

Survey Research in Low- and Middle-Income Countries (LMICs) During the COVID-19 Pandemic

AAPOR Webinar

April 23, 2021

Charles Lau, PhD

Senior Survey Methodologist

RTI International



Today's Webinar

Outline

1. **Introduction (5 minutes)**
2. **Survey Examples (5 minutes)**
3. **Mobile Phone Design (15 minutes)**
 1. Modes
 2. Sampling
 3. Data Collection
4. **Survey Error (20 minutes)**
 1. Coverage and Non-Response
 2. Measurement
5. **Ethics (5 minutes)**
6. **Discussion (30 minutes)**

Learning Objectives

1. Understand challenges with coverage, response, and measurement for mobile phone surveys in low- and middle-income countries
2. Learn emerging best practices to maximize the quality of mobile phone surveys in low- and middle-income countries
3. Learn about examples of successful mobile phone surveys in low- and middle-income countries on COVID-19 and other topics



1. Introduction

Evolution of Modes

WEB SURVEYS

A REVIEW OF ISSUES AND APPROACHES

MICK P. COUPER

As we enter the twenty-first century, the Internet is having a profound effect on the survey research industry, as it is on almost every area of human enterprise. The rapid development of surveys on the World Wide Web (WWW) is leading some to argue that soon Internet (and, in particular, Web) surveys will replace traditional methods of survey data collection. Others are urging caution or even voicing skepticism about the future role Web surveys will play. Clearly, we stand at the threshold of a new era for survey research, but how this will play out is not yet clear. Whatever one's views about the likely future for survey research, the current impact of the Web on survey data collection is worthy of serious research attention.

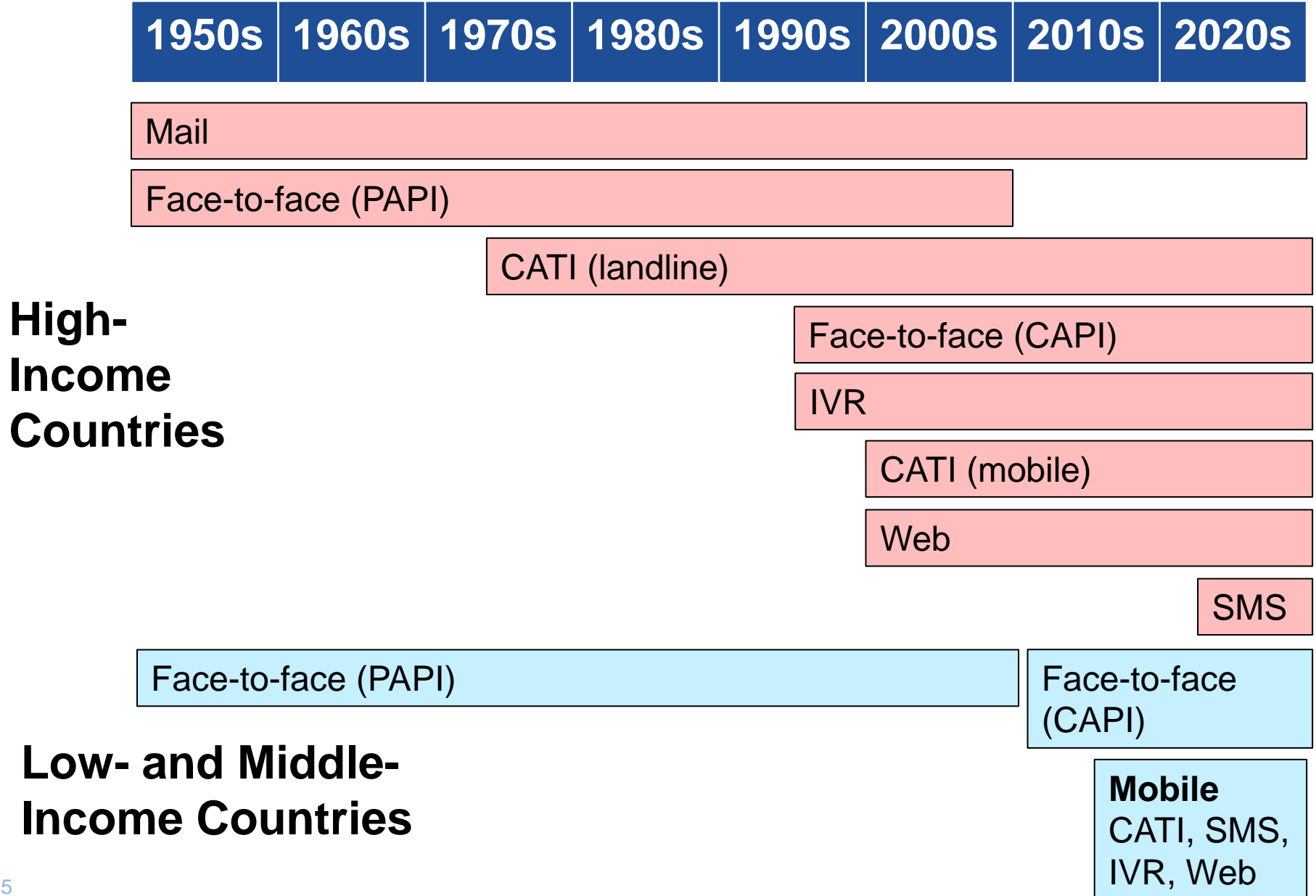
Given the rapidly growing interest in Web surveys,¹ it is important to distinguish among different types of Web surveys. The rubric "Web survey" encompasses a wide variety of methods, with different purposes, populations,

Web surveys in ~2000 in high-income &
Mobile phone surveys in ~2020 in LMICs

- Disruptive to industry
- Seen as threat and savior
- Low cost, fast
- Concerns about representativeness and measurement quality
- Present technical challenges
- Raise legal and ethical questions

Couper, M. P. (2000). "Review: Web Surveys: A Review of Issues and Approaches." Public Opinion Quarterly, 64(4): 464-494.

Evolution of Modes



Distinctive Features of Mobile Phone Surveys in LMICs



Influence of international development community



Fragmented literature



Survey methodologists are later to the party



Rapid innovation



Data collection beyond *surveys*

Scope for the Webinar

Mobile phone modes

- Computer-assisted telephone interviewing (CATI)
- Short message service (SMS)
- Interactive voice response (IVR)
- Mobile web

Not covered: Face-to-face, landlines, chatbots

National surveys

Less prominent: Beneficiary surveys, monitoring and evaluation

Selected low- and middle-income countries

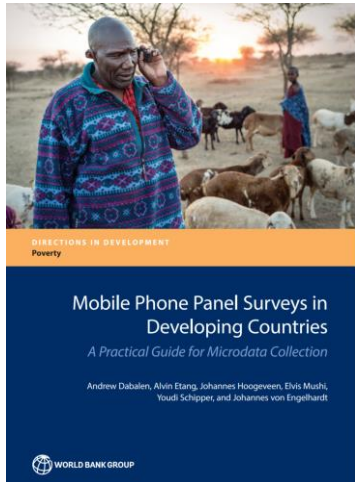
A taste of relevant work!



2. Survey Examples

Examples of Mobile Phone Surveys: Pre-COVID

World Bank [High-Frequency Mobile Phone Surveys](#)

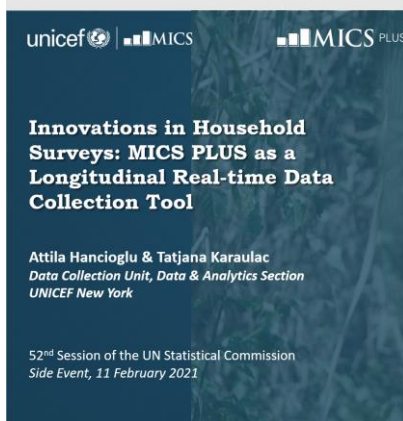


CDC/RTI/InSTEDD [Non-communicable disease surveillance](#)



GeoPoll [Surveys: Conflict Zones](#), Post-Disaster

UNICEF [MICS Plus](#)



Examples of Mobile Phone Surveys: Focus on COVID-19

WAPOR [list of Mobile Phone Surveys on COVID-19](#)

The screenshot shows the WAPOR (Worldwide Access to Public Opinion Research) website. The header includes navigation links: ADULT WAPOR, MEMBERSHIP, EVENTS, PUBLICATIONS, RESOURCE LIBRARY, WAPOR NEWS, and a search bar. A 'MEMBER LOGIN' button and a 'DONATE' link are also present. Below the header, a dark banner reads 'COVID-19 Public Opinion Research'. The main content area features a 'Resource Library' sidebar on the left with a list of WAPOR Webinars from September 2019 to March 2021. The main text area states: 'Public opinion research on COVID-19 is being conducted all over the world. We have compiled a list of public opinion reports from around the world that have been posted on our listserve, as well as those that have been submitted to us by other research institutions. Please feel free to submit any links to research studies or documents to us at waporoffice@gmail.com. This page was last updated 17 December 2020 at 2:24 PM EDT.' Below this is a table with two columns: 'Country' and 'Links to Research'.

