that began with questions about the salient COVID-19 pandemic, or the same survey with those questions placed toward the end. Those assigned to the former condition were only slightly less likely to break off, though the treatment effect widens when we exclude those who do not believe that coronavirus is a serious concern. We find a similar breakoff pattern in a survey experiment in Ecuador. We then add data from Mexico and Peru to our analyses, and show a correlation between concern about the pandemic and lower likelihood of breaking off in questionnaires that begin with the pandemic. Our findings show that - at least in times of crisis - adjusting survey content and structure to place timely topics upfront may increase completion rates, but only at the margins.

Keywords: breakoffs, non-response, survey topic, experiment, Haiti, COVID-19

Do Question Topic and Placement Shape Survey Breakoff Rates?

Surveys have become a dominant method of research in many social scientific fields (Rossi et al. 2013) and, as a tool for measuring popular will, serve an essential role in democratic governance (Berinsky 2017). Unfortunately, many surveys suffer from high rates of breakoff – that is, instances in which respondents prematurely and deliberately terminate their participation. Meta-analyses of web surveys have found median breakoff rates of 7.5% (Mavletova and Couper 2015), 16% (Musch and Reips 2000), and 34% (Lozar-Manfreda and Vehovar 2002), while a major telephone survey (the Panel Study of Income Dynamics) reported a breakoff rate of 23% (McGonagle 2013).

Survey researchers aim to minimize 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, 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 (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 add an argument about question topic and placement: placing interesting and relevant questions at the beginning of a survey will generate greater engagement with the remainder of the survey. Though survey professionals often advise this,¹ 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 this particular type of topic-induced motivation ought to 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.

To examine how breakoff rates are shaped by question topic and placement, we leverage the onset of the COVID-19 pandemic in 2020. More specifically, we include a small set of questions about a highly salient and nationally important issue – the pandemic – on phone surveys about democratic governance in four countries: Haiti, Ecuador, Mexico, and Peru. We adopt both experimental and observational approaches to assess the extent to which these questions (and their placement) motivate greater overall engagement – and, thus, fewer breakoffs. Our core test consists of an experiment embedded within the survey in Haiti, conducted April-June 2020. Individuals were randomly assigned to answer questions related to the pandemic at the start or toward the end of the survey. We repeat this experiment in Ecuador, and also use non-experimental data from all four studies to assess the relationship between concern about COVID and breakoff tendencies.

We find, in brief, mild evidence that placing interest-arousing topics at the start of the survey can minimize breakoff behavior, and that 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 the most important 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, survey practitioners around the globe reported anecdotal accounts of individuals being more willing to participate in interviews when the survey topic was related to the pandemic (e.g., Ambitho et al. 2020).

Existing scholarship offers theoretical backing for the idea that the salience of the pandemic shaped 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) posits that the outcome of each survey request is influenced by multiple attributes, including, among other factors, the survey topic. According to this theory, survey topic (and other factors like monetary incentives or advanced notice) can affect propensity to cooperate via the interaction of two factors: leverage (the importance that individuals or subgroups assign to the topic) and salience (how the topic is introduced or emphasized in the request protocol). Groves, Presser, and Dipko (2004) show that sampled individuals are more likely to cooperate when the survey topic is of interest to them, but the effect is moderated by the presence of monetary incentives.

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, questionnaire length, and so on. Additionally, much of the literature linking survey topic and response rates, including the leverage-saliency theory, assumes that the survey has a singular, unified

topic that can be succinctly expressed in the request to participate. There are numerous circumstances in which this assumption does not hold. These circumstances include: a survey has an omnibus structure so that it includes questions on a range of topics; a survey on an overarching topic (e.g., health) has questions with multiple subtopics for which respondents have varying levels of interest/comfort (e.g., diet, COVID-19, mental health, drug usage); respondents' expectations about the topic of the survey are different than what is actually asked (e.g., they are told it is about current events but are not asked about the issue that is most important to them).

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. Specifically, we posit a type of topic-induced motivation that is generated by placing interesting and relevant questions at the start of the survey. We offer three non-rival mechanisms through which question placement ought to shape survey engagement. First, respondents may experience a bump in interest when they engage with a personally relevant topic. If thought-provoking or intriguing questions pique respondent interest at the start of the survey, that should create an initial "bank" of interest. A slow decay in that banked interest could then keep engagement comparatively high even as the survey moves on to other topics.

Second, it may be that initiating the survey with questions on topics that are particularly relevant convinces respondents that the broader research effort is worthwhile, thus increasing willingness to stay engaged. Prior research has established that those who are active in their communities are more likely to accept solicitations for survey participation (Groves, Singer, and Corning 2000). Others have found that belief in the importance of scientific studies predicts a lower breakoff rate in web-based surveys (Roßmann, Blumenstiel, and Steinbrecher 2015). Scientific surveys offer an opportunity to voice opinion, assist in research, and potentially help influence public dialogue or policy on issues. It follows that if a survey opens on topics that participants consider as important and timely, they may be more likely to complete the entire survey compared to a survey on more trivial issues or one designed

for commercial/marketing purposes.

