

Remote Surveying in a Pandemic: Handbook¹

Executive Summary

Innovations for Poverty Action (IPA) is a research and policy nonprofit that discovers and promotes effective solutions to global poverty problems; since our founding in 2002, IPA has worked with over 600 leading academics to conduct over 830 evaluations in 51 countries. As part of IPA's response to COVID-19, many existing and new data collections have shifted to remote data collection modes including computer-assisted telephone interviews (CATI), interactive voice response (IVR) and SMS surveys.

This transition has required new protocols, new tools, and new workflows to ensure that data can be collected at similarly high-quality across remote survey modes. This is compounded by the logistics and health challenges associated with a global pandemic. This document contains tips and best practices for shifting face-to-face surveys to remote survey modes, as a response to pandemic conditions where person-to-person contact risks virus transmission.

IPA's Global Research and Data Support (GRDS) team has created technical tools and protocols that fall into four categories of tasks for remote surveying. We summarize the major changes and available tools for CATI, the predominant choice for remote data collection mode, in this document according to these tasks:

- **Data collection**
 - [Deciding when to delay data collection](#) is the first step to implementing a study. The transition to remote survey modes may require rethinking what a data collection can measure.
 - [Deciding an appropriate remote survey mode](#) is a direct result of the different characteristics of various modes.

¹ This document was prepared by Steve Glazerman, Michael Rosenbaum, Rosemarie Sandino, and Lindsey Shaughnessy of Innovations for Poverty Action. This document incorporates many thoughts, and owes its name, from the crowd-sourced *Phone Surveying in a Pandemic* document. We reserve special thanks on that document to Ala' Alrababa'h and the team at the Immigration Policy Lab, Amanda Beatty at Mathematica Policy Research, Charles Lau at RTI, Rohit Nainpally at J-PAL, Chris Robert at Dobility, and Rachel Steinacher at IPA as well as many other commenters on the document including Hasina Badani, Ellen Bates-Jefferys, Willie Blackmon, Ashraf Haque, Erik Jorgenson, Prathap Kasina, Sarah Kopper, Yuna Liang, Pat Malone, Samuel Kembou, Teresa Martens, Phillip Okull, María Juliana Otalora, Laura Steiner, Dayana Lorena Téllez Galeano, Jorge Luis Tipisman, Shana Warren, Rachel Wells, Zin Nwe Win, and Arsene Zongo.

- **Measurement**
 - **IRB and Data Protection** protocols will need to be adjusted as the location of data storage and frequency of data transfers increases with remote modes.
 - **Questionnaire Design** is dependent on the remote data collection mode, and will likely require substantive changes to the size and content of a questionnaire designed for a face-to-face survey.
 - **Sampling** is an important concern for representativeness of the survey due to reduced response rates among existing and new survey populations compared to face-to-face data collection.
 - **Infrastructure** used for face-to-face data collection may be inadequate for conditions in a pandemic such as lock down, or to increase interview fidelity over imperfect cell networks.
- **Field logistics**
 - **Protocols** require a different set of standards for setting up and conducting virtual phone banks compared to either physical phone banks for face-to-face surveys.
 - **Interviewer training** is complicated by remote implementation and a different set of tools and survey topics than many interviewers are used to.
- **Data quality**
 - **Tracking and case management** require different tools and connectivity levels for managing contact attempt tracking and case assignments for interviewers who are working in their own homes.
 - **Monitoring** data quality and interviewer performance requires checking a different set of indicators and can be supplemented by more information from dialing procedures.

These categories do not fully or exhaustively capture learnings and protocols for remote survey modes during COVID-19. They focus on the concerns that we expect most data collection efforts to face.

This handbook was compiled from a wide variety of sources and contributors. In many ways it will remain a work in progress as we continue to learn methods to increase the fidelity of data collection across remote survey modes. We invite clarifications and updates by [email](#).

Table of Contents

<u>Introduction</u>	5
<u>Choosing to Continue Data Collection during COVID-19</u>	6
<u>Choosing an Appropriate Mode</u>	7
<u>Research Goals</u>	8
<u>Characteristics of Remote Survey Modes</u>	9
<u>Recommended Tools</u>	10
<u>Specific Considerations by Mode</u>	10
<u>IRB and Data Protection</u>	13
<u>Interviewer Data Security</u>	14
<u>Informed Consent and Respondent Privacy</u>	14
<u>Links to Consent Forms and Additional Resources:</u>	15
<u>Questionnaire Design</u>	16
<u>Survey Content</u>	16
<u>Data Quality</u>	17
<u>Sampling</u>	20
<u>Existing Samples</u>	20
<u>New Samples</u>	20
<u>Infrastructure</u>	22
<u>Equipment</u>	22
<u>Software</u>	23
<u>Protocols</u>	25
<u>Physical Set-up</u>	25
<u>Call Scheduling</u>	25
<u>Coding Call Statuses</u>	27
<u>Collecting Contact Information</u>	28
<u>Incentives</u>	29
<u>Interviewer Training</u>	31
<u>Interviewer Recruitment</u>	31
<u>Training Preparation and Prerequisites</u>	31
<u>Training format and engagement</u>	31
<u>Training content</u>	32
<u>Training quizzes</u>	32

<u>Tracking & Case Management</u>	<u>34</u>
<u>Respondent Contact Information</u>	<u>34</u>
<u>Daily Workflows and Call Assignment</u>	<u>34</u>
<u>Protocols for Special Cases</u>	<u>35</u>
<u>Reporting</u>	<u>35</u>
<u>Monitoring</u>	<u>37</u>
<u>High Frequency Checks</u>	<u>37</u>
<u>Backchecks</u>	<u>38</u>
<u>Sensor Data and Audio Audits</u>	<u>38</u>
<u>References</u>	<u>40</u>

Introduction

COVID-19 has resulted in many changes, including in survey implementation. As the pandemic has stretched on, many of IPA's projects have transitioned to remote survey modes. IPA also launched its Research for Effective COVID-19 Responses ([RECOVR](#)) initiative in an effort to generate rigorous evidence through remote surveys, advise governments, and bring together partners across the research-to-policy sector. IPA's work has continued during the pandemic to ensure that relevant evidence from survey data can be delivered to policymakers and researchers, including to rapidly provide decision-makers with rigorous data and evidence to mitigate the impacts of the crisis.

Ensuring that this survey research continues to produce high-quality data during a pandemic is vital. Inaccurate measurement can have profound consequences for the poorest individuals. It goes without saying that data collected remotely is currently informing many social protection programs, responses to the pandemic, and targeting from NGOs. If these data result in inaccurate conclusions or have increased measurement error, there is a real human cost.

Quarantine, lockdowns, and physical distancing have complicated IPA's evaluation work, but also resulted in a broad set of new protocols, tools, and best practices around safe and secure remote surveying. Although IPA has experience with conducting high-frequency remote survey implementations, the constraints imposed by COVID-19 on standard tasks have resulted in new necessary protocols, such as delivering tablets to interviewers and conducting trainings remotely.

This document summarizes the lessons that IPA has learned about high-quality data collection through more than 50 remote survey implementations. It focuses on ten different tasks for survey preparation and implementation. These tasks are loosely ordered following a standard survey implementation would follow. We categorize these tasks in four domains:

- **Data collection** includes the key decisions that investigators have to make about conducting survey research in a pandemic such as [deciding when to delay data collection](#), and if you do go ahead with remote data collection, [choosing an appropriate remote survey mode](#).
- **Measurement** includes pre-implementation planning including [IRB and data protection](#) adjustments, and changes in [Questionnaire Design](#) and [Sampling](#) as a result of remote modes. Finally, it also discusses the physical and digital [infrastructure](#) necessary for implementing remote data collection with fidelity.
- **Field logistics** includes [protocols](#) for survey implementation as well as extra components and structures for remote [interviewer training](#).

- **Data quality** includes changes to procedures for [tracking and case management](#) as well as suggested changes to more effectively [monitor](#) data quality and interviewer performance.

Many of IPA's projects have converted to phone surveys using CATI rather than other remote survey modes. As a result, much of this document is based on IPA's experiences with CATI data collections. It also concentrates on tools developed using SurveyCTO. Many of these lessons are generalizable to remote interviewer administered phone surveys, and even to other modes. However, specific details in this document focus on phone surveys using SurveyCTO. We present information on other modes, but cannot advise on IVR and SMS with the same level of detail as CATI.

Choosing to Continue Data Collection during COVID-19

Before shifting from face-to-face data collection to mobile phone-enabled interviewing, consider the potential data quality costs compared to delaying data collection. Although the key question of "when would a delayed data collection happen?" cannot be answered with certainty in the COVID-19 context, IPA recommends switching the primary mode of data collection to a remote mode only when data is needed immediately, and that data can be collected in a safe and ethical way.

Reasons for needing data immediately are likely obvious: timeliness of data collection close to an intervention, seasonality of outcomes, capturing COVID-19's effects, or external deadlines. However, limitations in remote data collection modes may mean that data collection is not feasible. Evidence on the efficacy of mobile phone-enabled data collection in developing countries find that response modes substantively [affect representativeness, response rates, and precision](#), although these effects vary in different contexts.