Country	Links to Research
Argentina	Link to research (Voices) Webinar + PDF (Voices Consultancy)
Australia	Link to research (Australian National University and the Social Research Centre)
Austria	Link to research (COVID-19)

World Bank High frequency CATI surveys ([dashboard](#), [data](#), [Nature paper](#))

The screenshot shows the 'COVID-19 HIGH FREQUENCY MONITORING DASHBOARD [BETA]' on the World Bank website. The dashboard has a navigation bar with tabs: MAIN, OVERVIEW, CROSSTAB, TRENDS, and SCATTERPLOT. The main content area features a large blue header with the title 'COVID-19 HIGH FREQUENCY MONITORING DASHBOARD [BETA]'. Below the header, a text box states: 'Learn more about the socioeconomic impacts of COVID-19 on households and individuals in 50 countries across all developing regions.' Another text box explains: 'This dashboard provides 96 harmonized indicators on 14 topics, allowing users to compare and analyze how COVID-19 impacts vary across countries, over time and by industry sector and regions. The data can be downloaded for further analysis.' A link to 'Technical Note' is provided. At the bottom, a text box says: 'To learn more about indicator construction and data availability by country, please refer to our [Data Dictionary](#). To see which countries are included in the latest data, refer to our [Country List](#).'

Innovations for Poverty Action [RECOVER](#)

The screenshot shows the IPA (Innovations for Poverty Action) website. The header includes navigation links: FOR RESEARCHERS, BLOG, NEWS, EVENTS, CONTACT, and a search bar. Below the header, a dark banner reads 'Research for Effective COVID-19 Responses (RECOVER)'. The main content area features a large image of a person wearing a green t-shirt with the IPA logo. Below the image, a blue bar contains the text: 'Overview RECOVER Research Hub RECOVER Survey Phone Survey Methods'.

In response to the COVID-19 pandemic, IPA has launched Research for Effective COVID-19 Responses, or RECOVER. Through this effort, IPA is generating rigorous evidence, advising governments, and bringing together partners across the research-to-policy sector to rapidly provide decision-makers with rigorous data and evidence to mitigate the impacts of the crisis.



3. Mobile Phone Survey Design

3.1 Modes

	Self-Administered	Interviewer-Administered
Voice	Interactive voice response (IVR)	Computer-assisted telephone interviewing (CATI)
Text	Short message service (SMS) Mobile web	<i>(Rare)</i>

Key concepts:

- Inbound/Outbound
- All modes typically free for respondents
- Language selection

3.1 Modes: CATI

Note: Many don't use "CATI" terminology; "voice surveys" or "phone surveys" is more common

How it Works

- Interviewer-administered phone surveys
- Permanent and temporary call centers
- Similar QC procedures as in high-income countries
- Questionnaire length: Optimal length is unclear, guidance ranges from 15-30 minutes.
- Virtual and in-person call centers



Advantages

- ✓ Doesn't require literacy
- ✓ Makes certain question types possible –e.g., collecting village can be challenging in self-administered modes
- ✓ Interviewers can tailor survey request and set appointments
- ✓ Interviewers can probe and clarify
- ✓ Longer questionnaires compared to other modes

Disadvantages

- ✗ Can be more expensive than IVR/SMS/web. BUT this depends on many factors: number of questions, number of interviews, whether professional or temporary call center used
- ✗ Interviewer effects (??? – more research needed)
- ✗ Worse for sensitive topics (??? – more research needed)
- ✗ Synchronous mode: requires respondent to pick up
- ✗ No visual content

3.1 Modes: IVR

How it Works

- Respondents listen to recorded questions and input a numeric response associated with their answer
- IVR recognition of a respondent's verbal response is rare
- ~20 questions (but limited research on optimal length), corresponds to about 10 minutes
- Surveys typically allow respondent to skip by pressing hash (#) or similar key
- Questions are often looped, to provide opportunities to listen again
- Typically, IVR systems hang up after ~3-5 error messages on same question or question repeats
- Language selector at beginning of survey for multi-lingual surveys
- Oftentimes female voice, but research is limited on gender-of-voice effects
- Some surveys have "call back" feature
- Close-ended questions only: Open-ended capture hasn't been successful (e.g., Morello and Leo, 2016)

Advantages

- ✓ Doesn't require literacy
- ✓ Eliminates interviewer effects in CATI (but standardization can have flaws)
- ✓ Potentially better for sensitive topics (???)

Disadvantages

- ✗ Unfamiliar mode to respondents
- ✗ Work required to record and manage audio
- ✗ Synchronous mode: requires respondent to pick up
- ✗ No visual content
- ✗ Suitable only for short surveys, concise questions
- ✗ Unclear who is responding

3.1 Modes: SMS

How it Works

- Short message service = SMS, or text messages
- Respondents receive text messages, one per survey question
- Texting available on virtually all phones
- This is self-administered **automated** mode: there aren't interviewers on other end (except in rare cases – e.g., [West, Ghimire, Axinn, 2015](#))
- Each question has 160-character limit for question and response options (including spaces)
- Multiple messages possible on some tools, but these can lead to errors
- Character limit depends on country and mobile network operator
- Sometimes special characters (é) can count double towards character limit (check the survey tool)
- ~20 questions (but limited research on optimal length)
- Language selector at beginning of survey for multi-lingual surveys

Advantages

- ✓ Asynchronous: people can respond at leisure
- ✓ Natural form of communication
- ✓ Less expensive than IVR
- ✓ Suitable for “momentary assessments” (real-time data)
- ✓ Potentially better for sensitive topics

Disadvantages

- ✗ Unfamiliar mode to respondents
- ✗ Requires literacy
- ✗ Questions must be short
- ✗ Problems with splitting up messages
- ✗ No visual content
- ✗ Small screen size often requires scrolling
- ✗ Suitable only for short surveys, concise questions
- ✗ Responses remain on phones, which can be a problem for surveys on sensitive topics (Furchow and Mac Ginty, 2020)
- ✗ Unclear who is responding

3.1 Modes: Web

How it Works

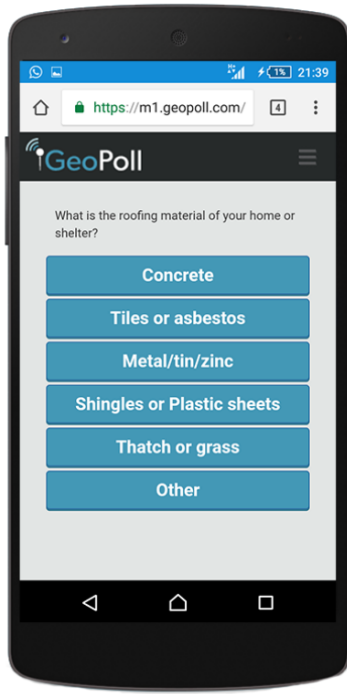
- Respondents receive an SMS invitation, click on link, and are taken to a web survey on their mobile device
- Respondents can also be emailed or mailed, but this is rarer in LMICs
- In LMICs, more people access Internet on mobile device rather than laptop

Advantages

- ✓ Asynchronous: people can respond at leisure
- ✓ Visual content
- ✓ Less expensive than IVR
- ✓ Better user experience for respondent
- ✓ Provides more credibility for sponsor (e.g., logo, branding)

Disadvantages

- ✗ Hard to provide free access to respondents
- ✗ Requires internet access
- ✗ Requires smartphone (leads to undercoverage)
- ✗ Requires literacy
- ✗ Questions must be short
- ✗ Small screen size often requires scrolling



Source:
[GeoPoll](https://www.geopoll.com/)

3.1 Modes - Comparisons

Source: Innovations for Poverty Action and Northwestern University Global Poverty Research Lab (2020). Remote Surveying in a Pandemic: Handbook. Available [here](#).

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.