Third, beginning a questionnaire by acknowledging a highly salient issue may help establish a rapport between interviewer and respondent that carries to the end of the interview. 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). These studies and others have considered how shared characteristics (like a common gender, race, or marital status) and interview experiences like laughter (see Lavin and Maynard 2001) can foster bonding between interviewer and subject, but few have considered how the questions themselves influence rapport. We theorize that during times of crisis (e.g., political upheaval, war, economic decline, health crisis), it may come across as insensitive or untimely to discuss irrelevant issues. Conversely, questions that acknowledge important local issues could foster interviewer-interviewee bonding, which ought to boost respondent cooperation and thus engagement.

Prior studies on breakoffs and respondent interest, which by and large rely on data from self-administered web-based questionnaires, 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 experimentally testing whether question topic and placement combine to elicit lower breakoff rates. In other words, in a survey that covers

multiple topics, can researchers optimize their questionnaire design by placing interesting questions first, or will breakoff rates “even out” once the survey moves on to other subjects?

A multitude of studies from a variety of fields have shown “order effects” (also called priming or context effects) with respect to question placement on substantive responses – in other words, they show prior questions or pieces of information influence responses to subsequent questions (see, for example, Kalton and Schuman 1982; Lee and Grant 2009; Schuman, Kalton, and Ludwig 1983; Strack 1992; Van de Walle and Van Ryzin 2011). In addition, some have found that by placing at the start of the survey various scripts that request the interviewee’s consent to merge their responses with additional (administrative, social media) data, researchers can induce higher rates of agreement from respondents (Sakshaug, Tutz, and Kreuter 2013; Mneimneh 2020).

Fewer studies assess the relationship among question order, topic, and engagement with the survey itself. Peytchev (2009) finds that “question characteristics” are related to likelihood of breakoff, but his focus is on the cognitive difficulty of the question rather than the actual subject matter, and many of his findings are specific to web-based surveys. Studies by Galesic (2006) and McGonagle (2013) similarly focus on characteristics of questions other than their subject domain. For reasons outlined above, we suspect that the nature of question placement and topic is important for motivating respondents to continue with the survey: frontloading the questionnaire with interesting, relevant, and important topics ought to cause engagement on the survey to become elevated, with that effect enduring for some period of time after the interviewer moves on to other topics.²

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. Within that context, we hypothesized the following about the relationship between question placement and topic, on the one hand, and respondent engagement, on the other hand:

- **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. 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:

- **H1a:** 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).

Further, if question topic and placement do in fact elicit greater engagement with the survey, then we also expect that the treatment will have an effect on other measures of engagement besides breakoffs. We consider two alternative dependent variables as operationalizations of respondent engagement in the survey: number of questions answered before breaking off and rate of item non-response.

- **H2:** Among those who break off, those who are asked questions about coronavirus first answer more questions before terminating than those who are first asked questions about other topics.
- **H3:** Respondents who are asked questions about the coronavirus first are less likely to give non-responses (don't know, refuse to answer) than those who are first asked questions about other topics.

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 23 to June 10, 2020. An experienced local survey firm drew the probabilistic sample and recruited interviewers, who were trained by the team responsible for the

survey.³ All interviews were recorded and audited for quality control: 100% were audited by the local firm's office according to pre-defined protocols, and about 20% were re-audited by the research team. Sampling relied on random digit dialing of a sample of active cell phone numbers, supplemented by frequency matching to realize census-derived targets for the survey on region, gender, and age cohorts. The purpose of the survey was to collect public opinion data on a variety of issues related to democratic governance.⁴ We embedded our study within this survey.

For this experiment, the selection of Haiti as a case was determined by survey objectives that are unrelated to the 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 moderation analysis). We believe this issue provides a test of the theory that is generalizable to other crisis situations, including natural disasters (such as the 7.2 magnitude earthquake that struck Haiti in August 2021), severe political upheavals (such as the assassination of Haitian president Jovenel Moïse in July 2021), wars, terrorist events, economic recessions or crashes, crime waves, or major scandals.

The 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 (see appendix for the module items). In the

second, the COVID-Late Condition, respondents were asked those 10 questions toward the end of the survey. No other module positions differed across these conditions. [Citation omitted for review] has already shown that this experiment has substantive effects: priming individuals to think about the pandemic influences certain democratic attitudes. 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, the survey software splits respondents into two branches (the treatment groups). Respondents then receive either the COVID module (10 questions) and then about 35 substantive survey items (exact number depends on branching) which we call the “core”,⁵ or vice versa. 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. In total, respondents are asked around 32 questions in the end block, though it takes much less time to complete than the core because the items are mostly short, straightforward recall questions. The main analyses consider only data within the COVID module, the core and the end block; those who break off or refuse during the initial screeners are dropped from the analysis because that module is delivered pre-treatment. During the end block, both groups will have already been treated with the COVID questions. Therefore, an alternative approach is to look only at data gathered *prior* to the end block. We implement this alternative research design as a robustness check.