Practical considerations are also relevant. Consider the logistics, feasibility, and ethics of surveying your sample remotely:

- Feasibility
 - Do you have or can you get reliable phone numbers for each respondent?
 - Do respondents have phones with airtime, power, and signal?
- Logistics
 - Is it possible to compensate respondents for their time, e.g. by transferring airtime minutes or mobile money?
 - Can you hire, train, and monitor interviewers remotely?
 - Can interviewers submit data over the internet with fidelity and in a timely manner?

- Ethics
 - Do you have permission to re-contact respondents (check the original consent language and local laws)?
 - Does the mode change put respondents at risk due to the content of the survey or compensation mode?

Some of these considerations can be ameliorated through effective surveying protocols such as compensation for airtime and sourcing contact information from central figures in villages, but many cannot. Inaccurate data is worse than no data: incorrect conclusions also cost money.

Remote data collection can supplement face-to-face survey data collection. In the COVID-19 context, study populations may drastically alter behavior and may increase attrition for face-to-face follow-ups. Intervention monitoring, contact information collection, and information dissemination can certainly be done remotely and may reduce attrition for an eventual face-to-face follow-up.

Choosing an Appropriate Mode

Once you have chosen to pivot to remote data collection, there are other study design decisions to make. This section describes mode choice and sampling, taking into account cost, timeline, availability of interviewers, and other factors.

When projects cannot delay surveys, IPA recommends computer-assisted telephone interviewing (CATI) for most projects due to data quality. Although interactive voice response (IVR), SMS, and Web surveys can be more scalable and less expensive, they can also be less representative and make it harder to elicit honest responses to many question types compared to live interviewing.

Mode choice is not necessarily an either-or decision. Mixed-mode data collection can be used to capture the best properties of each mode. For instance, live interview (CATI) baseline surveys can be supplemented with high-frequency follow-up by SMS or IVR once the respondent has been initiated into the study. Also, if automated survey modes are used for a large-scale survey effort and response rates are very low, then phone followup with a randomly selected

Remote Data Collection Modes

Computer-assisted telephone interviewing (CATI): any type of telephone survey using computerized data collection

IVR: recorded surveys that respondents self-administer using keystrokes or voice recognition

SMS: self-administered surveys delivered over SMS or other messaging platforms such as WhatsApp

Web: self-administered surveys on an internet platform such as Qualtrics

subset of nonrespondents can be used to characterize the nature of nonresponse and to construct adjustment weights.

Research Goals

Different remote data collection modes are better suited for accomplishing different research goals. The table below provides a brief overview of goals and how different modes may perform. We describe common goals where a **green +** means the mode is well suited to that research goal, a **yellow ~** means that it depends on other factors, and **red -** means it is not well suited for this goal:

Table 1: Remote Survey Modes and Research Goals

	CATI	IVR	SMS	Web
Tracking respondents	~	+	~	-
Updating contact information	+	-	~	~
Determining language	+	+	-	+
High-frequency data collection	~	+*	+*	+
Measuring sensitive outcomes	~	~	-	~
Achieving high response rates	+	-	-	-
Accumulating a large sample	-	+	~	+
Comparable attrition to face-to-face	-	-	-	-

* mode limitations change the eligible set of outcomes that can be collected.

Note: this table is a general guide to research tasks that we expect in the standard implementation of survey modes. This table will not be accurate for all implementations, but hopes to provide some guidelines to identifying which mode is appropriate for a given measurement task.

Characteristics of Remote Survey Modes

Before you adapt or write a survey for remote survey data collection, it's useful to consider limitations for survey design based on the mode, as well as costs that you may pay in data quality for moving away from a face-to-face mode. IPA [reviewed prior literature](#) on a

number of characteristics including cost, response rate, and representativeness. These results are summarized below, but are presented in more detail and have [source data](#) available.

Changes to the sampling frame, sample generation process, and contents of the survey will affect these guidelines. For example, a more cognitively or emotionally demanding questionnaire will reduce the effective length of the questionnaire.

Table 2: Remote Survey Mode Characteristics

	CATI	IVR	SMS
Length			
Length (Suggested)	<20 minutes	5-10 items	5-10 items
Length (Max)	30 minutes	15 items	15 items
Cost (Price/complete)			
	\$11.97 (\$8.78)	\$4.86 (\$5.40)	\$7.75 (\$1.71)
Response Rate			
New sample	32.8% (31.9%)	11.8% (14.3%)	4.5% (6.6%)
Existing sample	56.0% (18.9%)	16.8% (4.6%)	20.4% (1.8%)
Representativeness <i>In percentage points compared to a representative sample</i>			
Gender (Male)	+20	+19	+7
Urban	+27	+16	+36
Age (under 35)	+14	+12	+17
Education (Secondary)	+24	+7	+14

Recommended Tools

To deliver these surveys, IPA has tested a number of platforms for remote surveying modes. We only consider those platforms that have met our requirements around data

security and functionality. We also have recently tested, but not fully evaluated, new platforms for deploying remote surveys; they are also listed below.

Table 3: Recommend Survey Platforms by Remote Survey Mode

	CATI	IVR	SMS	Web
Recommended Platforms				
Supported by IPA	SurveyCTO CATI Toolkit			SurveyCTO Web
Tested by IPA	Twilio ; Exotel	Viamo ; engageSPARK	Twilio ; SampleSolutions	Qualtrics

Specific Considerations by Mode

CATI

CATI surveys substitute well for CAPI (computer-assisted personal interviews), but replacing longer modules may affect study and questionnaire design, as well as logistics:

- Statistical power may be reduced compared to a face-to-face survey.
- Survey compensation costs should factor in explicit costs in airtime and power due to delivering the survey, as well as the opportunity cost of time.
- During COVID-19, hiring, training, and providing resources to interviewers is an added logistical hurdle that may reduce fidelity of survey implementation or increase data security tasks.
- Due to the remote work, additional monitoring for data quality should be conducted, specifically around respondent identification, call completion, and data.
 - Consider increasing backcheck rates beyond 25 percent, and including non-connected calls in the backcheck sample.
- Ethical considerations around respondent time and use of resources such as electricity and airtime in a pandemic should be considered.
- Rates of partial response may increase as frictions to leaving post-consent are much lower.
- Perceptions of privacy and trust may be affected by the change in mode. This may increase responses to sensitive questions due to perceived anonymity or reduce responses due to increased social distance.
- Surveys delivered by phone have the advantage of being administered by an interviewer—including the ability to ask for clarification, and other skill-related benefits.

Although CATI approaches are the “next best” mode for many projects, existing literature may not be a credible guide for response rates or representativeness, even when ignoring the COVID-19 context. Many studies that plan CATI data collections provide mobile phones

and SIM cards to respondents at the onset of the study and tie cash transfers or survey compensation to the SIM card ([Suri et. al., forthcoming](#); [Dabalén et. al., 2016](#)).

Attrition rates from other samples will be substantially higher without consistent compensation to update the survey team when contact numbers change. These incentives can be added to a project at any point.

IVR

One benefit IVR provides is much lower cost, due to the lack of interviewers. However, qualitative and anecdotal research on the form suggests a variety of design considerations for any intervention:

- Statistical power may be further reduced compared to a CATI or face-to-face survey in part due to increased measurement error.
- Ensuring that the correct respondent is answering the survey must be automated in an IRB compliant way.
- Ethical considerations around respondent time and use of resources such as electricity and airtime in a pandemic should be considered.
- IVR systems may frustrate respondents.
- Touch tone response interpretation [is a skill \(that can be learned\)](#).
- Voice response may not be usable for respondents due to cell connectivity and voice recognition on the IVR platform.
- Extensive piloting on question phrasing is necessary, with piloting occurring using multiple modes.
- Spotty cell phone signal can disrupt responses, and may be coded as a partial response instead of a non-connect depending on the IVR service.
- IVR systems may not accurately code non-connects.
- There is suggestive evidence of [order effects in response options](#).
- Perceptions of privacy & trust may be affected by the change in mode. This may increase responses to sensitive questions due to perceived anonymity or reduce responses due to increased social distance.
- Clarification on question meaning is harder for respondents, outside of repeating the question. Soft checks are encouraged to reduce measurement error.
- Presentation of recorded voice may change how respondents respond, depending on factors like recording quality or how computerized the recording sounds.
- Your options may be limited by the number of languages or dialects required for your survey.

SMS

SMS interventions share many of the same limitations and benefits of IVR, but also have added length and literacy limitation:

- Statistical power may be further reduced compared to a CATI or face-to-face survey in part due to increased measurement error.
- Ensuring that the correct respondent is answering the survey must be automated in

- an IRB compliant way.
- Ethical considerations around respondent time and use of resources such as electricity and airtime in a pandemic should be considered.
- There is anecdotal evidence that fatigue due to volume of COVID-19-related SMS may reduce response rates in certain areas.
- SMS length constraints make it harder to determine language for survey introductions and run a risk for increasing non-response.
- Survey design should avoid [select multiple items, “don’t know” response options as well as indicate substantive ordering effects](#).
- Items must be less than 160 characters, as receipt order is not guaranteed.
 - Take this into account when designing your consent process.
- Extensive piloting on question phrasing is necessary, with piloting being done across multiple modes.
- Responses leave a record on respondents’ devices, which may increase risk for some respondents.