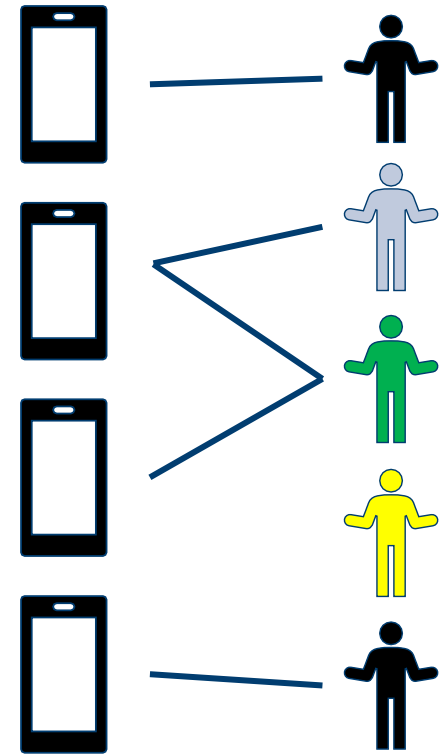
3.2 Sampling - Frames for Mobile Phones

For more information, see [Himelein et al. \(2020\)](#)

Advantage		Disadvantage	
Random digit dial	<ul style="list-style-type: none"> ✓ Large sample size ✓ Easy and inexpensive 	<ul style="list-style-type: none"> ✗ Inefficient (67 calls for eligible contact in Ghana IVR study) and varies by mobile operator ✗ Numbering system typically begins with mobile operator and doesn't have geography ✗ Designs typically limited to stratification by operator ✗ Lower response rates (vs. re-contact) 	
Re-contact from previous survey	<ul style="list-style-type: none"> ✓ Higher response rates than RDD ✓ Lots of data for non-response adjustments 	<ul style="list-style-type: none"> ✗ Limited sample size ✗ Many people switch phone numbers (after 6 months, 43% of numbers in Liberia switched off) 	
List sample from vendor	<ul style="list-style-type: none"> ✓ Rich auxiliary information for sample targeting and non-response adjustment 	<ul style="list-style-type: none"> ✗ More expensive ✗ Probability ✗ Panel conditioning effects ✗ Potential for coverage error when lists exclude certain mobile network operators 	
List sample from mobile network operator	<ul style="list-style-type: none"> ✓ Large sample size ✓ Potential for auxiliary information 	<ul style="list-style-type: none"> ✗ Nearly impossible to get for most researchers (exceptions: 2009 Mexico H1N1 study) 	

3.2 Sampling - People and Mobile Phone Numbers

- Many people in LMICs share phones (SIMs)
- Many phones in LMICs have multiple phones (SIMs)
- Shared and multiple phones affect an individual's probabilities of selection
- [Labrique et al. \(2017\)](#) simulations
- Solutions: Collect data on multiple/shared phones in surveys, use those data for post-survey adjustments
 - Limited impact of this method on survey estimates for CDC NCD Program (unpublished data)



3.2 Sampling - Working and Non-Working Numbers

- In high-income countries, non-working numbers return a clean, standardized error code (“this number is not working”)
- Challenges with error codes in LMICs
- Non-working numbers often are classified as working numbers, which can artificially depress response rate calculations
 - Many RDD response rates in the literature are underestimated
- Solution: Screen RDD sample through database to filter out non-working numbers (as well as numbers out of the country)

3.3 Data Collection

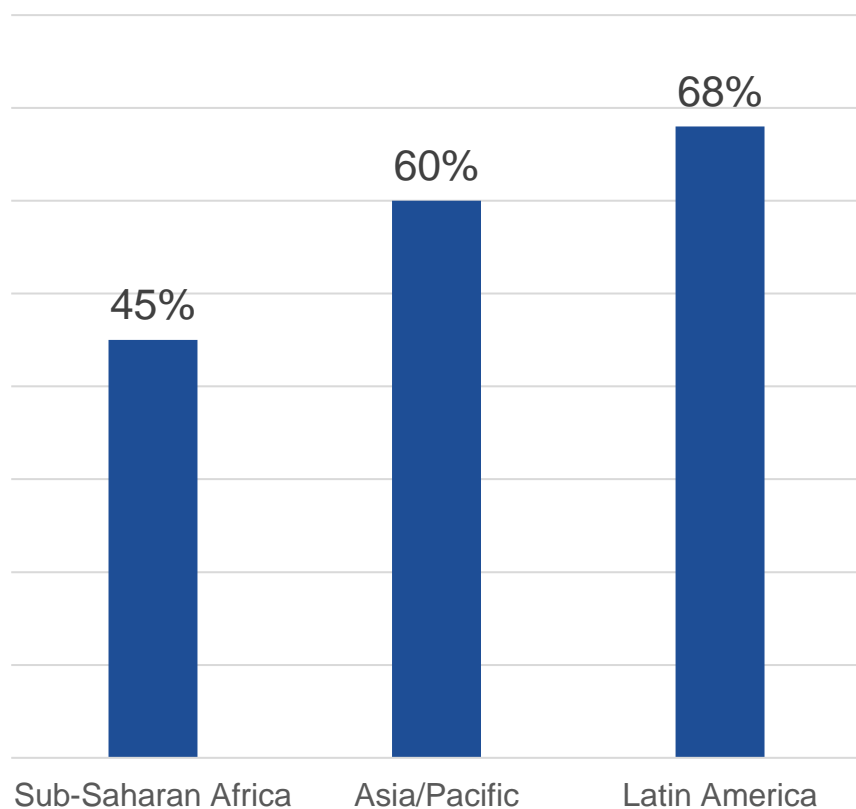
- Calls typically made ~8am to ~8pm, 6-7 days
- Number of recontacts varies, but typical is 4-8 recontacts
 - [World Bank Solomon Islands CATI survey](#): 70% completes on first call, 4% completes on fourth+ calls
 - Less research about optimal number of recontacts in LMICs
- Most countries don't regulatory approval for RDD (even for IVR/SMS), but others do (e.g., Uganda)
- For more practical information, see World Bank's [handbook](#) and other resources at the end of the presentation



4a. Survey Error: Coverage and Non-Response

4.1 Coverage Error

Penetration of Mobile Phones
(2018; unique subscribers)

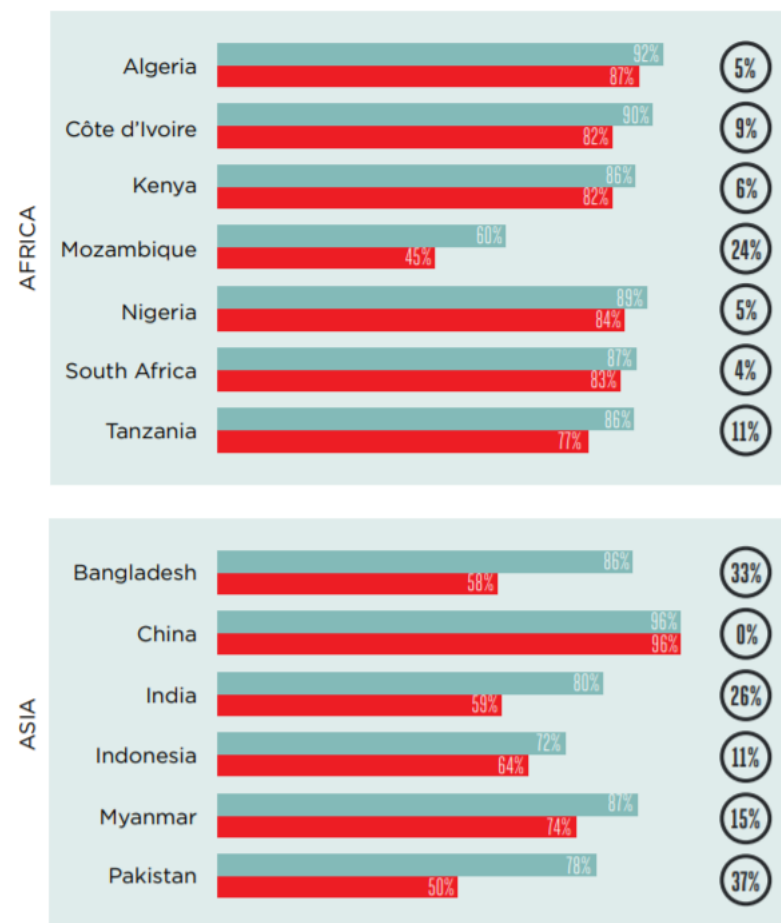


Source: GSMA, 2020



X% = Gender Gap

MOBILE OWNERS (%)



Source: GSMA, 2019

4.1 Coverage Error - Complexities

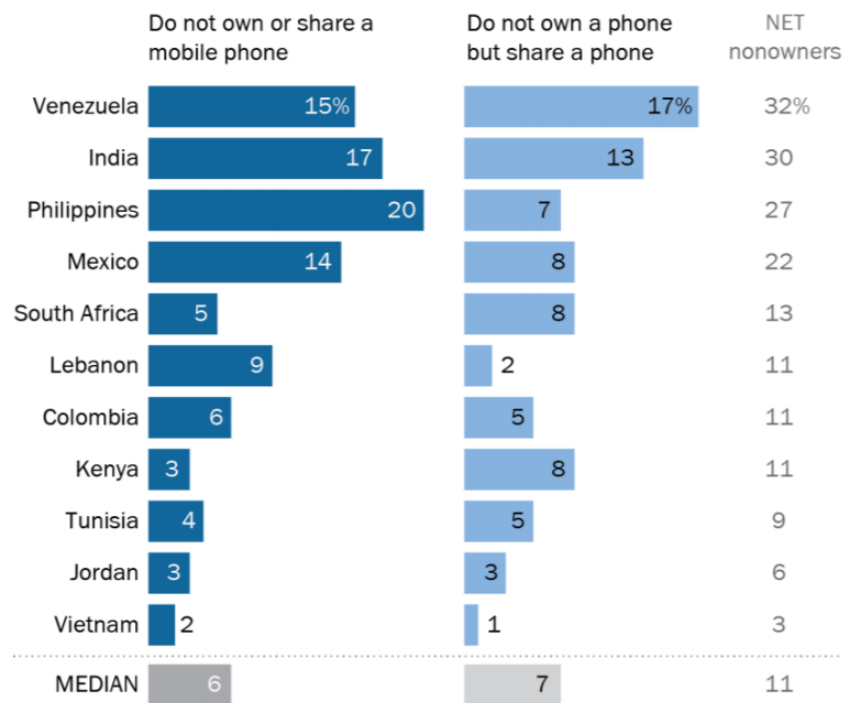
- Phone sharing is common, and varies by country
- Women and other groups more likely than men to **have access** but not own phone

Complexities in Measuring Coverage

1. Coverage statistics based on **ownership** may overstate coverage errors because of sharing
2. Family members don't share phones equally
3. Some people who share phones may not accept incoming calls

In some emerging economies, many do not own – or even share – mobile phones

% of adults who ...