[Figure 1 Here]

Analysis Strategy

Since this design is experimental and treatment assignment is random, our main 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 use independent sample *t*-tests to assess treatment effects on questions answered before breakoff (H2) and item non-response rates (H3). Since the hypotheses are directional, we use one-tailed p-values. We do not use survey weights in any analyses.⁶

In analyses that permit a more precise test of the argument (as specified in Hypothesis 1a), 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 (see appendix for full question wording for this variable, COVID1). In discussion below, we refer to this as a measure of concern about the seriousness of the pandemic. Per our theoretical framework, we do not expect the treatment to have an effect on those who are not concerned; in this way, concern about COVID acts a moderator for the relationship between treatment and respondent engagement.⁷ A total of 439 individuals who reached the core gave one of these responses, compared to 1,407 that said the outbreak was “very” or “somewhat” serious. There were a total of 274 non-responses (i.e. answered “don’t know” or “no response”) and 109 N/As (i.e. broke off before being asked the question). We include these non-answers in the analyses because we do not know how seriously they take the pandemic. This should make our test more conservative; these cases add noise to the data, making it more difficult to detect if the treatment effects differ between the full and limited samples.

We further explore the topic-induced motivation theory using a series of follow-up analyses. The first repeats the main analyses using an alternative research design that uses the same data but excludes analysis of behavior during the end block of the questionnaire (“Demographic & Sample Info” block in Figure 1), at which point both groups have received the treatment. Second, we repeat the same experiment and analyses in another study in Ecuador. Third, using non-experimental data from these two studies and two additional ones in Mexico and Peru, we examine an observable implication of our argument: when COVID questions are at the beginning of a survey, there should be a correlation between the

extent to which one believes that COVID is a serious issue and their likelihood of breaking off. We test this relationship using a logistic regression.

Descriptive Statistics and Sample Characteristics

The target number of complete interviews on the survey was 2,000. There were 3,327 calls that connected to potential participants. Nearly a third are excluded from our analyses because of ineligibility, quality control issues, unwillingness to give consent, or drop-out during the screening phase. Participants are assigned to treatment groups randomly by the survey software at the final screen of the initial block (i.e., just before the COVID-First group receives the first COVID question). For our analyses, we consider only the 2,290 interviews that passed the “Screeners & Consent” block and thus were assigned to a treatment group.

The software used for the study did not record a reliable call log, which is necessary for calculating an overall response rate. However, for this survey experiment, we do not believe that response rates are necessary or even meaningful, for a few reasons. First, response rate definitions like AAPOR codes RR1-4 measure 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 only interested in complete (and partial) interviews, but rather we are focused on breakoffs among those who participate. Second, as a study of cooperation among survey participants, we are not interested in generalizing to the overall Haitian population (the purpose of weighting) 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. Lastly, research has shown that for survey experiments, non-probability convenience samples can produce treatment effect estimates that are remarkably similar to those found in nationally representative population-based samples (Mullinix et al. 2015).

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.⁸ We add the excluded 61 cases back in for the item non-response analysis since the reason for early termination is not particularly relevant to that assessment.

Using the narrow definition, there are 211 breakoffs and 2,018 complete interviews, meaning the overall breakoff rate (breakoffs divided by [breakoffs plus completes]) is 9.5%. The interviewer team consisted of 11 staff, each conducting between 167 and 229 interviews. Breakoff rates between interviewers ranged from 2.7% to 17.2%. To account for the clustering of observations by interviewer, we use robust standard errors on our tests of treatment effects. In-depth analysis of interviewer effects is beyond the scope of this paper; we have no theoretical expectations about how the treatment would interact with interviewer characteristics, nor do we have enough information about interviewers or enough observations per interviewer to leverage for further exploration.

The data show that the two treatment groups are balanced on all observable demographic characteristics including wealth, education, gender, age, region, and urban/rural residence. It is difficult to say whether breakoffs are associated with these demographic variables because of the inherent censoring problem: those who break off may not provide demographic data before leaving the interview. A naïve 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. Date of interview also matters: those who took the survey later on (late May and into June 2020) were significantly less likely to

break off ($p < 0.01$), in line with the timing of the spread of the virus. Age, urban/rural residence, region, and time of interview (day or night, weekend or weekday) are not significantly associated with breakoff rate.

Main Results: Pandemic Module Placement in Haiti

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 ($p = 0.14$). To test H1a, 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 ($p < 0.1$), in line with our expectation from H1a.

[Table 1 Here]

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 our 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.

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). Though not a definitive test, this result is in line with our motivating hypothesis that people became more interested in surveys about COVID as the issue became omnipresent.⁹

According to the second hypothesis, those who break off should stay on the phone longer if they receive the COVID questions first. Table 2 shows, among breakoffs, the average number of questions that a respondent was asked by the interviewer before exiting the survey.¹⁰ Directionally, the results are in line with H2. In the full sample, the treatment effect is 1.9 questions ($p = 0.11$), while in the sample excluding those who take the pandemic less seriously, it grows just slightly to 2.3 ($p = 0.15$). To make this more concrete: those in the COVID-First group who broke off were asked around two more questions before exiting than those who were a part of the COVID-Late group and dropped out. These patterns are in line with the expectation stated in H2, yet not significant given the large standard errors.