Web

Web surveys are less commonly used in low and middle-income countries (LMICs). There is an extensive literature on design and implementation of self-administered web surveys. Here, we concentrate on specific considerations for web surveys in LMICs that will likely be taken on smartphones:

- Ensuring that the correct respondent is answering the survey must be automated in an IRB-compliant way.
- Compensation should include data costs as well as the opportunity cost of time.
- Question and response formatting should be targeted towards mobile devices. Even smartphones may have small screens used vertically.
- Typed responses should have soft checks, accessible formatting, or confirmations due to difficulties in using mobile keyboards.
- Interruptions in page submissions due to coverage and data limits may interfere with system timers.
- Images and videos may load differentially across respondents as a function of data speed, cell service, and computer ownership.

IRB and Data Protection

This section is written for research studies governed by [IPA's Institutional Review Board](#) (IRB), but the principles should apply to any human subjects research governed by a different IRB. Check with your own study's IRB to set expectations about how to document protocols that mitigate the risks of virus transmission.

IRB protocol changes

IPA's IRB has provided specific guidance for IPA studies that will transition its data collection mode during COVID-19. Two types of amendments are commonly expected from studies that decide to shift data collection mode from face-to-face surveying to phone surveying:

- Adding staff to studies to allow production-smoothing, e.g. by giving data cleaning work to RAs who need coverage while their field studies are halted.
- Changing questionnaires to adapt to phone mode; any changes to consent or assent procedures should be flagged.

Some minor study changes do not require amendments, but instead can provide bundled amendments or a progress report at the time of study renewal with the IPA IRB. If a study anticipates changes to its existing protocols, surveys, and/or consent forms, amendment submissions WOULD NOT need to be submitted if:

- The study is only shifting its protocol to reflect a data collection mode shift from face-to-face survey to phone survey.
- The only informed consent script change reflects the study's data collection mode shift from face-to-face survey to phone survey and minor resulting procedural shifts.
 - If the project has not yet already obtained consent from respondents, then the research team can use the approved consent form and document consent over the phone.
- The only questionnaire change is eliminating questions.

Amendment submissions WOULD need to be submitted if:

- The project is federally funded (even for minimal changes).
- Questions are edited on, or being added to, the questionnaire,
- There are additional changes to consent forms.
- The incentive changes.
 - Example: The project team wants to move from a bag of sugar or a bar of soap to an electronic gift such as airtime or a gift card. However, if the



electronic gift does not match the cost of the tangible object, then there's potential for the respondent to respond differently. Additionally, the IRB is concerned about the electronic incentive going to the right person.

For IPA studies with specific questions because of the nature of the study, please email humansubjects@poverty-action.org. IPA's IRB recognizes that there will be situations where the study's questionnaire or other procedures may be impacted by COVID-19 that won't fall into the streamlined amendment category.

Interviewer Data Security

The unique challenges of COVID-19 requires balancing health risks with data security risks. IPA's IRB has provided some guidance on potential changes to data security on remote phone banks:

- **Studies should avoid having interviewers use their personal devices for calling respondents and sending them SMS messages.** You should consider health risks and make the best and safest possible effort to provide IPA devices for calling (such as a SIM-card enabled IPA tablet, an IPA tablet with data and an internet connection for voice-over-IP calls, or an IPA-issued mobile phone or SIM card) before resorting to using interviewers' devices. If none of these is possible, then you may rely on interviewers' personal devices, while using an IPA tablet for data collection. If possible, use a masking service on the interviewers' device to mask phone numbers, such as Exotel.
- **interviewers cannot use their personal devices for entering survey data.** Discuss with the IRB early in the process if a government lockdown prevents you from delivering IPA devices to interviewers. For any case where personal devices are used to collect survey data, your IRB application or amendment should detail additional procedures to ensure strong password protection and wipe protected data from their devices at the end of data collection.

More details can be found in this [FAQ](#), including a detailed protocol for device security.

Informed Consent and Respondent Privacy

Refer to IPA's standard [informed consent checklist and template](#) to generate or modify your consent forms. You may need to revise a typical informed consent script for a face-to-face research project to reflect remote surveying protocols and/or the realities of surveying in a pandemic; some of those changes are outlined in this [phone survey consent checklist](#). Some of these revisions might include:

- Asking whether the respondent is in a location where they can privately answer the questions.



- Describing how the incentive works and will be transmitted.
- Requesting consent to audio record the conversation.

Ensure that your consent form follows best practices. One common problem, especially relevant for CATI surveys, is understandable language around data security of audio recording.

Links to Consent Forms and Additional Resources:

- Example consent forms:
 - [CATI](#): an example consent form used for a fielded CATI survey
 - [Web](#): an example consent form used for a fielded web survey
- [Phone survey consent checklist](#): a checklist of items to include in a CATI survey consent form
- [IPA IRB FAQ](#): frequently asked questions from IPA's IRB on when and how to submit IRB amendments for changes that involve switching response modes to a remote survey mode

Questionnaire Design

Many principles of questionnaire design for face-to-face surveys apply for phone surveys. This section describes two areas to consider around questionnaire design when transitioning from an existing face-to-face survey to a phone survey: survey content and data quality.

Survey Content

There are three ways that survey content should be adapted when moving from CAPI to phones:

- Reduce survey duration.
- Allow time to confirm the respondent's identity.
- Create a language reassignment protocol

Survey Duration

Surveys over the phone, especially automated surveys, should not be as long as traditional face-to-face surveys. Breakoffs or loss of concentration can occur after 20 minutes. Surveys of more than 30 minutes may be challenging for respondent attention, even if respondents remain on the line (Suri, 2020). One hour should be the upper limit. Although surveys can be conducted and completed with modules longer than 30 minutes, data quality may suffer in ways that are hard to notice or quantify.

When longer forms are deemed necessary, consider conducting interviews over two sessions, with adequate incentives attached to each session. The informed consent should announce the expected or typical duration of the survey at the outset, so it can be rescheduled if necessary. As part of this design process, it's important to ensure that the questionnaire is structured to reduce attrition. One approach is to set up a tracking system with appointment-making that allows you to save and resume. This especially helps for respondents with spotty cell service/Internet or with dying batteries.

Velthausz et al. (2015) found strong resistance to an approximately 18-minute phone survey with modules on phone and internet use, with respondents complaining that the questionnaire was too long and questions were "invasive" at times (Velthausz et. al., 2015). These questions addressed mobile financial and communications products as well as general demographics.

Phone surveying makes it possible to increase the frequency of measuring a smaller set of outcomes. High frequency surveys can measure outcomes that change rapidly over time, like time use, dietary intake, mental health symptoms, or consumption which may not be feasible in face-to-face surveys. They substitute for long and detailed modules which do not vary day-to-day or week-to-week such as asset ownership.

Respondent Identification

Remote surveying also provides challenges to ensure that the correct respondent is identified. Even if the respondent does not pick up a phone, the respondent may be reachable through passing the phone or collecting contact information of family members or neighbors. For panel samples, ensure that a questionnaire collects multiple contact numbers including non-respondent individuals who may be able to contact the respondent.

It is also important to verify that the respondent is located in a private place where they could complete a survey. This will increase respondent privacy and data quality. Remember to confirm with the respondent whether they are in a location where they can privately answer the questions. For surveys that contain sensitive items, include a protocol to reschedule the survey if the respondent cannot respond to sensitive items with sufficient privacy.

Language Reassignment

Compared to face-to-face surveys, it is possible to build language reassignment directly into the questionnaire. In RDD samples, this can delay or result in low survey completion rates. If an interviewer or mobilizer determines the language, automatic routing of cases can result in call backs by interviewers who speak the relevant language within a short amount of time. Although virtual phone banks delay this process, SurveyCTO's case management allows [integrated case reassignment](#) within the form allowing for language reassignment to be more automated in settings with many languages.

Data Quality

Remote surveying also raises concerns about data quality. The main concerns are that the sample will not be representative and the quality of responses suffer since interviewers cannot use body language and other nonverbal cues to establish rapport and trust with respondents or to keep them engaged.

Representativeness

Samples obtained by phone may be unrepresentative for several obvious reasons: response rates tend to be lower and individuals (more often women than men) and

households (more often poor and rural than wealthier or urban) without phones or with limited power or airtime will be under-represented compared to face-to-face surveys. In developing countries, sample composition often skews urban, younger, and more highly educated by substantive margins (Heath et. al., 2017; Lau et. al., 2019; Kastelic & Kastelic 2015; L’Engle et. al., 2018; Leo et. al., 2015; Bloomberg Philanthropies Data for Health Initiative: Zambia NCD Mobile Phone Survey, 2020; Bloomberg Philanthropies Data for Health Initiative: Philippines NCD Mobile Phone Survey, 2020; Lau et. al., 2019a). If ground truth data exists, compare sampled individuals on various dimensions to these data by asking the same questions. Otherwise, consider using demographic question formats that are exactly comparable to a recent nationally representative survey so that representativeness information can be accurately estimated. Age, gender, education, and urbanicity are especially relevant questions. However, items on education and urbanicity should be aligned directly to question phrasings in the ground truth data.