Source: Mobile Technology and Its Social Impact Survey 2018. Q4 & Q5.
"Mobile Divides in Emerging Economies"

PEW RESEARCH CENTER

[Source: Pew Research Center, 2019](#)

4.2 Non-Response Error: Reasons for Non-Response

Non-contact

- Poor network connection
- People keep phones off
- Miss calls (IVR, CATI)
- Differential access to phone (phone sharers)
- Can't navigate to website (mobile web)

Refusal

- Worry about spam
- Distrust of surveys
- Concerns about data charges, even if supposed free

Language barrier

- Language not offered in survey
- Problem with language selection question

Breakoff

- Long surveys
- Lack of digital literacy
- Difficulty inputting numeric answers
- Problematic questions

Literacy

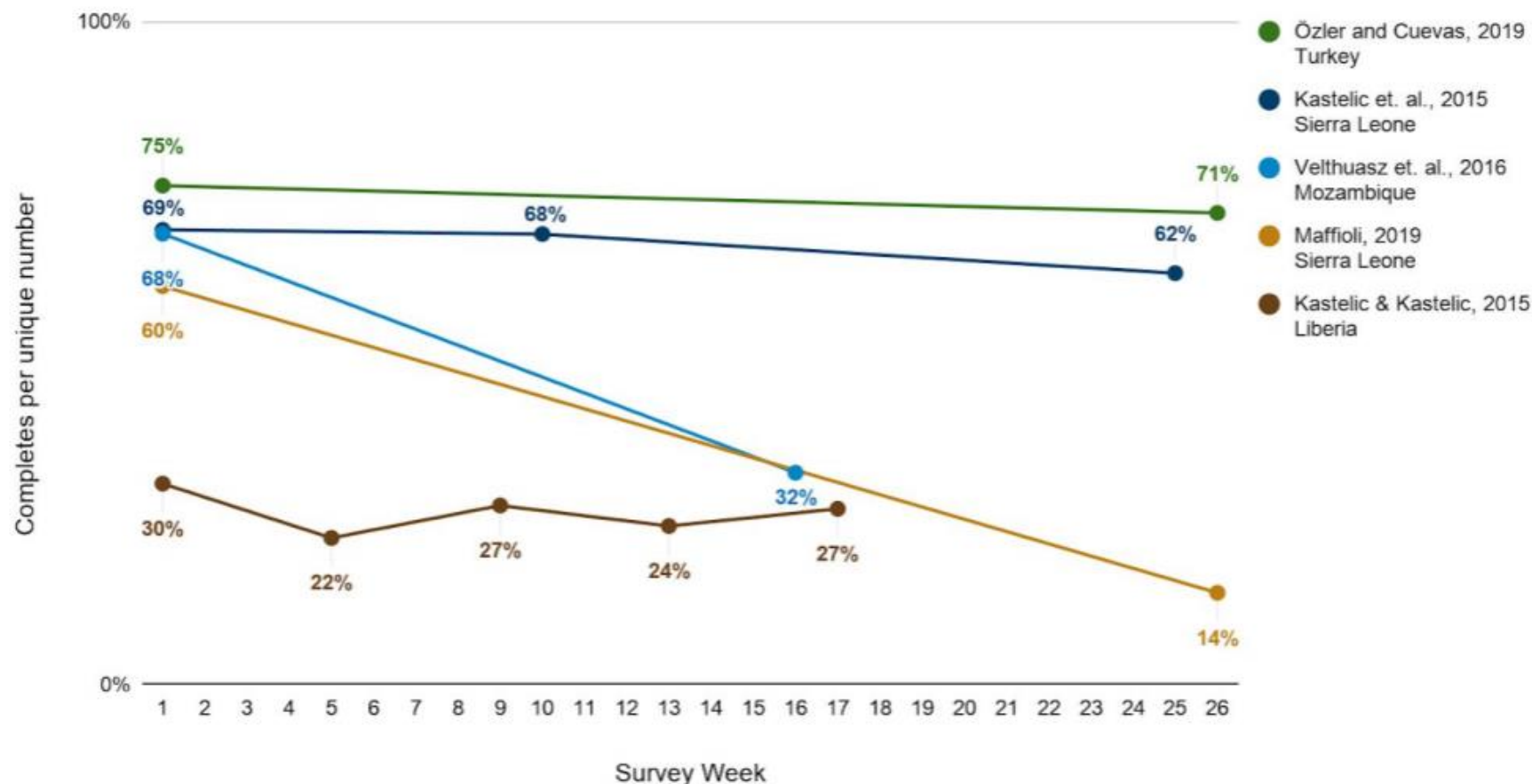
- Literacy barrier for SMS/web

4.2 Non-Response Error: Response Rates

	Cross-section	Re-contact
CATI	<p>~ 5-25%</p> <p>Lau et al. (2019): 15% Pariyo et al. (2017): 6%, 9% WB Solomon Islands: 46% (high)</p>	<p>~ > 50%</p> <p>Pariyo et al. (2017): 54%, 50% WB Covid Survey: >75% Ballivian et al. (2014): 39%-88%</p>
IVR	<p>~ 1-10%</p> <p>Leo et al. (2015): 0.7%-1.1% Pariyo et al. (2017): 0.9%, 1% Lau et al. (2019): 3% Greenleaf et al. (2020): 10% L'Engle et al. (2018): 21% (high)</p>	<p>~ 30%</p> <p>Pariyo et al. (2017): 31%, 42% Ballivian et al. (2014): 40%, 20%</p>
SMS	<p>~< 5%</p> <p>Lau et al. (2019): 0.2% Lau et al. (2019): 0.3-14%</p>	<p>~ 10-50%</p> <p>Ballivian et al. (2014): 30%, 45%</p>

4.2 Non-Response Error: High-Frequency Phone Surveys

Figure 5: Response Rates Across Panel Waves



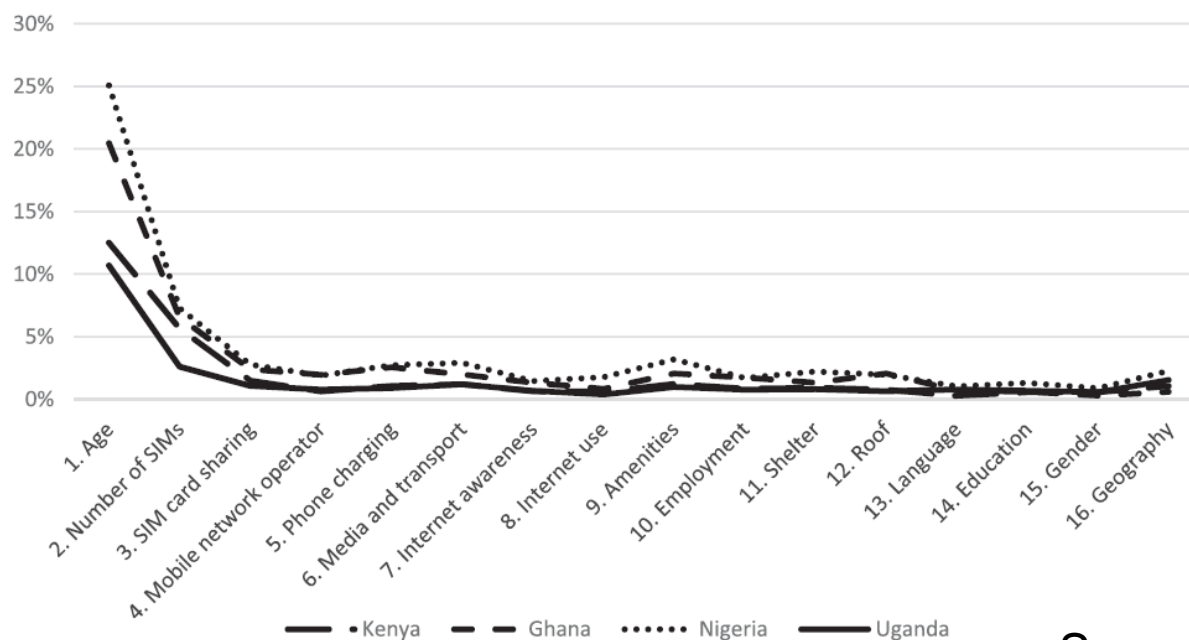
Source: [Innovations for Poverty Action \(2020\)](#)