[Table 2 Here]

To assess the third hypothesis, we consider item non-response as an alternative operationalization of respondent engagement. In Table 3, we show the average item non-response rate, defined as the percentage of “don’t know” or “no response” answers given out of all questions heard, among each treatment group. The empirical evidence provides little support for H3. In both the full and reduced (removing those who take the virus less seriously) samples, item non-response rates differ little between the two treatment groups.

[Table 3 Here]

Secondary Analyses

To extend the findings from our principal study, we perform three additional sets of analyses. First, we assess the hypotheses for the Haiti experiment with an alternate research design that excludes behavior in the end block within the analysis. Second, we included a variant of the experiment in a second national phone survey, conducted in Ecuador, which we assess in this section. And, as well, we investigate correlations between concern about the pandemic and breakoff behaviors among respondents to the surveys in Haiti and Ecuador, plus two additional national phone surveys, in Mexico and Peru.

Excluding Post-Treatment Module

The above analysis considered behavior that occurred at any time during the questionnaire after treatments were assigned, including the end block. That is an appropriate analytical approach since the intention of the study is to compare a condition in which a particularly relevant topic is addressed first, to one in which it not asked first. At least one mechanism we describe in our theoretical motivation – that asking about important topics first establishes a rapport between respondent and interviewer – requires that the salient questions are placed at the very beginning of the questionnaire. However, it could be argued that after the COVID-Late module, data from the end block is irrelevant because both groups have received the treatment. Therefore, as a stricter test of our theory, we repeat our analyses excluding data from the post-treatment module; that is, we analyze only the data that formed part of the “core” questionnaire plus the COVID modules (see Figure 1). This means that individuals who break off during the end block – in other words, those who complete the core but do not finish the whole survey – flip from a 1 (breakoff) in our earlier analysis to a 0 (complete) in this analysis.

[Table 4 Here]

Table 4 shows the results for the dataset that excludes the end block. The results are

largely similar to those found in the main analyses. The breakoff rate decreases by 1.6 percentage points when the COVID module is placed near the end of the questionnaire (not statistically significant). The treatment effect doubles when excluding those that take the virus less seriously ($p < 0.1$). Again, those who receive the COVID-First Condition answer more questions before breaking off, though the difference is not statistically significant. With respect to H3, the difference in the non-response rate for the two treatment groups is minimal.

Ecuador Study

We were able to repeat the experiment in a probabilistic national phone survey of Ecuador, carried out December 4, 2020 to January 4, 2021. As with Haiti, the choice of Ecuador as a case was driven by objectives unrelated to the present study. The sample design and protocols were broadly the same – random digit dialing of cell numbers targeting voting-age citizens. However, the target sample size in Ecuador was only 800 adults. Two local firms were employed to collect data, following the same guidelines, training, oversight, quality control, and other protocols managed by the research team.

The questionnaire covered similar topics as the one in Haiti – attitudes toward democratic governance, elections, institutions, the economy, and crime were major topics – though specific question wordings varied due to other survey objectives and adjustments for country context.¹¹ In Ecuador, following the screening questions, all participants were first asked an ice breaker question (“what is the most serious problem facing the country?”), and then were asked either nine questions related to the pandemic (COVID-First Condition), or a question about interpersonal trust (COVID-Late Condition). The first four questions of the COVID module were the same as in Haiti, though the next five questions differed. The questionnaire in Ecuador was slightly longer, with an average of 41 questions in the core (compared to 36 in Haiti). Average levels of concern about the coronavirus outbreak were much higher in Ecuador, with an average seriousness rating of 4.65 on the 5-point scale (4.23 in Haiti).

Table 5 shows results from Ecuador. Compared to Haiti, breakoffs were very infrequent

in Ecuador: out of 889 attempted interviews (among those who consented to participate), 822 were successful, complete interviews. Using the criteria described in the Data & Methods section above, we code only 13 cases as breakoffs (other cases were omitted due to refusal, quality control issues, or inadvertent early terminations). Nevertheless, we find, as expected by H1, a significant treatment effect on breakoff rates: the breakoff rate is 1.7 percentage points higher for those that were assigned to the COVID-Late Condition ($p < 0.05$), coincidentally the same effect size found in Haiti. Also, those who broke off and were in the COVID-First group were asked, on average, 3 more questions before breaking off than those who received the COVID questions toward the end (difference was not significant due to very low n). As in Haiti, the item non-response rate differed little between the two conditions. Overall, then, though breakoff and item non-response rates are substantially lower in Ecuador than Haiti, the pattern of results with regard to treatment effects are remarkably similar between the two countries.

[Table 5 Here]

Four-Country Correlational Analysis

In addition to the surveys in Haiti and Ecuador, we were able to gather additional data via nationally representative phone studies in Mexico (July-August 2020) and Peru (July 2020). These studies were also focused on various issues related to democratic governance and current events, including attitudes toward democracy, political institutions and leaders, elections, crime, and economic security. As in Haiti and Ecuador, a local survey firm was employed to carry out interviews, and our research team carried out training procedures and enforced quality control using the same standardized protocols to ensure quality and comparability. Samples consisted of voting-age adults who were reached via RDD of cell phone numbers. The target number of completes was 2000 in each study.