Attention Checks

Data quality may also be affected by reduced attention in remote surveys. Over the phone, respondents may have distractions like children or other family members needing attention, chores or driving, or just street or animal noises. Although CATI surveys are not self-administered, interviewers have reduced ability to notice respondent fatigue, distractions, and effort over the phone.

Additional checks can be included to try to explicitly measure when and how much attention loss occurs during the survey. In interviewer administered surveys, attention check questions can be included. These ask the interviewer to record their observations of respondent attentiveness at various points in the survey, based on unusual patterns of background noise, pauses, or requests to repeat questions. These questions are only completed by the interviewer. They are not asked of the respondent. We recommend placing them during module transitions, where completing these items will not affect the flow of the survey. Asking verbal questions to estimate respondent attention can provide additional information, but results in a more explicit tradeoff between time spent on the survey content and time spent measuring data quality

In self-administered surveys, Instructional Manipulation Checks (IMC) serve a similar role. An IMC is a device that you can include in self-administered questionnaires to make sure respondents are paying attention and reading instructions, if applicable (Oppenheimer, Meyvis & Davidenko, 2009). To include an IMC:

- Design a question with a correct answer that is obvious to respondents if they read the instructions, but also an incorrect answer that is obvious/tempting if they skip the instructions.

- Program the survey to require a correct response (with a reminder to read the instructions).

IMCs can prompt respondents to pay greater attention, which will improve response quality or they can be used to screen out respondents who don't or won't read instructions, or a combination of both. There is a long history of IMC usage, including some evidence suggesting effectiveness may be design or context dependent (Liu & Wronski, 2018; Curran, 2016; Vanette, 2016).

Sampling

We consider sampling frames for phone surveys that are either new (“cold calls”) or from pre-existing lists. The source of the sampling frame will influence response rates and representativeness.

Within this section we provide advice on contacting existing samples, as well as considerations for using new samples.

Existing Samples

Using phone numbers gathered in earlier rounds is an ideal way to conduct a phone survey because you have a well-defined population that you already selected. Be sure that the way you intend to use the numbers is consistent with the consent you gathered when collecting them. See this document’s [IRB](#) section for further guidance.

Existing samples may be less representative as different subpopulations vary in their access to cell phones, electricity, or transience on contact numbers. Data that contains multiple contact numbers and/or contact information from friends, family, or village authorities can be a useful way to recover these contact information.

Short CATI surveys or even SMS surveys can be used to update this contact information to ensure that delayed face-to-face follow-up surveys can be successfully implemented without large attrition rates. In the COVID-19 context, internal migration may be higher, and as a result, contact information may change. Incentivizing respondents to update contact information with the study can be a strategy for many existing sampling frames.

New Samples

Approaches to create new samples for remote survey modes often trade off cost and statistical power for representativeness. Most costs come from increased airtime and interviewer time in screening deactivated numbers or dialing unresponsive numbers. There are two primary approaches for developing a new sampling frame for phone surveys remotely:

- **Unscreened Random Digit Dialing (RDD):** Numbers are randomly generated in alignment with target countries’ mobile number formations and subsequently contacted. While RDD tends to be representative of those with working phones, it inevitably generates a large proportion of ineligible numbers such as non-connected and non-residential numbers. This requires more call attempts for the same size sample.

- **Screened RDD: Mobile Network Operator (MNO):** A list of active numbers is procured through a Mobile Network Operator, generally through the services of a third party or the MNO. These numbers are prepulsed--checked for activity --by the telecom and generally are approximately randomly collected from the universe of telecom customers. For MNO to be representative, the sample must reflect the MNO's subscriber share, which depends on the third party or vendor and their relationship with all national operators. Although number lists are supposed to be up-to-date with active numbers, slightly outdated lists may include disconnected numbers and more seriously, be missing new numbers, which could result in a biased sample composition especially among migrants, younger individuals, and people who have recently experienced a positive income shock.

RDD may present additional data collection challenges in multilingual settings. Interviewers either need to be multilingual or deliver a language screener. [Mobilizers](#) can be used to language screen all respondents and then route cases to a language-matched interviewer, who remains on standby, if budget is available, or calls back at a later time.

Infrastructure

Tools and infrastructure can vary across remote survey contexts. This section describes suggested digital and physical infrastructure used for CATI surveys. In most cases, infrastructure for virtual phone banks is not dissimilar in type from that used in physical phone banks. However, variance in interviewer surroundings can cause variations in call quality for reasons related to connectivity, background noise, and productivity. This can be partially ameliorated through equipment (as well as through smart questionnaire design and training).

Equipment

Depending on data collection and IRB approvals, there are two main strategies to conduct virtual phone banks: (1) both calling and entering data on one device, or (2) calling on one device and entering data on a separate device. A data entry device is typically an Android phone or tablet, while a calling device might be any kind of cellular phone, a WiFi-enabled device capable of making voice calls over the internet. This results in three typical approaches for data entry.

- **Provide one cellular device (typically an Android tablet or phone):** Collect data using a cellular Android phone or tablet. If interviewers have very good internet connectivity for Voice over IP dialing, you could replace a cellular device with a WiFi-enabled device, but this is unlikely in many of IPA's survey locations.
- **Provide one cellular device and one non-cellular device:** Collect data using a device that cannot make cellular calls. Also provide interviewers with any kind of cellular phone to make calls.
- **Provide one non-cellular device and also rely on an interviewer's personal cellular device:** In some conditions, it may not be possible to provide a cellular phone or tablet for making calls, such as in cases where previous device procurement assumed face-to-face interviews and non-cellular devices. In this case, continue to use non-cellular tablets to read questions and enter data, while interviewers call on their personal phone using a call masking service.

It is NOT recommended that projects rely exclusively on interviewer personal devices for both calls and data entry. In this case, you should consult with your IRB about very detailed data security protocols and respondent confidentiality.

Calling on a phone-enabled Android tablet is the preferred method, as it (1) allows for maximum data security, (2) allows for recording both sides of the conversation with the smallest amount of data management, and (3) offers a large screen size for data entry.

In addition, projects can provide the following assets to interviewers to increase call fidelity and productivity:

- Headset
- Solar charger
- Mobile hotspot
- Protective case and screen cover

Software

Data Collection

IPA recommends using [SurveyCTO](#) to collect data during phone surveys, and has long relied on its extensive features, security, and customizability. SurveyCTO has released a number of updates and materials to allow for increased data security and ease of use during phone surveys. In particular, these features have been used to create a data collection platform specialized for CATI:

- [CATI Starter Kit](#): SurveyCTO has released guidance on two workflows for conducting phone surveys as part of their CATI starter kit. The advanced kit allows for multiple phone numbers that can be dynamically updated, stores previous attempt information, and includes the option to set a time for a rescheduled call.
- [Field Plug-ins](#): SurveyCTO's growing catalogue of field plug-ins allows for dialing, sending SMS, and even conducting training through the Collect app.
- [SurveyCTO Collect app](#): SurveyCTO's Collect app offers a number of additional data security and call monitoring features including password requirements as well as support for two-way call recording in Android versions 4.4-7.
- [SurveyCTO templates](#): IPA has created a number of templates for more specialized tasks in SurveyCTO. These include a template to include all attempts in the same submission for low-internet contexts, a template for mobilizers to schedule languages, and random assignment of cases to interviewer conditional on language with Stata or SurveyCTO.

Survey Implementation

Additional paradata and other characteristics such as call masking can be used securely through the use of other third-party software companies:

- **Exotel:** Exotel and similar companies such as Viamo or Twilio can mask the number that phones dial and receive to allow for increased respondent privacy. They can also record calls and deliver batch texts. J-PAL South Asia has created [a SurveyCTO plug-in and workflow](#) to integrate with Exotel within the SurveyCTO Collect app.
- **Third-party airtime transfer:** To ensure airtime are transferred, a number of services exist to batch transfer incentives to many numbers at once. These services are generally regional or country-specific.

Protocols

Physical Set-up

The COVID-19 context makes physical phone banks impossible due to risks of virus transmission between interviewers. If any physical interaction is required, ensure that social distancing is followed, individuals are separated by at least two meters, and appropriate hygiene practices are followed (such as washing hands with soap and water and frequently sanitizing surfaces).

For virtual phone banks, interviewers' home environments may have different conditions that may affect phone surveying such as levels of background noise, reliability of power, and internet connectivity. When you provide interviewers with tablets for data collection, consider including solar chargers and confirm that interviewers have regular access to the internet. If interviewers cannot predict their access to the internet, ensure that your case assignment protocol allows for irregular internet access.

If you cannot provide interviewers with a tablet and/or phone to record data due to quarantine or lockdown conditions, there are some software based solutions to preserve respondent confidentiality. If you are using SurveyCTO, a combination of the Collect app and third-party software can provide sufficient confidentiality. Use of Mobile Device Management (MDM) and Mobile App Management (MAM) technologies may allow tighter control of data collection applications installed on a personal device, but these options should be discussed and approved before implementing with both your IRB(s) and IT team. For any case where personal devices are used to collect survey data, your IRB application or amendment should detail additional procedures to ensure strong password protection and wipe protected data from their devices at the end of data collection (such as visiting the surveying organization's office after lockdown). It is also important, whether using provided or personal devices, to make sure every device has *its own SurveyCTO login*, rather than using one login across all devices. The SurveyCTO login cannot be a login that would allow an interviewer to access the server.