4.2 Non-Response Error: Sources of Non-Response

- Most non-response is due to non-contact, refusal rates are low
- Breakoff rates in [Lau et al. \(2019\)](#) by country: 46% Nigeria, 38% Ghana, 20% Uganda, 13% Kenya
- Breakoff more common among older people

Figure 1

Question-level breakoff rate



Source: [Lau et al. \(2019\)](#)

4.3 Impact of Coverage and Non-Response on Estimates - Summary

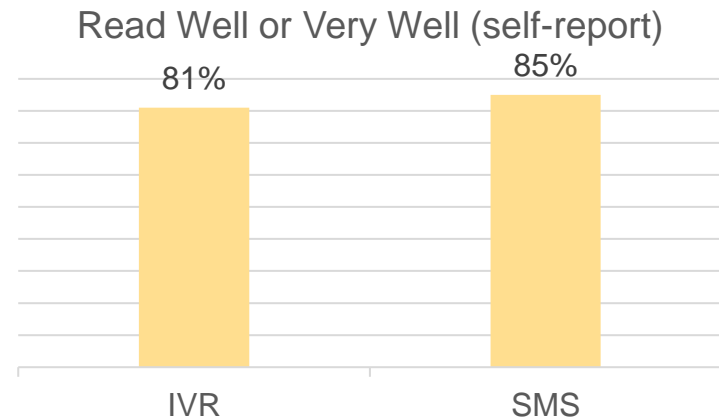
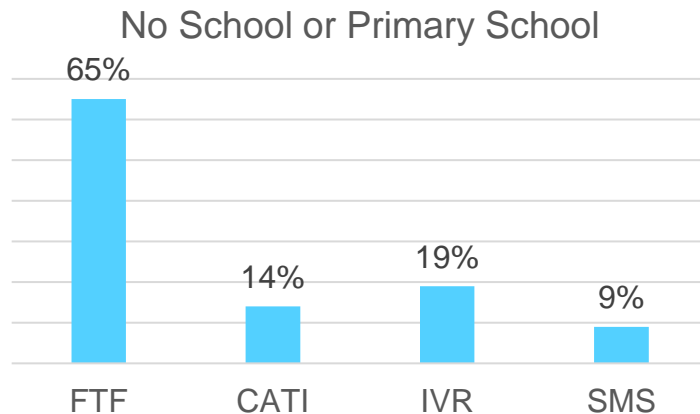
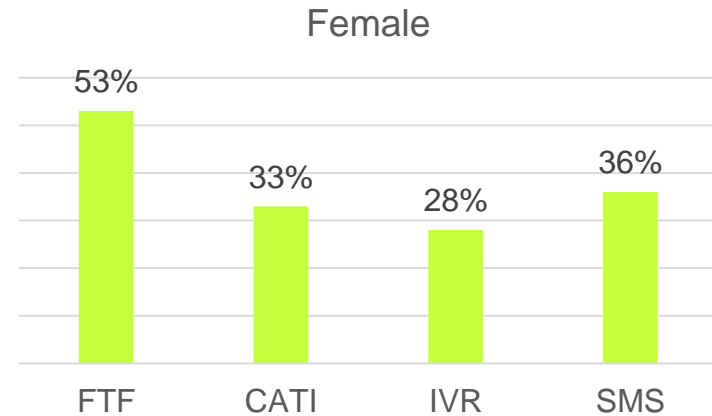
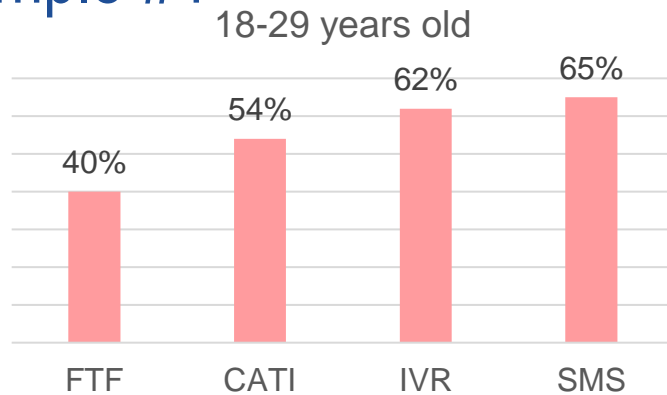
Groups typically underrepresented in cross-sectional and longitudinal mobile phone surveys in LMICs

- Women (especially women over 45)
- Older people
- Lower socioeconomic status: Less educated, lower income, assets
- People who share phones
- People who report difficulty charging phones

Separating Coverage and Non-Response

- [Lau et al. \(2019\)](#)
- [Greenleaf et al. \(2020\)](#)

4.3 Impact of Coverage and Non-Response on Estimates – Example #1

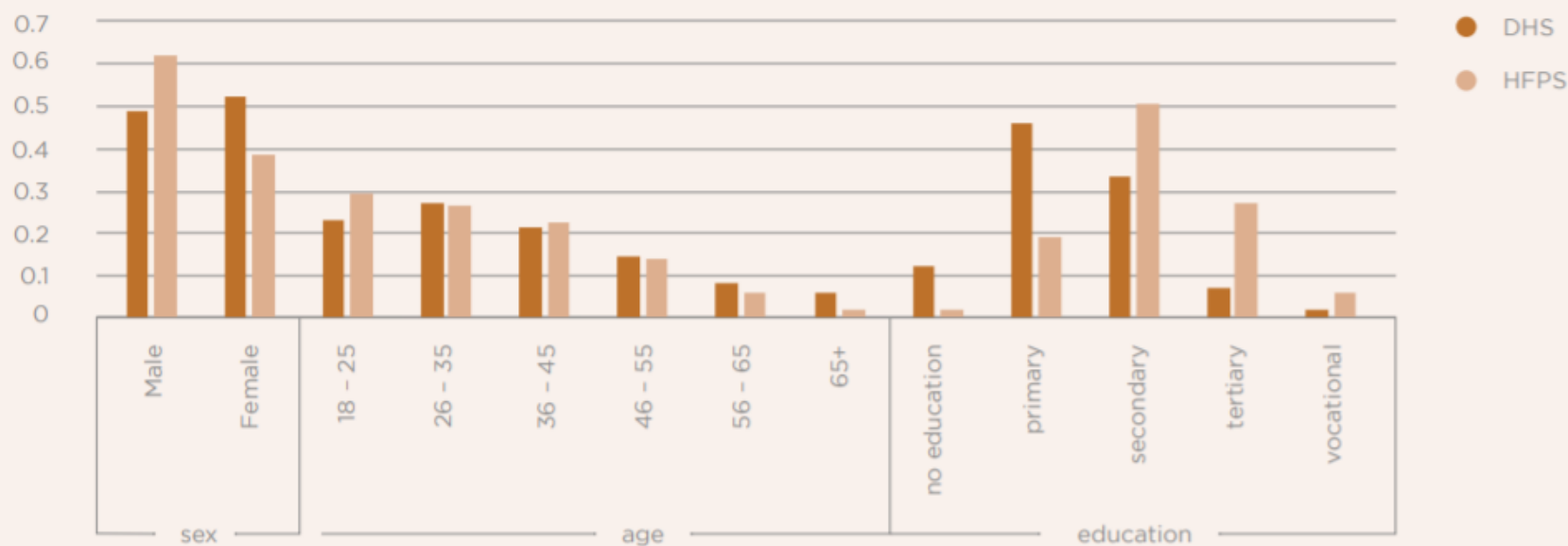


[*Lau et al. \(2019\)*](#)

Notes: The Nigeria Demographic and Health Survey is the reference face-to-face (FTF) survey. CATI, IVR, SMS used RDD samples.