The questionnaires in Mexico and Peru included a COVID module similar to the ones used in Ecuador and Haiti, though its placement in the questionnaire was not randomized.

Rather, the module was included in all interviews at the beginning, where the COVID-First Condition placed the questions. We use these additional data to test an observable implication of the main theory. If our argument that salient, relevant, and important topics motivate engagement is correct, then wherever the COVID module begins the survey, we should find a correlation between concern about the coronavirus and likelihood of breakoff.

To measure concern, we again turn to the question about how serious the pandemic is in the country; the question was included as the first item in the pandemic module in each country (see appendix). Table 6 displays the breakoff rate for each response to the question: “very serious”, “somewhat serious”, “not so serious”, “not at all serious”, or “haven’t thought much about this”. The columns from Haiti and Ecuador exclude data from the COVID-Late group, as we cannot expect breakoff to be substantially influenced by their reported concern about the outbreak (which comes later in the questionnaire). This ensures consistency in comparing results across the four different countries. The overall breakoff rate in Peru is 0.7%, while it is 6.6% in Mexico, 3.3% in Ecuador (COVID-First group only), and 15.9% in Haiti (COVID-First group only).¹²

The results are consistent with the topic-induced motivation theory. Breakoff rates are generally higher the less seriously one views the coronavirus pandemic. To assess the nature of the relationship, the table shows the result of a logistic analysis in which breakoff (1/0) is regressed on the 5-point concern measure (a “haven’t thought about it” response is treated as expressing the least level of concern). In Haiti, Peru, and Mexico, there is a negative relationship, and the coefficient is significant in analyses of the data for Haiti and Peru. The coefficient sign flips in Ecuador, though it is near zero and nowhere close to statistically meaningful; this is an artifact of there being very little variation to analyze, as there are only three cases of breakoff in the COVID-First group in which the respondent answered the COVID seriousness question.

[Table 6 Here]

Conclusion

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 and related behaviors. 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 phone studies in less developed democracies.

Our principal study is an original experiment carried out in a national phone survey in Haiti. We find a pattern of results that overall supports 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 and led respondents to stay on the phone slightly longer (although, contrary to expectation, there were no differences item non-response rates). Importantly, in line with the topic-induced motivation argument, the treatment effect widens when those who are unconcerned with the pandemic are removed

from the analysis. In a repeated experiment in a different context, Ecuador, the results fit the same pattern. Further, we find correlational evidence consistent with the argument, when considering the general pattern of connections between perceptions of seriousness of the pandemic and breakoff behavior, among a set of surveyed individuals with four country questionnaires that placed COVID-19 questions at the start.

Based on the consistent differences across experimental conditions, we conclude that surveys in a context marked by a health crisis will be slightly more efficient if participants are asked questions related to the crisis at the top of the questionnaire: due to a lower breakoff rate, fewer respondents will need to be interviewed to reach the target sample size. The size of the effect is not extremely large, so we do not necessarily recommend adding “ringer” or “throwaway” questions¹³ solely to increase cooperation rates, but all else equal, it is better to start with questions on the most important issues. It is important to note, though, that the effect size – and hence, the strength of our recommendation – might change depending on the salience of the issue at hand. Our Haiti survey began just as COVID-19 was beginning to take shape, so it could be that the treatment effect would be more acute in a context in which more people were concerned about the virus (or another issue). On the other hand, in cases where there is no single overarching issue or broadly interesting topic to discuss, it may be useless or even disadvantageous to ask about an issue that seems strangely out of place within the rest of the survey.

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 (H1a), 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). We also 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. Finally, in this study, we also found substantial variation in breakoff rates by both country and interviewer. As scholarship on breakoff behaviors continues to expand, additional avenues of research ought to include work that considers what factors drive variation across these two variables, and how this relationship might be affected by questionnaire characteristics.

References

- 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/>
- 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).
- Berinsky, Adam J. "Measuring Public Opinion with Surveys." *Annual Review of Political Science* 20 (2017): 309-329.
- 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.
- 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.
- 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.
- Holbrook, Allyson L., Melanie C. Green, and Jon A. Krosnick. "Telephone versus Face-to-face Interviewing of National Probability Samples with Long Questionnaires: Comparisons of Respondent Satisficing and Social Desirability Response Bias." *Public Opinion Quarterly* 67, no. 1 (2003): 79-125.