Call Scheduling

It may take many contact attempts to complete the interview for each respondent. Develop a protocols that standardized the following characteristics for survey call scheduling:

Table 4: Example Call Protocol Components

Category	Example
<i>Call frequency</i>	
<i>Minimum spacing between calls</i>	Call every other day
<i>Call hours</i>	Between 8 am and 8 pm, Mon. - Sat.
<i>Variation in call times for subsequent calls</i>	Call in the morning if the initial call was made in the afternoon
<i>Maximum number of calls</i>	5 calls
<i>Minimum coverage</i>	Called at least once in the weekend and once in the evening over a 10 day period
<i>Call assignment to interviewers</i>	Random order of respondent dialing
<i>Pre-survey notification</i>	Pre-survey SMS one day before the call
<i>Dialing protocol</i>	Let ring for at least one minute
<i>When to switch dialed numbers</i>	Only contact a call once if the number is deactivated; if the number does not connect, dial it four times over two weeks: (1) Two days later (2) Over the weekend (3) In the following week (4) In an evening in the following week
<i>Language reassignment</i>	Automated reassignment to a different interviewer
<i>Respondent identification</i>	Unmatched respondent numbers are removed from the pool
<i>Rescheduling</i>	Explicitly offer call rescheduling to respondents
<i>Dropped call attempts</i>	Try for three attempts before coding as a partial survey: (1) After the call drops; (2) Try again in 30 minutes; (3) attempt in the following day
<i>Respondent call backs</i>	Tell the respondent that an interviewer will call them shortly, and inform your supervisor to confirm case assignment
<i>Respondent airtime delivery</i>	Delivered by a third-party vendor at the end of each working day
<i>Case escalation</i>	Team leaders respond to interviewer's escalating cases and escalate the case to the field manager if necessary

Depending on the sample, it may be useful to increase the number of calls and time of day. Both Ozler & Ceivas and Suri recommend calling households at varying times of day across multiple days and weeks (2019; 2020). Varying call times and call days will increase chances of reaching populations that have regular periods without access to phone service—for example, agricultural workers who may be planting or harvesting during the day.

Additional protocol components may be relevant based on specific survey context. This is not an exhaustive list of examples, but is recommended for all projects.

Coding Call Statuses

Call statuses should reflect both scenarios that can result from a call for functional as well as response rate reporting guidelines. Ensure that many different scenarios are captured by your call status code -- especially if survey workflow components, such as call assignment and reassignment, have been automated. These statuses should be able to be used directly on import into data monitoring processes.

A good starting place for standardized disposition codes/call statuses is the WAPOR (World Association for Public Opinion Research) [standard definitions](#). SurveyCTO's [CATI starter kit](#) also provides ready-made forms with relevant choice lists for anyone conducting phone surveys.

We recommend including a subset of these statuses that correspond to AAPOR disposition codes. These codes should each align to a functional protocol or process or be necessary for reporting response rate. We report a combined list of codes used across RDD and existing samples, as well as projects that have complex reassignment protocols for their specific context. These codes may not all be relevant for your project:

1. **Not attempted:** The case is not attempted by any interviewer at the end of surveying.
2. **Invalid number:** The network does not recognize the number or the number belongs to a fax or data number which cannot receive calls.
3. **Inactive number:** The network recognizes the number, but the number is disconnected or temporarily out of service. For example, the SIM exists but is not currently in a phone, or the phone is off.
4. **Rings, no answer:** The attempt rings until the phone or network ends the call. This does not reach voicemail.
5. **Voicemail:** The attempt reaches a voicemail. Include information on if a voicemail is left if that is part of protocol.
6. **Busy tone:** The phone is busy during the duration of the attempts.

7. **Incorrect respondent:** The number does not correspond to the correct respondent (based on a pre-existing list). This will not be applicable for RDD lists.
8. **Rescheduled – busy:** The case is rescheduled because the respondent is busy.
9. **Rescheduled – language:** The case is rescheduled because the interviewer cannot speak the respondent’s language fluently enough to deliver the interview.
10. **Ineligible participant:** The participant is not eligible for the survey.
 - If ineligibility can occur for multiple reasons, include the specific reason why. Common ineligibility characteristics include business number, death, not part of household, unsupported language, cannot be contacted, quota filled, or incorrect age.
11. **Call drop:** The respondent picks up the phone and answers any question, but the call drops or is terminated before the consent script
12. **Inadequate audio quality:** The call is ended due to inadequate audio quality.
13. **Refused:** The respondent verbally refuses to participate in the survey before the first module begins.
14. **Partial:** The respondent consents to the interview, but the interview cannot be completed
15. **Complete:** The survey is successfully completed

A smaller subset of status codes can be used in simplified implementations:

2. **Invalid number**
4. **No answer**
8. **Incorrect respondent**
9. **Rescheduled - busy**
10. **Rescheduled - language**
11. **Ineligible participant**
14. **Refused**
15. **Partial**
16. **Complete**

In most cases, more detailed information on break-off and other response dispositions can be collected from skip patterns in the survey data and added for response rate reporting.

Collecting Contact Information

Collecting contact information for use in panel surveys is an important component of many questionnaires. If possible, collect at least one primary and one backup phone number, and potentially more than just two numbers. Numbers of friends and extended family can be used if the whole household moves. This allows an interviewer to contact someone who can potentially provide updated contact information or a location for a migrant household.

It’s important to ensure that contact data is collected with the highest fidelity. Interviewers should ask for the number twice and read it back if collecting information aloud.

[SurveyCTO's advanced CATI toolkit](#) allows for dynamic updating of numbers from preloaded data. If you are not using SurveyCTO, ensure that interviewers can update contact data, and that interviewers cannot cut and paste the same number into different fields. For example, put phone number entry fields on separate pages.

In terms of programming these data, ensure that the programming is done in such a way that in form validation can reduce chances for mistakes:

- Use field validation to ensure you have the right number of digits. Country codes and formats can be added as separate fields and combined when dialing.
 - Standardize collection of the “+” and validate fields on entry.
 - Train interviewers on a standard way of entering phone numbers.
- Be aware that the country codes may vary considerably even in within-country samples.
 - Example: a survey of Syrian refugees in Lebanon found respondent phone numbers from 60 countries, notably Latvia, where low-cost black market numbers originated. It was noted that maintaining mobile phone service in Lebanon (with a Lebanese number) is very expensive.
 - Variation in country-of-origin for telecom services could complicate your plans for transferring minutes, data, or funds via mobile transfer for respondent payments or reimbursement of air time. Ensure that you can deliver survey compensation to your pilot sample.
- Contact type may vary by number—WhatsApp, phone, or SMS—be sure that collection can record what type of contact is used for each number.

Incentives

Standardizing incentive protocols is also necessary. This includes both incentive format and delivery methods. There is some consensus that monetary incentives, the most widely tested, increase response rates by reducing refusal rate, but do so with diminishing returns as the size of incentives increases (Singer & Ye, 2013). These findings apply to the developing world, based on a number of experiments. Survey incentives, ranging in size from \$0.10 to \$3.00, increase response rate by around 5 percentage points, although this effect ranges from 2 to 10 percentage points depending on context and incentive size (Gibson et al., 2019; Lau et al., 2018b; Leo & Morello, 2016; Velthausz et al., 2016; Ballivan, Azevedo & Durbin, 2015). Irrespective of response rate increases, it is important to compensate respondents both for their opportunity cost of time as well as the cost in airtime and electricity that a phone conversation entails. Incentive size and format will vary with study context.

Common methods to deliver incentives over the phone include individual payments from the interviewer, centralized payments from the survey firm, and using a third-party delivery to deliver. It is often easiest to deliver payment through airtime, as it guarantees that

respondents can receive and use the incentive, as well as directly compensates for time spent on the phone. In cases where centralized remote delivery is impossible, purchasing airtime cards and reading the codes aloud on the phone is an option. If you use this approach, include an auditing mechanism to ensure that respondents are compensated.

Interviewer Training

Interviewer Recruitment

During a pandemic, interviewers should be trained remotely. Ideally, these interviewers are selected from a pool of interviewers with previous field experience working on research studies where they have built skills in interviewing, monitoring, tracking, app-based data collection, research ethics, and potentially the survey questionnaire itself. Ensure that you have considered interviewer fluency in the training and interviewing language(s).

Training Preparation and Prerequisites

Training should not begin until the survey questionnaire has been programmed and piloted, so that the training reflects the best possible version of the questionnaire. Later changes to the questionnaire will require re-training sessions. The duration of the training can also depend on the length and complexity of the questionnaire, as well as the interviewers' familiarity level with the questionnaire.

Also before the training begins, ensure that interviewers have the necessary set-up and equipment to complete the training, which may include: laptops, tablets, phones, stable internet connections, stable electricity, mobile airtime, headsets, and access to a quiet room.