4.3 Impact of Coverage and Non-Response on Estimates – Example #2

Figure 29. Comparison of characteristics between mobile phone survey (unweighted) and the 2015 DHS (weighted)



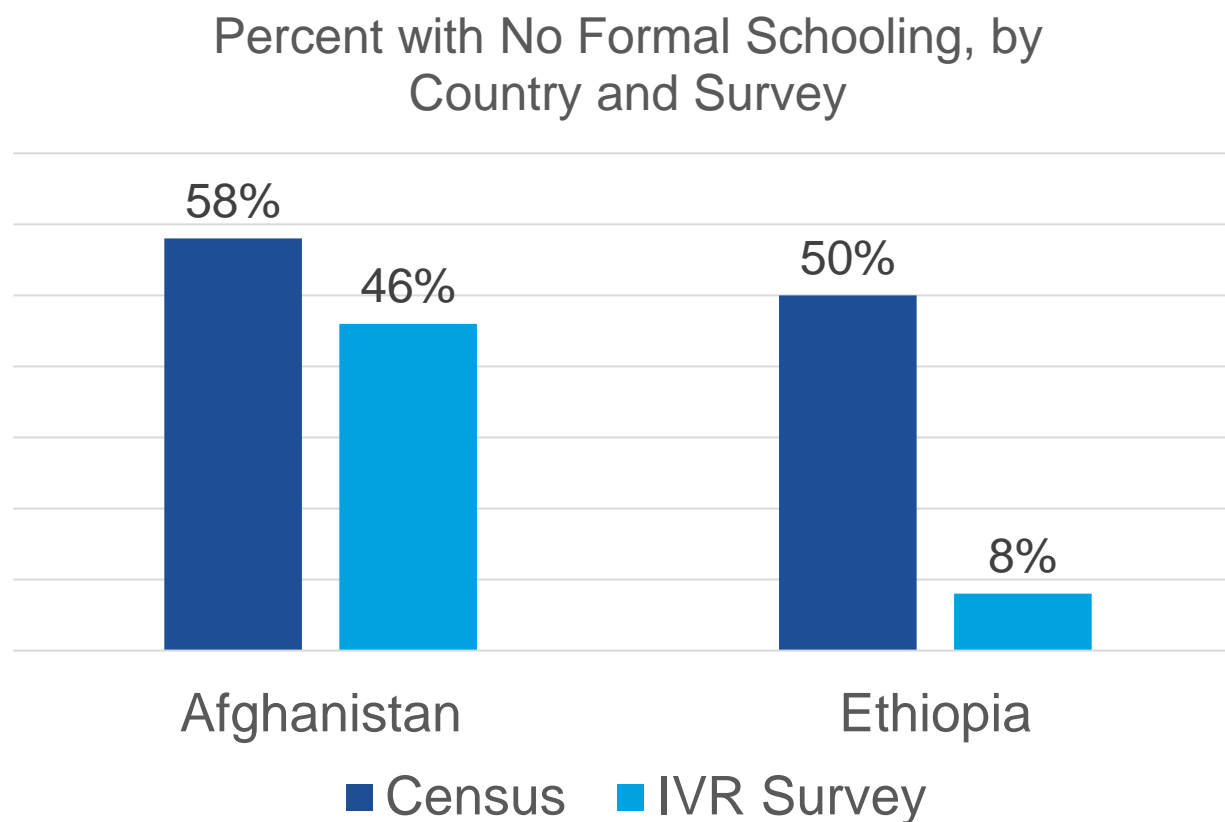
Source: [WB Solomon Islands RDD Survey \(2020\)](#)

DHS = Demographic and Health Survey (reference face-to-face survey)

HFPS = High Frequency Phone Survey (CATI)

4.3 Impact of Coverage and Non-Response on Estimates – Beware of Blind Focus on Response Rates!

[Leo et al. \(2015\)](#): RDD IVR Surveys in Afghanistan and Ethiopia have roughly the same response rate (1.11% and 0.80%), but representativeness is quite different.



4.4 Reducing Coverage Error

- Conduct surveys in countries with higher mobile phone penetration
 - [Leo et al. \(2015\)](#) shows that there's better demographic representativeness in countries with higher mobile penetrations (but coverage and non-response are confounded)
- Distribute phones to people without phones
 - Example: [South Sudan Experimental Phone Survey](#)

4.4 Reducing Non-Response Error

1. Incentives
2. Shorter surveys
3. SMS prenotification
4. Concise introductions
5. Better language selector
6. Optimizing day/time contacts
7. Mixing modes
8. Targeted sampling
9. Weighting: Calibration
10. Weighting: Multilevel Regression and Poststratification (MRP)

4.4 Reducing Non-Response (1/10): Incentives

Incentives are typically provided for IVR, SMS surveys: less common for CATI

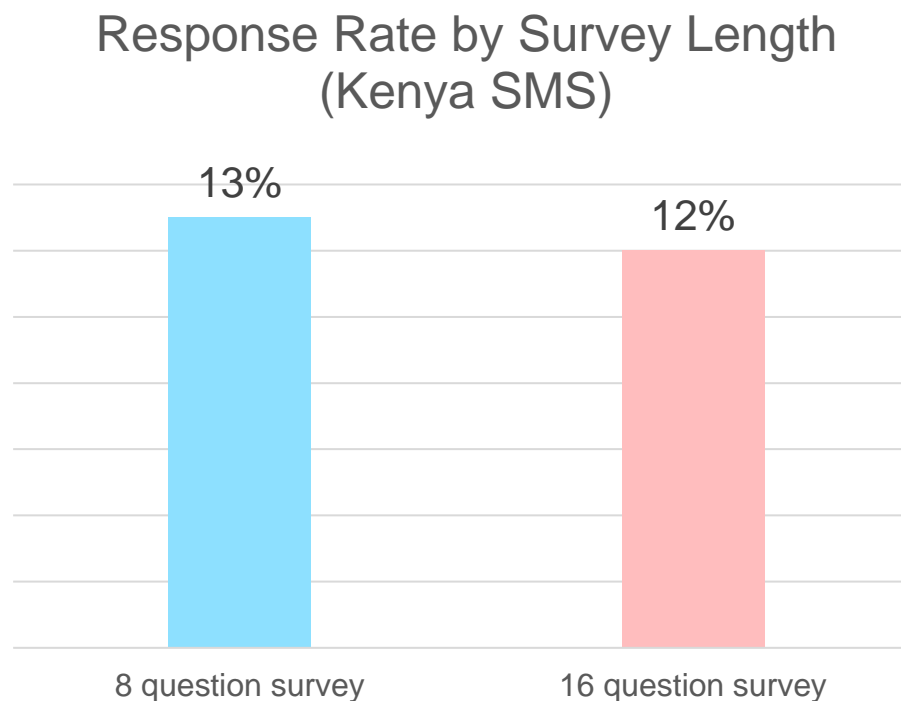
Key conclusions from experiments (e.g., [Gibson et al., 2019](#))

- For cross-sectional surveys
 - Offering any incentive improves response rates for IVR/SMS
 - Offering any incentive has unclear effect on response rates for CATI
 - Increasing amount of incentive (e.g., 0.5 USD to 1.25 USD) doesn't matter
 - Mixed evidence about pre-paid airtime versus lottery
- For longitudinal surveys: Offering incentive may have weaker effects

4.4 Reducing Non-Response (2/10): Shorter Surveys

Available evidence suggests that decreasing survey length is **not effective** for boosting response rates

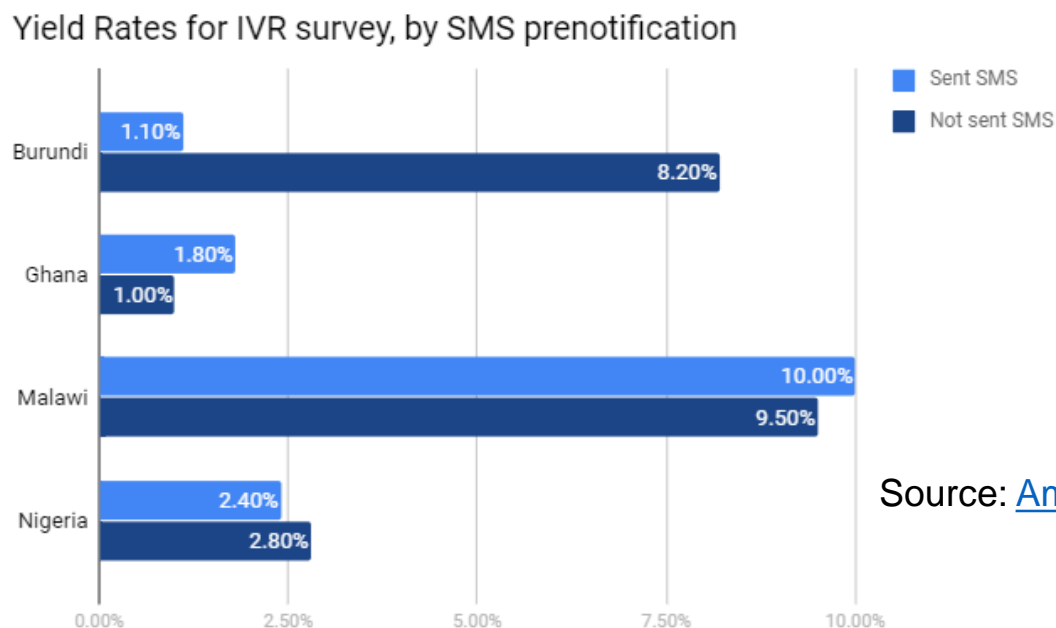
More research needed – especially for CATI



Source: [Lau et al. \(2019\)](#)