- 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.
- Kalton, Graham, and Howard Schuman. "The Effect of the Question on Survey Responses: A Review." *Journal of the Royal Statistical Society, Series A* 145, no. 1 (1982): 42-57.
- 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.
- Keusch, Florian. "The Role of Topic Interest and Topic Salience in Online Panel Web Surveys." *International Journal of Market Research* 55, no. 1 (2013): 59-80.
- 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.
- Lavin, Danielle, and Douglas W. Maynard. "Standardization vs. Rapport: Respondent Laughter and Interviewer Reaction during Telephone Surveys." *American Sociological Review* (2001): 453-479.
- Lee, Sunghee, and David Grant. "The Effect of Question Order on Self-rated General Health Status in a Multilingual Survey Context." *American Journal of Epidemiology* 169 no. 12 (2009): 1525-1530.
- Martin, Charles L. "The Impact of Topic Interest on Mail Survey Response Behavior." *Journal of the Market Research Society* 36, no. 4 (1994): 327-339.
- 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.
- McGonagle, Katherine A. "Survey Breakoffs in a Computer-Assisted Telephone Interview." *Survey Research Methods* 7, no. 2 (2013): 79-90.
- 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.
- Mullinix, Kevin J., Thomas J. Leeper, James N. Druckman, and Jeremy Freese. "The Generalizability of Survey Experiments." *Journal of Experimental Political Science* 2, no. 2 (2015): 109-138.
- Mneimneh, Zeina. "Consent to Link Twitter Data to Survey Data: A Comprehensive Assessment." Presentation at BigSurv20 Conference, Virtual, 11 November, 2020.
- Peytchev, Andy. 2009. "Survey Breakoff." *Public Opinion Quarterly* 73 (1): 74-97.
- Rossi, Peter H., James D. Wright, and Andy B. Anderson, eds. Preface to *Handbook of Survey Research*. p. xv. New York: Academic Press, 2013.
- 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.
- Rouzier, Vanessa, Bernard Liautaud, and Marie Marcelle Deschamps. *Facing the Monster in Haiti*. New England Journal of Medicine 383, no. 1 (2020): e4.
- Sakshaug, Joseph W., Valerie Tutz, and Frauke Kreuter. 2013. "Placement, Wording, and Interviewers: Identifying Correlates of Consent to Link Survey and Administrative Data." *Survey Research Methods* 7, no. 2 (2013): 133-144.

- Schaeffer, Nora Cate, and Jennifer Dykema. "Questions for Surveys: Current Trends and Future Directions." *Public Opinion Quarterly* 75, no. 5 (2011): 909-961.
- Schuman, Howard, Graham Kalton, and Jacob Ludwig. "Context and Contiguity in Survey Questionnaires." *Public Opinion Quarterly* 47, no. 1 (1983): 112-115.
- 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.
- Strack, Fritz. "'Order Effects' in Survey Research: Activation and Information Functions of Preceding Questions." In *Context Effects in Social and Psychological Research* eds. Norbert Schwarz and Seymour Sudman, 23-34. New York: Springer, 1992.
- 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.
- 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.
- Van de Walle, Steven, and Gregg G. Van Ryzin. "The Order of Questions in a Survey on Citizen Satisfaction with Public Services: Lessons from a Split-Ballot Experiment." *Public Administration* 89, no. 4 (2011): 1436-1450.
- 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.

Notes

¹See recommendations on question order from Pew Research (<https://www.pewresearch.org/our-methods/u-s-surveys/writing-survey-questions/>) and Qualtrics (<https://www.qualtrics.com/experience-management/research/question-sequence-flow-style/>).

²This framework could offer a partial or additional explanation for differences in substantive order effects; systematic differences between individuals that exit the survey early and those who complete after receiving the “treatment” can influence response distributions on subsequent questions.

³The authors are part of the team responsible for the survey. Funding for the survey came from [source omitted for blind review]. Survey research is legal in Haiti, and the survey was reviewed and determined exempt from full review by an Institutional Review Board at [omitted for blind review]. This is also true for the other surveys included in analyses in this paper.

⁴The study information (consent) script at the start of the interview did not mention the pandemic, but instead only generally referenced that the study was about “the situation in Haiti.” We do not mention the coronavirus issue because our theory is about both topic and question placement, and this design avoids pre-treating all respondents.

⁵The modules in the core of the survey are as follows: pandemic (10 questions), economic situation (2), services (2), interpersonal trust (3 questions), democratic attitudes (4), crime and insecurity (6), attitude toward TV (1), system support and trust in institutions (6), presidential approval / vote intention (2), perceptions of corruption (1), health / medical services (6), political interest and knowledge (2), and welfare (1).

⁶There are three main reasons to not use weights. First, as a practical matter, weighting is difficult for a study of breakoffs precisely because respondents may drop out before they give information about their personal or household characteristics; applying weights would eliminate a large portion of the dataset. Second, the treatments are assigned at random, and we have no theoretical reason to expect that the proposed treatment effects would vary according to any particular sampling or demographic variable(s). Third, research finds that, for survey experiments, the benefits of using weights (decreased bias) are small while the costs (loss of statistical power) are substantial (Miratrix et al. 2018).

⁷The more traditional method for testing for moderators, an interaction term, is inappropriate in the present study because of a data censoring issue. Most respondents who are in the COVID-Late group and who break off do not stay on long enough to answer the COVID questions near the end of the survey. Those that reach that point are unlikely to break off, meaning the breakoff rate for that group will be quite low. Therefore, comparing the breakoff rate between treatments among the COVID-unconcerned (or the COVID-concerned) bias the results in favor of our hypothesis (a higher breakoff rate for the COVID-First group). We avoid this problem by adopting the more conservative approach of keeping the NA cases (nonresponders) in the analysis and excluding only those that think coronavirus is not a serious issue.