Training format and engagement

During a pandemic, interviewer training should be entirely remote. However, if you attempt to live-stream a standard, all-day, week-long, lecture-heavy, face-to-face training, you will quickly lose the attention of your interviewers. You may also have logistical difficulty with live-streaming in areas with poor internet connectivity and/or high prices for bandwidth. You might consider pre-recording videos, uploading them to YouTube, and using [IPA's YouTube plug-in for SurveyCTO](#), which has worked well for streaming videos within the SurveyCTO Collect app in low-bandwidth environments.

A "mixed" approach to remote training is recommended. Here is an example agenda:

- Early morning: Group check-ins to set up the day's agenda.
- Late morning: Individual asynchronous review of training materials that incorporate comprehension quizzes.
- Lunch time: Re-group to set afternoon agenda, take lunch break.
- Early afternoon: One-on-one break-outs to discuss what was learned and/or practice skills and interviewing.

- Late afternoon: Group close-out sessions that incorporate overviews of the day's materials, lessons learned, and breakout discussions.

Consider hiring a higher number of trainers than normal for a face-to-face training, so that the ratio of trainers to trainees allows for breakouts, side chats, and troubleshooting.

Training content

While the core of the remote training should focus on the questionnaire that is unique to the context and research project, several training components should appear in most remote trainings. They include:

- Introduction to organization and research project
- Critical information on logistics, payroll, HR, management
- Ethics of human subjects research + informed consent
- Phone surveying protocols and workflows
- Phone conversation best practices and roleplaying
- How to use device(s) securely
- How to use data collection applications and tools
- Call assignment and tracking procedures and processes
- Data quality control procedures and processes
- Mental health awareness, such as how to respond to the effects of vicarious trauma
- Symptom checks and referral protocols as approved by appropriate IRBs and health institutions

It is important to give special consideration to interviewer mental health during a pandemic, and this starts during training. Not only will interviewers be facing a new work environment during a pandemic, they will also be interviewing respondents who are themselves potentially facing crises and traumas. Interviewers may experience “vicarious trauma” through their repeated exposure to the trauma of respondents, and may in turn experience burnout or other mental health challenges. It is important to create space to discuss these challenges during your training program, and also to build long-term opportunities for interviewers to focus on their mental health -- such as through a buddy program, breathing exercises, individual debriefs, and sick leave policies.

Training quizzes

Since it is more difficult to monitor interviewer attention to training and there are more distractions during training, it is important to include quizzes and other comprehension checks throughout training to ensure interviewers understand:

- **Phone surveying protocol:** How to handle all call outcomes, what to do when someone wants to speak to a supervisor, and how to mark a final call status.

- **Consent and respondent privacy:** How to read consent clearly and how to properly collect consent. Interviewers should also understand why it is important to ensure they are in a quiet place and to respect phone numbers as PII that should not be shared or used outside of the survey.
- **Survey-specific protocol:** How to categorize difficult questions, how to probe effectively without making the respondent uncomfortable, and how to explain questions.

Quizzes should be administered ideally at the end of each training day to ensure that interviewers understand all concepts covered. This will also help you recognize areas that you should cover again if many interviewers do not get the correct answer.

Ideally, you should also have interviewers go through the survey entirely during training so you can see how they categorize responses and what the data looks like for each interviewer.

Tracking & Case Management

Respondent Contact Information

Whether the study is collecting contact information for the first time or relying on existing contact information for respondents, here are some pieces of information that may facilitate calls to, and identification of, the right respondent:

- Full name
- Nickname
- Primary phone number
- Alternative phone numbers
- Preferred day(s) and time(s) for contact
- Address
- Names, nicknames, and phone numbers of household members (if permitted by IRB and if helpful for identifying the respondent)
- Unique ID (remember that a respondent may change phone numbers regularly, so even a primary phone number may not be a reliable unique ID)

Alternatively, the study may rely on Random Digit Dialing (RDD) to generate and dial random numbers without knowing any information about the respondent beforehand. In this case, ensure there is a robust system or platform in place for assigning and logging calls.

Daily Workflows and Call Assignment

Every interviewer should be trained on a standard work day and calling protocols, while also understanding that unexpected call results or challenging calls can be escalated to a manager. While every project context varies, see some considerations and recommendations for daily workflows and call assignments during remote surveying in the [Protocol](#) section of this document. Ensure that the study team receives feedback on these protocols during piloting and adds to them as necessary.

Survey software or platforms can help to automate this process, such as through SurveyCTO's very well-documented [CATI starter kit](#). The kit takes advantage of two very useful SurveyCTO features: "case management" to automate and organize the call attempt process, and a dialer "plug-in" to automate phone calls to specified numbers. IPA has also produced [two modifications of SurveyCTO's starter kits](#) to facilitate phone surveys in contexts with many languages and contexts with intermittent internet connectivity. One template reduces the number of form submissions by documenting multiple contact attempts in one submission; the other template kicks off respondent engagement through

a survey “mobilizer” who determines a respondent’s preferred language and appointment time for the full interview.

Protocols for Special Cases

Not every phone call will result in a respondent answering the phone, consenting to the survey, and completing the survey. You should plan ahead for special cases, including:

- Respondents provide “soft refusals” to participate.
- Call is dropped in the middle of the survey.
- Someone answers the call, but it is the wrong person.
- Interviewer and respondent do not speak the same language as expected.
- Interviewers must end their employment and calls must be re-assigned.
- Interviewer requests escalation of the case to the manager for any other reason.

Dropped Calls/Incomplete Surveys

Incomplete surveys are inevitable with CATI surveys. Interviewers should be trained on how to categorize dropped calls and incomplete surveys, and your questionnaire should incorporate this situation into its programming.

Interviewers should still be able to submit questionnaires if they are not able to complete them due to a dropped call or the respondent’s need to reschedule the rest of the interview. While interviewers can save the questionnaire and complete it later, consider including checkpoints before each module of the questionnaire that ask the interviewer if the respondent is still on the phone. If the answer is no, the interviewer should be taken to the end of the questionnaire and still be allowed to submit the questionnaire.

You should also create a data flow for a situation where an interviewer has completed a new questionnaire after a previous attempt was incomplete and submitted. Interviewers can go through the entire questionnaire again with the respondent, or you can consider programming your questionnaire so they only have to complete the unanswered questions. Either way, you should include a workflow that combines these two attempts into one observation.

Reporting

Ensure that every member of the research team can see the data they need to make informed, timely decisions. Platforms such as SurveyCTO can integrate with reporting tools like Google Sheets to create simple [dashboards](#) that provide timely or even real-time information to team members. These dashboards can be used to supplement standard high-frequency checks run in Stata. For example, a Field Manager should be able to quickly



see which cases are in need of review, while a Principal Investigator should be able to get an overall picture of survey progress.

Your survey protocols should be reflected in the information you report. Information should be presented to the relevant team member in a usable format, and highlight action items. For example, if you are concerned about productivity, ensure that your reporting looks at success rates by call number, potentially disaggregated by time of day or initial call result. Similar to face-to-face surveying, ensure that you regularly schedule calls to digest this information and modify protocols that are not effective. It's not worth setting up reporting if this information is not used.

Monitoring

Remote data collection may make it impossible to run standard implementation quality protocols including spotchecks -- where a field supervisor randomly audits an on-going interview without informing the field officer -- and accompaniments -- where a field supervisor accompanies a subset of field officers' interviews to monitor field officer performance and to check for survey issues. However, phone surveys allow for data quality checks that are not possible in face-to-face surveys. In particular consider the following:

- Increase the use of high-frequency checks on metadata, especially to flag responses for audio auditing.
- Backcheck time invariant data or collect audit data at higher frequencies.
- Use audio audits to substitute for spotchecks and accompaniments, or consider using three-way calls with supervisors as spotchecks.
- Collect "text audit" data and analyze the amount of time interviewers spend on each question; consider implementing SurveyCTO "speed violations" which trigger audio recordings when questions are completed too quickly.

High Frequency Checks

Standard high-frequency checks (HFCs) should be run daily or at another regular frequency. IPA recommends using [IPA's data management system](#) (DMS) to run these checks in Stata. Although the standard set of DMS checks covers many data quality areas, some additional checks can be included for CATI surveys. These include call status codes, time of day, and day of week trends. Otherwise, consider more in-depth reviews of partial responses and attention checks.

To easily track implementation information outside of HFCs, also consider setting up a dashboard on Google Sheets with information on response rates and attempts. SurveyCTO can [publish form data directly to Google Sheets](#). This data can then be used to generate a quick snapshot of the data through the use of a pre-coded summary statistics. Setting up this instrument is advantageous for project management in a remote context, especially around survey progress and protocol management. This does not replace the granularity that HFCs run through Stata provide.

When collecting data using SurveyCTO, the large size of media files that contain audio may result in long download times. To increase download speed, consider downloading data for high-frequency checks using an [API in Stata](#), or setting up a user in SurveyCTO desktop to [download without attached media files](#).

Comparisons with Baseline

If baseline or ground-truth data on demographics exist, consider using these data to confirm questions that should be immutable as part of standard high-frequency checks. Similar to flagging backchecks with errors on [Type 1 questions](#), comparisons with existing data can allow you to flag potential erroneous entry or other data quality problems such as a misidentified respondent.