4.4 Reducing Non-Response (3/10): SMS Prenotification

- Prenotifications are typically 1 day before survey, but could also be 5 minutes ([Morse et al., 2016](#))
- SMS prenotifications increase IVR response rates by ~6-8 pps 😊 but also skew samples towards wealthier and more educated 😞 ([Morello and Leo, 2016](#))
- Effect may also depend on survey topic (see below): Burundi survey was on politics, others on less sensitive topics



Source: [Amaya et al. \(2018\)](#)

4.4 Reducing Non-Response (4/10): Concise Introductions

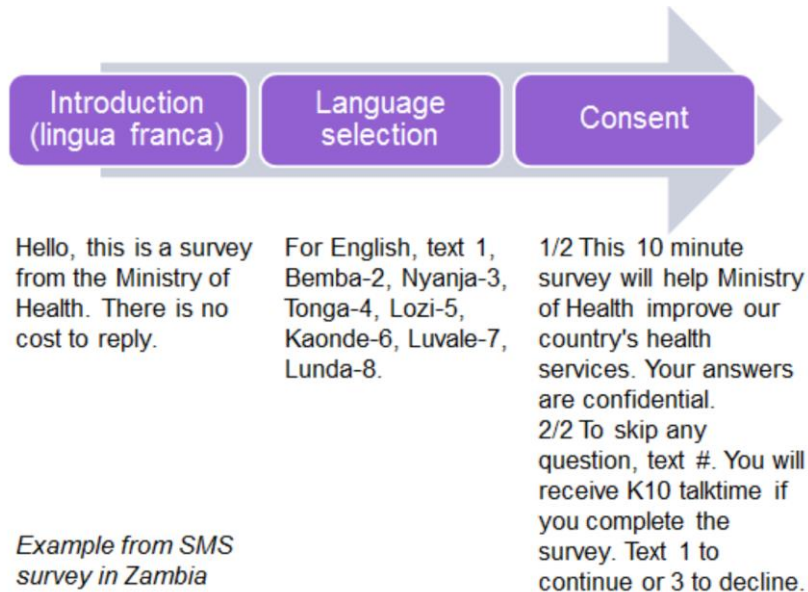
Preliminary conclusion: Shorter introductions increase response rates

[L'Engle et al. \(2018:3\)](#): *“Second, a shorter, straightforward message introducing the survey and providing essential elements of informed consent yielded higher call continuation than longer introduction messages that included an emotional appeal for participation.”*

CDC/RTI/InSTEDD Project in Zambia (unpublished data):

“You will receive K10 talktime if you complete this 10 minute survey. Your answers are confidential and will help improve our country’s health . Press 1 to continue or press 3 to decline.”	“You will receive K10 talktime if you complete this 10 minute survey. Your answers are confidential. Press 1 to continue or press 3 to decline.”
Response Rate: 0.42%	Response Rate: 0.62%

4.4 Reducing Non-Response (5/10): Better Language Selector



- Common problem: Surveys over-represent English speakers
- Potential reason: Primacy effect, with English often presented first
- [L'Engle et al. \(2018\)](#) experiment showed that randomizing order of languages led to greater diversity of region 😊
- More research needed on language selection

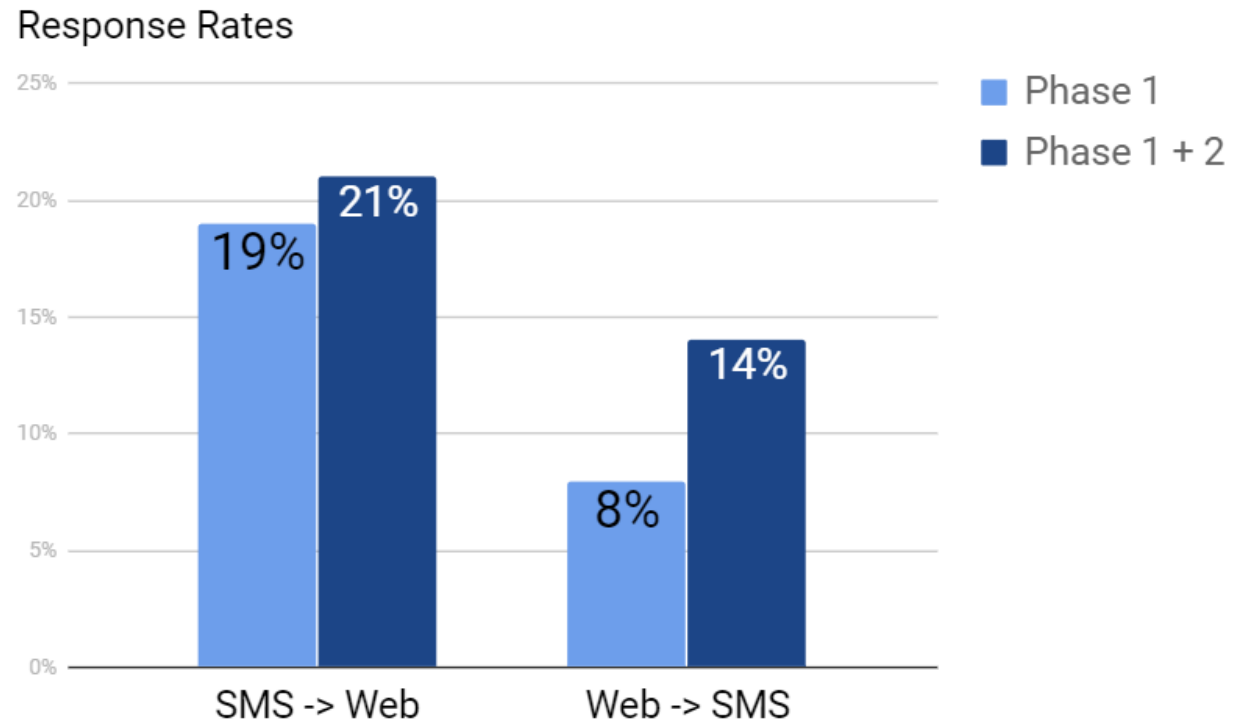
4.4 Reducing Non-Response (6/10): Optimizing Day/Time

- Limited literature
- Limited data from synchronous modes (CATI and IVR)
- For SMS, no effect of time day in Kenya ([Johnson, 2016](#))
- Unpublished data from Kenya SMS surveys (Lau)

<i>Kenya</i> ¹	Invitation Time				
	800	1030	1300	1630	1900
Monday	6.2%	7.9%	7.1%	6.9%	6.1%
Tuesday	6.3%	7.6%	5.5%	6.2%	5.4%
Wednesday	7.5%	7.0%	7.8%	7.3%	5.8%
Thursday	7.8%	7.3%	7.8%	8.1%	5.6%
Friday	5.4%	7.3%	7.1%	7.1%	6.7%
Saturday	5.6%	7.2%	6.2%	7.6%	6.6%
Sunday	6.6%	7.4%	8.2%	6.9%	6.5%

4.4 Reducing Non-Response (7/10): Mixing Modes


- Goal of mixing modes in LMICs is increasing response and sample quality: less about reducing cost (like in high-income countries)
- Adding follow-up mode can increase response, but depends on mode sequence (see right)




Source: [Lau et al. \(2018\)](#) study tracking job training program participants in South Africa

4.4 Reducing Non-Response (8/10): Targeted Sampling

- Papua New Guinea
- World Bank partnered with Digicel (operator)
- CATI Survey

WORLD BANK  BLOGS


HOME ALL BLOGS TOPICS CONTACT COVID-19 

Published on Development Impact


Improved targeting for mobile phone surveys: A public-private data collaboration: Guest post by Kristen Himelein and Lorna McPherson


KRISTEN HIMELEIN | JANUARY 19, 2021

This page in: [English](#)

 Mobile phone surveys have been rapidly deployed by the World Bank to measure the impact of COVID-19 in nearly 100 countries across the world. Previous posts on this blog have discussed the sampling and implementation challenges associated with these efforts, and coverage errors are an inherent problem to the approach. The survey methodology literature has shown mobile phone survey respondents in the poorest countries are more likely to be male, urban, wealthier, and more highly educated. This bias can stem from phone ownership, as mobile phone surveys are at best representative of mobile phone owners, a group which, particularly in poor countries, may differ from the overall population; or from differential response rates among these owners, with some groups more or less likely to respond to a call from an unknown number. In this post, we share our experiences in trying to improve representativeness and boost sample sizes for the poor in Papua New Guinea (PNG).