⁸We did assess the data using a more inclusive definition of breakoff (analysis not shown in paper), and the results do not meaningfully change.

⁹We 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.

¹⁰We also examined interview duration – the actual amount of time spent on the phone – as an alternative outcome variable. Results were similar. Among those who broke off, those in the COVID-First group stayed

on 17 seconds longer on average, though the difference is not statistically significant. Among complete interviews, the COVID-First group also had a slightly longer average interview duration (19 seconds; again not significant).

¹¹Full questionnaires are available upon request.

¹²We do not include cases in which there is a non-response for the COVID seriousness question, explaining why the results may differ from those described above and why those shown in the table do not average to the overall breakoff rate. The reason for the widely varying rate of breakoff is not immediately clear. We suspect it may have to do with experience with social scientific surveys among firms, interviewers, and respondents, or possibly other factors like the socioeconomic and demographic makeup of each country, cultural differences, or the timing of the surveys with respect to COVID-related lockdowns. This wide range of breakoff rates, though, is in line with meta-analyses of breakoffs (Lozar-Manfreda and Vehovar 2002; Mavletova and Couper 2015; Musch and Reips 2000), so we leave this question as an opportunity for future research.

¹³These throwaway questions are sometimes suggested as a means to boost cooperation rates in guides like this one: <https://www.qualtrics.com/experience-management/research/question-sequence-flow-style/>

Appendix

A: Haiti Experiment COVID-19 / Pandemic Question Module

The COVID-19 Module consists of 10 questions. Those in the COVID-First Condition were randomly assigned to be asked these 10 questions at the start of the questionnaire, following a short module of screening questions (age, citizen, gender, department, municipality (commune), and urban vs. rural place of residence). The Module is presented here in English; the survey was conducted in Haitian Creole.

I. Concern (5 questions)

- **COVID1.** How serious of a problem do you think the coronavirus outbreak is for Haiti?
 - (1) Very serious (2) Somewhat serious (3) Not so serious (4) Not serious at all (5) Have not thought much about this
- **COVID2A.** How worried are you about the possibility that you or someone in your household will become sick with the coronavirus?
- **COVID2B.** How worried are you about the possibility that your household economic situation will be negatively affected by the coronavirus?
- **COVID2C.** How worried are you about the possibility that your household will have difficulty accessing basic goods, such as food or medicine, due to the coronavirus?
 - (1) Very worried (2) Somewhat worried (3) A little worried (4) Not at all worried
- **COVID3.** How would you rate the performance of the national government in handling the coronavirus outbreak?
 - (1) Very good (2) Good (3) Neither good nor bad (fair) (4) Bad (5) Very bad

II. Responsibility (1 question)

- **COVID4.** Coronavirus is spreading in Haiti. Who is the most responsible for the increased number of infections? [Open-ended; interviewers recorded one response]

III. Behavior (3 questions)

Please tell me if you have taken any of the following actions because of the recent coronavirus outbreak. [Randomize items]

- **COVID6.** Have you canceled any plans to attend large gatherings such as social events or church?
- **COVID7.** Have you kept a greater distance between you and others when out in public?
- **COVID8.** Have you washed your hands more often with water and soap or sanitizer?
 - (1) Yes (2) No

IV. Stigma (1 question)

- **COVID14.** Considering those who get the coronavirus, would you say that they deserve to be sick because of the way they live, or that they could not help becoming sick?
 - (1) Yes, they deserve to be sick (2) No, they could not help becoming sick