Backchecks

Backchecks are shortened questionnaires where a respondent is interviewed by an auditor to check reliability as well as survey implementation quality. We suggest increasing the percentage of surveys backchecked. Standard [IPA protocol](#) suggests backchecking at least 10% of all surveys to ensure reliability of survey questions and measure interviewer performance. Tavneet Suri recommends backchecking at much higher rates, up to 100% (2020). Since all data collection is happening remotely, it is important to confirm that respondents have been called and that numbers marked as not working have been checked. Include higher rates of [Type 1 questions](#) to ensure interviewers are calling respondents and information is correct, as well as backcheck soft refusals or other numbers.

Concerns about panel fatigue are a valid concern. In the case that panel attrition is being managed carefully, some studies have substituted backcheck surveys with audio audits of surveys and used double entry to compare this recording. Although this provides measurement of entry error and potentially flagging other data quality or enumeration concerns, this protocol makes it impossible to measure test-retest reliability. If this protocol is used, also ensure that data quality measures are tested in detail.

Sensor Data and Audio Audits

Audio recordings of the interviews, audio audits, can be used to allow for lighter touch review of data. Ensure that the devices you are using can collect audio audits. Audio recording is enabled by the SurveyCTO Collect app and other third party applications like Exotel or Twilio. If devices cannot record calls natively, it is also possible to use speakerphone on their calls to record both sides of the call, although that can disrupt survey enumeration and affect response quality. If so, you should also ensure that interviewers are conducting interviews in a location where they can use speakerphone if audio audits cannot capture the respondent's voice.

Audio recording must also be in your IRB application and in your informed consent to record calls. If audio audits are included in an IRB protocol, ensure that "call may be

recorded for quality control” is part of the consent script, or that audio recording is a separate consent that respondents can opt out of. The latter approach, a separate audio consent, is recommended in most cases.

Responses to some items may be affected by audio audits. This is especially relevant around items that the respondent may feel more comfortable responding to in private contexts such as items referencing topics like intimate partner violence. Consider in-depth piloting and/or gathering feedback from interviewers and respondents to confirm that respondents feel comfortable answering items around these topics.

If interviewers cannot use a call recording app or two-sided audio recording in SurveyCTO, one-sided audio audits, that only record the interviewer’s side of the audio, can be used for data quality checks. This reduces the effectiveness of audio audits considerably.

To review audio audits, auditors can collect and analyze multiple forms of data. Auditors can analyze interviewers’ performance equivalent to a spotcheck or accompaniment as well as using double entry on the same form to check for entry error. These data can be used as part of backchecks or to inform retraining on questions that are not performing as expected from pilot data.

SurveyCTO can also collect [sensor data](#) on light, movement, and sound measured through the sensors in the Android device. Sensor data provides information on the physical context under which the survey took place. Light and movement data may not be useful for CATI surveys, but statistics on sound levels and predicted conversation can be used to target supplemental audio audits or to understand the environment in which interviewers are making calls. These fields can be included in HFCs. We recommend including four sensor statistics: mean sound level, standard deviation of sound level, percent quiet, and percent conversation.

References

1. Amaya, Ashley, Charles Lau, Yaa Owusu-Amoah, and Jocelyn Light. "Evaluation of gaining cooperation methods for IVR surveys in low-and middle-income countries." *Survey Methods: Insights from the Field* (2018): 1-8.
2. Arthi, Vellore, Kathleen Beegle, Joachim De Weerd, and Amparo Palacios-López. *Not your average job: measuring farm labor in Tanzania*. The World Bank, (2016).
3. Ballivian, A, João Pedro Azevedo and Will Durbin. "Using Mobile Phones for High- Frequency Data Collection." In *Mobile Research Methods: Opportunities and Challenges of Mobile Research Methodologies*, 21–39. London: Ubiquity Press, 2015.
4. Bloomberg Philanthropies Data for Health Initiative. "Executive Summary: Philippines NCD Mobile Phone Survey 2018." (2020).
5. Bloomberg Philanthropies Data for Health Initiative. "Executive Summary: Zambia NCD Mobile Phone Survey 2017." (2020).
https://0df04920-08a1-4b70-a12c-25629cb5ffc4.filesusr.com/ugd/7a26dc_8d7efc17f3d649fd8f3db4c8fa7c0bcc.pdf
6. Cantor, David, Barbara C. O'Hare, and Kathleen S. O'Connor. "The use of monetary incentives to reduce nonresponse in random digit dial telephone surveys." *Advances in telephone survey methodology* (2008): 471-498.
7. Curran, P. G. (2016). Methods for the detection of carelessly invalid responses in survey data. *Journal of Experimental Social Psychology*, 66, 4-19.
8. Dabalen, Andrew, Alvin Etang, Johannes Hoogeveen, Elvis Mushi, Youdi Schipper, and Johannes von Engelhardt. *Mobile phone panel surveys in developing countries: a practical guide for microdata collection*. The World Bank, (2016).
9. Daftary, Amrita, Yael Hirsch-Moverman, Getnet M. Kassie, Zenebe Melaku, Tsigereda Gadisa, Suzue Saito, and Andrea A. Howard. "A qualitative evaluation of the acceptability of an interactive voice response system to enhance adherence to isoniazid preventive therapy among people living with HIV in Ethiopia." *AIDS and Behavior* 21, no. 11 (2017): 3057-3067.
10. Dal Grande, Eleonora, Catherine Ruth Chittleborough, Stefano Campostrini, Maureen Dollard, and Anne Winifred Taylor. "Pre-survey text messages (SMS) improve participation rate in an Australian mobile telephone survey: An experimental study." *PloS one* 11, no. 2 (2016).
11. Dillon, Brian. "Using mobile phones to collect panel data in developing countries." *Journal of international development* 24, no. 4 (2012): 518-527.
12. Edwards, Michelle L., Don A. Dillman, and Jolene D. Smyth. "An experimental test of the effects of survey sponsorship on internet and mail survey response." *Public Opinion Quarterly* 78, no. 3 (2014): 734-750.
13. Edwards, Phil, Ian Roberts, Mike Clarke, Carolyn DiGuseppi, Sarah Pratap, Reinhard Wentz, and Irene Kwan. "Increasing response rates to postal questionnaires: systematic review." *Bmj* 324, no. 7347 (2002): 1183.
14. Feng, Shuo, Karen A. Grépin, and Rumi Chunara. "Tracking health seeking behavior during an Ebola outbreak via mobile phones and SMS." *NPJ digital medicine* 1, no. 1 (2018): 1-8.
15. Ferreira, Aline Dayrell, Cibele Comini César, Deborah Carvalho Malta, Amanda Cristina de Souza Andrade, Cynthia Graciane Carvalho Ramos, Fernando Augusto Proietti, Regina Tomie Ivata Bernal, and Waleska Teixeira Caiaffa. "Validity of data collected by telephone survey: a comparison of VIGITEL 2008 and 'Saude em Beaga' survey." *Revista Brasileira de Epidemiologia* 14 (2011): 16-30.
16. Francisco, Priscila Maria Stolses Bergamo, Marilisa Berti de Azevedo Barros, Neuber José Segri, Maria Cecília Goi Porto Alves, Chester Luiz Galvão Cesar, and Deborah Carvalho Malta.