List of Tables and Figures

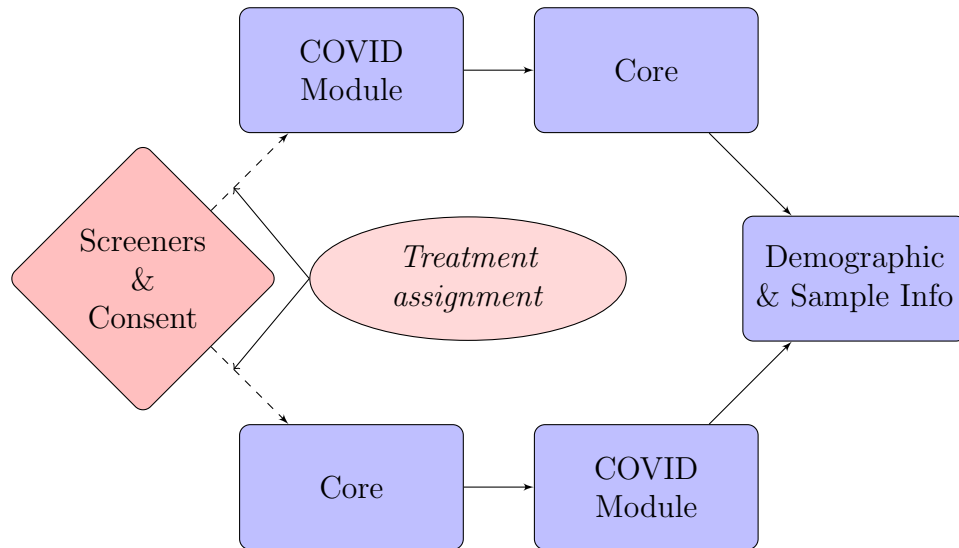


Figure 1: Questionnaire Structure

Table 1: Breakoff Rate (%), by Question Order Treatment

	Full Sample	Limited Sample
COVID Questions First	8.6 (0.8)	8.7 (1.0)
COVID Questions Late	10.3 (0.9)	11.8 (1.0)
<i>Net Difference</i>	-1.7 (1.6)	-3.1 (2.2)
<i>p</i> -value	0.16	0.09
<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 and are clustered by interviewer. *p*-value is one-tailed.

Table 2: Average Number of Questions Asked Before Breakoff, by Question Order Treatment

	Full Sample	Limited Sample
COVID Questions First	18.9 (1.6)	18.4 (1.8)
COVID Questions Late	17.0 (1.4)	16.1 (1.4)
<i>Net Difference</i>	1.9 (1.5)	2.3 (2.1)
<i>p</i> -value	0.11	0.15
<i>n</i>	203	177

Note: Standard errors in parentheses. Standard errors for “Net difference” are calculated from two-sample independence of means *t*-test and are clustered by interviewer. *p*-value is one-tailed.

Table 3: Item Non-Response Rate (%), by Question Order Treatment

	Full Sample	Limited Sample
COVID Questions First	11.9 (0.5)	12.6 (0.6)
COVID Questions Late	12.1 (0.4)	12.6 (0.5)
<i>Net Difference</i>	-0.2 (0.7)	0.0 (0.9)
<i>p</i> -value	0.37	0.52
<i>n</i>	2,279	1,829

Note: Standard errors in parentheses. Standard errors for “Net difference” are calculated from two-sample independence of means *t*-test and are clustered by interviewer. *p*-value is one-tailed. 2290 subjects were assigned to a treatment group (1840 in the limited sample), but 11 broke off on the first question, leaving an *n* of 2,279 (1,829 in limited sample).

Table 4: Results for Haiti Experiment Excluding End Block

	Breakoff Rate (%)		Qs Asked Before Breakoff		Item Non-Response Rate (%)	
	Full Sample	Limited Sample	Full Sample	Limited Sample	Full Sample	Limited Sample
COVID Qs First	8.0 (0.8)	8.0 (0.9)	15.9 (1.2)	14.8 (1.4)	16.4 (0.6)	17.5 (0.7)
COVID Qs Last	9.7 (0.9)	11.3 (1.0)	14.3 (0.9)	14.0 (0.8)	16.8 (0.5)	17.2 (0.6)
Net Difference	-1.6 (1.5)	-3.2 (2.2)	1.5 (1.4)	0.8 (1.8)	-0.4 (0.8)	0.3 (1.0)
p-value	0.15	0.09	0.15	0.33	0.39	0.36
<i>n</i>	2,229	1,790	190	167	2,279	1,829

Table 5: Results from Ecuador Experiment

	Breakoff Rate (%)	Qs Asked Before Breakoff	Item Non- Response Rate (%)
COVID Qs First	0.7 (0.4)	33.6 (8.3)	3.5 (0.2)
COVID Qs Late	2.4 (0.7)	30.7 (6.9)	3.7 (0.3)
Net Difference	-1.7 (0.8)	3.0 (12.1)	-0.2 (0.3)
p-value	0.03	0.41	0.29
<i>n</i>	835	13	864

Table 6: Breakoff Rate by Country and Rating of COVID Seriousness

	Haiti	Peru	Mexico	Ecuador
Very serious	6.83% (1.10)	0.48% (0.17)	3.53% (0.49)	0.61% (0.43)
Somewhat serious	5.38% (1.99)	1.43% (0.72)	4.51% (1.04)	1.96% (1.96)
Not so serious	4.52% (1.57)	4.54% (4.55)	8.03% (2.58)	0%
Not at all serious	12.12% (5.78)	0%	11.11% (5.31)	0%
Haven't thought much about this	19.23% (5.52)	0%	4.70% (2.31)	0%
<i>Logit Coefficient</i>	-0.20	-0.28	-0.20	0.05
<i>(p-value)</i>	(0.04)	(0.296)	(0.029)	(0.945)
<i>n</i>	919	2,099	2,076	408

Notes: Data from Haiti and Ecuador excludes those who received “COVID-Last” treatment condition. Top five rows display breakoff rate among each country and rating of seriousness of coronavirus outbreak in the country. Standard errors are in parentheses. The last row shows coefficients of logistic regressions of breakoff (0/1) on COVID concern level (interval, 1-5), with p-value in parentheses. The COVID seriousness variable was heavily skewed toward the more serious end, explaining the large standard errors among the less serious categories.