- "Comparison of estimates for the self-reported chronic conditions among household survey and telephone survey-Campinas (SP), Brazil." *Revista Brasileira de Epidemiologia* 14 (2011): 5-15.
17. Gallup. 2011. Mobile Phone Access Varies Widely in Sub-Saharan Africa. <https://news.gallup.com/poll/149519/mobile-phone-access-varies-widely-sub-saharan-africa.aspx>
 18. Gallup. "The World Bank Listening to LAC (L2L) Pilot Final Report." (2012).
 19. Garlick, Robert, Kate Orkin, and Simon Quinn. "Call me maybe: Experimental evidence on frequency and medium effects in microenterprise surveys." *The World Bank Economic Review* 34, no. 2 (2020): 418-443.
 20. Gibson, Dustin G., Adaeze C. Wosu, George William Pariyo, Saifuddin Ahmed, Joseph Ali, Alain B. Labrique, Iqbal Ansary Khan, Elizeus Rutebemberwa, Meerjady Sabrina Flora, and Adnan A. Hyder. "Effect of airtime incentives on response and cooperation rates in non-communicable disease interactive voice response surveys: randomised controlled trials in Bangladesh and Uganda." *BMJ global health* 4, no. 5 (2019): e001604.
 21. Gibson, Dustin G., Amanda Pereira, Brooke A. Farrenkopf, Alain B. Labrique, George W. Pariyo, and Adnan A. Hyder. "Mobile phone surveys for collecting population-level estimates in low-and middle-income countries: a literature review." *Journal of medical Internet research* 19, no. 5 (2017): e139.
 22. Greenleaf, Abigail R., Dustin G. Gibson, Christelle Khattar, Alain B. Labrique, and George W. Pariyo. "Building the evidence base for remote data collection in low-and middle-income countries: comparing reliability and accuracy across survey modalities." *Journal of medical Internet research* 19, no. 5 (2017): e140.
 23. GSMA. 2020. Connected Women: The Mobile Gender Gap Report 2020.
 24. Heath, Rachel, Ghazala Mansuri, D. Sharma, B. Rijkers, and W. Seitz. "Measuring Employment: Experimental Evidence from Ghana." (2017).
 25. Heward-Mills, Nii Lante and Yaa Amankwaa Owusu-Amoah. "USAID Communicate for Health: Mobile Cohort Study - Follow-up Survey 2018." USAID Communicate for Health. (2018).
 26. James, Jeannine M., and Richard Bolstein. "Large monetary incentives and their effect on mail survey response rates." *Public Opinion Quarterly* 56, no. 4 (1992): 442-453.
 27. Janevic, Mary R., Amparo C. Aruquipa Yujra, Nicole Marinec, Juvenal Aguilar, James E. Aikens, Rosa Tarrazona, and John D. Piette. "Feasibility of an interactive voice response system for monitoring depressive symptoms in a lower-middle income Latin American country." *International journal of mental health systems* 10, no. 1 (2016): 59.
 28. Johnson, Douglas. "Collecting data from mHealth users via SMS surveys: A case study in Kenya." *Survey Practice* 9, no. 1 (2016): 2824.
 29. Kastelic, Kristen Himelein, and Jonathan G. Kastelic. "The socio-economic impacts of Ebola in Liberia: Results from a high frequency cell phone survey round five." *Washington, DC: The World Bank Group* (2015).
 30. Kastelic, Kristen Himelein, Mauro Testaverde, Abubakarr Turay, and Samuel Turay. "The socio-economic impacts of Ebola in Sierra Leone: Results from a high frequency cell phone survey round three." (2015).
 31. Kasy, Maximilian, and Anja Sautmann. "Adaptive treatment assignment in experiments for policy choice." (2019).
 32. Liu, M., & Wronski, L. (2018). Trap questions in online surveys: Results from three web survey experiments. *International Journal of Market Research*, 60(1), 32-49.
 33. L'Engle, Kelly Eunice Sefa, Edward Akolgo Adimazoya, Emmanuel Yartey, Rachel Lenzi, Cindy Tarpo, Nii Lante Heward-Mills, Katherine Lew, and Yvonne Ampeh. "Survey research with a random digit dial national mobile phone sample in Ghana: methods and sample quality." *PLoS one* 13, no. 1 (2018).
 34. Lamanna, Christine, Kusum Hachhethu, Sabrina Chesterman, Gaurav Singhal, Beatrice Mwongela, Mary Ng'endo, Silvia Passeri et al. "Strengths and limitations of computer assisted

- telephone interviews (CATI) for nutrition data collection in rural Kenya." *PloS one* 14, no. 1 (2019).
35. Lau, Charles Q., Eric Johnson, Ashley Amaya, Patricia LeBaron, and Herschel Sanders. "High stakes, low resources: what mode (s) should youth employment training programs use to track alumni? Evidence from South Africa." *Journal of International Development* 30, no. 7 (2018a): 1166-1185.
 36. Lau, Charles Q., Ansie Lombaard, Melissa Baker, Joe Eyerman, and Lisa Thalji. "How representative are SMS surveys in Africa? Experimental evidence from four countries." *International Journal of Public Opinion Research* 31, no. 2 (2018b): 309-330.
 37. Lau, Charles, and Nicolas di Tada. "Identifying Non-Working phone numbers for response rate calculations in Africa." *Surv Pract* 11 (2018c).
 38. Lau, Charles, Ehud Gachugu, Eric Johnson, and Leenisha Marks. "Using SMS Technology to Survey Low-Income Youth: Lessons from a Vocational Education Tracking Study in Kenya." *Journal of International Development* 30, no. 6 (2018d): 1060-1063.
 39. Lau, Charles Q., Alexandra Cronberg, Leenisha Marks, and Ashley Amaya. "In Search of the Optimal Mode for Mobile Phone Surveys in Developing Countries. A Comparison of IVR, SMS, and CATI in Nigeria." In *Survey Research Methods*, vol. 13, no. 3, pp. 305-318. (2019a).
 40. Lau, Charles Q., Herschel Sanders, and Ansie Lombaard. "Questionnaire Design in Short Message Service (SMS) Surveys." *Field Methods* 31, no. 3 (2019b): 214-229.
 41. Leo, Ben, and Robert Morello. "Practical Considerations with Using Mobile Phone Survey Incentives: Experiences in Ghana and Tanzania." *Center for Global Development Working Paper* 431 (2016).
 42. Leo, Benjamin, Robert Morello, Jonathan Mellon, Tiago Peixoto, and Stephen T. Davenport. "Do mobile phone surveys work in poor countries?." *Center for Global Development Working Paper* 398 (2015).
 43. Li, Tracey. UN Mobile Data Training Workshop: Mobile data analysis in low-income country settings.
<https://unstats.un.org/unsd/bigdata/conferences/2019/workshops/mobile-phone/day2/04%20MPD%20-%20Resource-scarce%20settings%20-%20Flowminder.pdf>
 44. Maffioli, Elisa M. "Relying Solely on Mobile Phone Technology: Sampling and Gathering Survey Data in Challenging Settings." (2019).
 45. Mahfoud, Ziyad, Lilian Ghandour, Blanche Ghandour, Ali H. Mokdad, and Abba M. Sibai. "Cell phone and face-to-face interview responses in population-based surveys: how do they compare?" *Field methods* 27, no. 1 (2015): 39-54.
 46. Medway, Rebecca. "Beyond response rates: The effect of prepaid incentives on measurement error." PhD diss., (2012).
 47. Meuleman, Bart, Arnim Langer, and Annelies G. Blom. "Can incentive effects in web surveys be generalized to non-western countries? Conditional and unconditional cash incentives in a web survey of Ghanaian university students." *Social Science Computer Review* 36, no. 2 (2018): 231-250.
 48. Morse, Ben, Karen A. Grépin, Robert A. Blair, and Lily Tsai. "Patterns of demand for non-Ebola health services during and after the Ebola outbreak: panel survey evidence from Monrovia, Liberia." *BMJ global health* 1, no. 1 (2016): e000007.
 49. Oppenheimer, D. M., Meyvis, T., & Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of experimental social psychology*, 45(4), 867-872.
 50. Özler, Berk, and P. Facundo Cuevas. "Reducing Attrition in Phone Surveys." World Bank Blogs, November 21, 2019.
<https://blogs.worldbank.org/impacitevaluations/reducing-attrition-phone-surveys>.
 51. Pariyo, George W., Abigail R. Greenleaf, Dustin G. Gibson, Joseph Ali, Hannah Selig, Alain B. Labrique, Gulam Muhammed Al Kibria et al. "Does mobile phone survey method matter? Reliability of computer-assisted telephone interviews and interactive voice response

- non-communicable diseases risk factor surveys in low and middle income countries." *PloS one* 14, no. 4 (2019).
52. Pattnaik, Anooj, Diwakar Mohan, Sam Chipokosa, Sautso Wachepa, Hans Katengeza, Amos Misomali, and Melissa A. Marx. "Testing the validity and feasibility of using a mobile phone-based method to assess the strength of implementation of family planning programs in Malawi." *BMC health services research* 20, no. 1 (2020): 1-9.
 53. Singer, Eleanor, and Cong Ye. "The use and effects of incentives in surveys." *The ANNALS of the American Academy of Political and Social Science* 645, no. 1 (2013): 112-141.
 54. Stecklov, Guy, Alexander Weinreb, and Calogero Carletto. "Can incentives improve survey data quality in developing countries?: results from a field experiment in India." *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 181, no. 4 (2018): 1033-1056.
 55. Suri, Tavneet. 2020. Phone Surveys: Use and Adaptation to COVID-19. Adaptations for phone surveys: A Webinar with Tavneet Suri. J-PAL.
<https://www.povertyactionlab.org/blog/3-20-20/best-practices-conducting-phone-surveys>
 56. Vannette, David L. 2016. "Testing the Effects of Different Types of Attention Interventions on Data Quality in Web Surveys. Experimental Evidence From a 14 Country Study." Paper presented at the 71st Annual Conference of the American Association for Public Opinion Research in Austin, TX.
 57. Velthausz, Daan, Rotafina Donco, Hannah Skelly and Marga Eichleay. "Mozambique Mobile Access and Usage Study: Computer-Assisted Telephone Interview (CATI) Survey Results." United States Agency for International Development. (2016).
 58. Vogel, Lara D., Levi Goertz, Suhuyini S. Shani, Mark Boots, Louis Dorval, and N. Ewen Wang. "A mobile-based healthcare utilization assessment in rural Ghana." *Procedia engineering* 159 (2016): 366-368.
 59. West, Brady T., Dirgha Ghimire, and William G. Axinn. "Evaluating a modular design approach to collecting survey data using text messages." In *Survey research methods*, vol. 9, no. 2, p. 111. NIH Public Access, (2015).
 60. World Bank. "Listening to Tajikistan - Household Survey: Background, Implementation, and Methods (English)." Washington, D.C. : World Bank Group. (2017).
<http://documents.worldbank.org/curated/en/624621538136672723/Listening-to-Tajikistan-Household-Survey-Background-Implementation-and-Methods>