[link](#)

Round 1

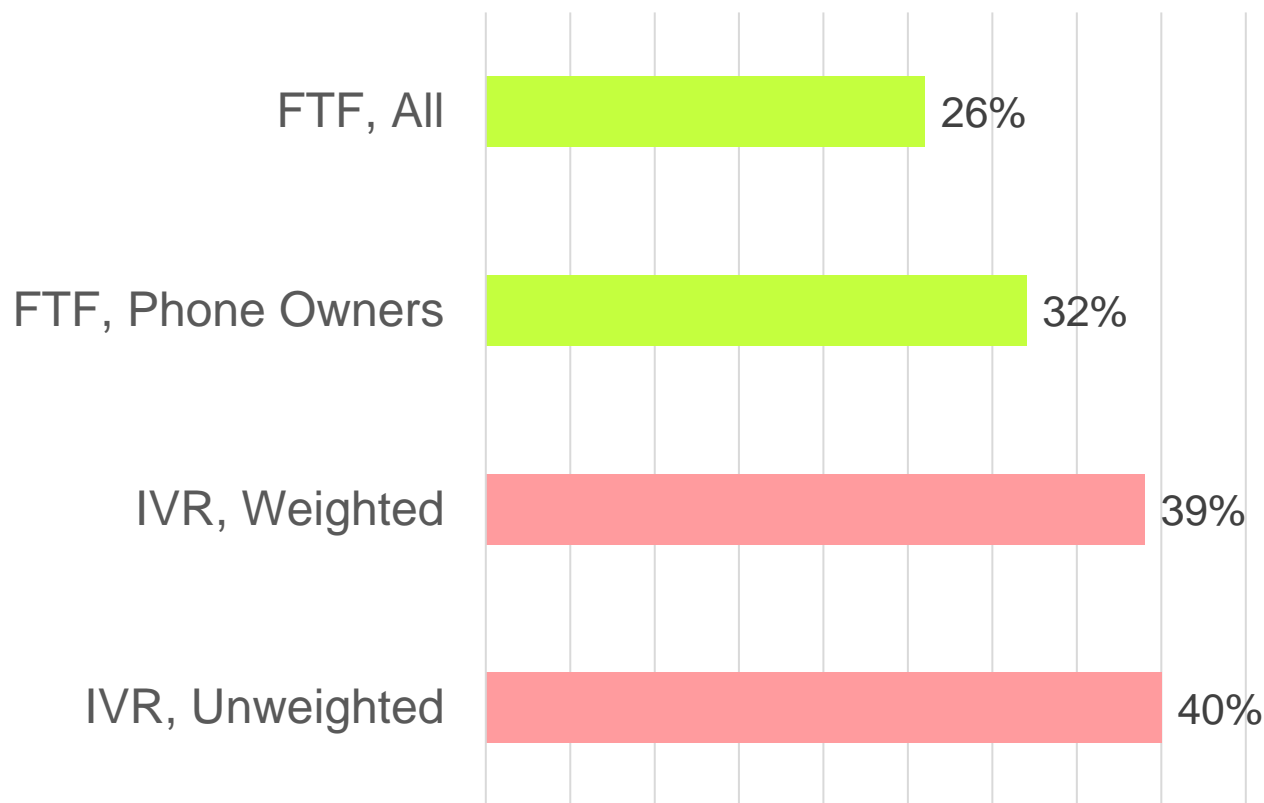
- RDD, stratification based on last user ping
- Skewed wealthy, relative to benchmark data

Round 2

- Follow-up of Round 1
- Also targeted users that didn't send SMS (not literate) or received only incoming calls (poorer)
- Targeting captures less wealthy sample, improving sample quality

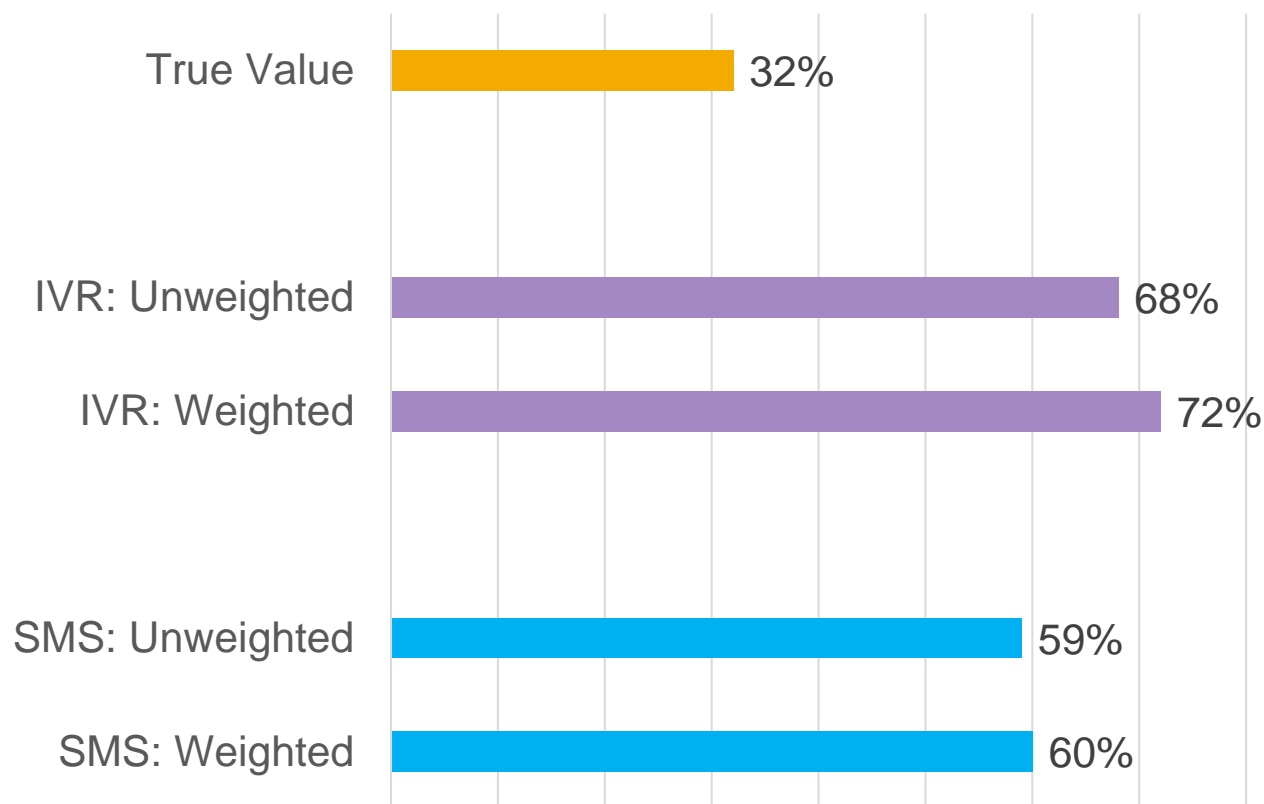
4.4 Reducing Non-Response (9/10): Weighting - Calibration

[Greenleaf et al. \(2020\)](#) IVR overstates contraceptive use among women in Burkina Faso, weighting has minor impact



4.4 Reducing Non-Response (9/10): Weighting - Calibration

[Lau et al. \(2019\)](#): IVR, SMS overestimate voting, and calibrating to control totals doesn't decrease bias



4.4 Reducing Non-Response (10/10): Weighting - MRP

Multilevel Regression and Poststratification (MRP): Weighting method for correcting model-based estimates:

- [WB Solomon Islands](#) application to estimating job loss in CATI survey
- [Gellar et al. \(2021\)](#) application to financial inclusion in Uganda SMS surveys



4b. Survey Error: Measurement

4.5 Measurement Error: Challenges

All Modes

1. No in-person interviewer
2. Questions are shorter
3. Maybe not first language
4. May not be reaching correct person (wrong respondent)
5. Hard to measure urban/rural
6. Potential for primacy or recency effects

IVR and SMS

1. No interviewer
2. Unfamiliar mode
3. Limited number of questions
4. Limited space for questions, instructions
5. No visuals

SMS-specific

1. ~ 160 characters for question and response options
2. Small screens and old phones, scrolling is hard

4.5 Measurement: Reliability

- Very few examples of studies that identify measurement error explicitly
- Most studies examine reliability

[Pariyo et al. \(2017\)](#) – Bangladesh, Tanzania

- High reliability for demographics, simple health risk factor questions (e.g., tobacco, alcohol).
- Low reliability for complex questions (e.g., fruit/vegetable consumption)
- No differences between CATI and IVR

[Mafoud et al. \(2015\)](#) - Lebanon

- High reliability of face-to-face survey and CATI follow-up survey, limited to simple health questions

[Ballivian et al. \(2014\)](#) – Peru, Honduras: IVR appears to be more reliable than SMS, CATI

4.5 Measurement: Sensitive Topics

Do mobile phone surveys produce more reports of socially undesirable behaviors compared to face-to-face surveys?

- Literature is limited
- [Mafoud et al. \(2015\)](#) find higher reports of alcohol consumption in CATI compared to face-to-face (Lebanon)

Which mobile phone modes produce more reports of socially undesirable?

- Literature is limited
- [West, Ghimire, Axinn \(2015\)](#): SMS has more reports of sensitive topics (e.g., smoking marijuana, mental health) than CATI (Nepal)

4.5 Reducing Measurement Error: Primacy Effects

Minimize primary effects through randomizing response options or combining response options

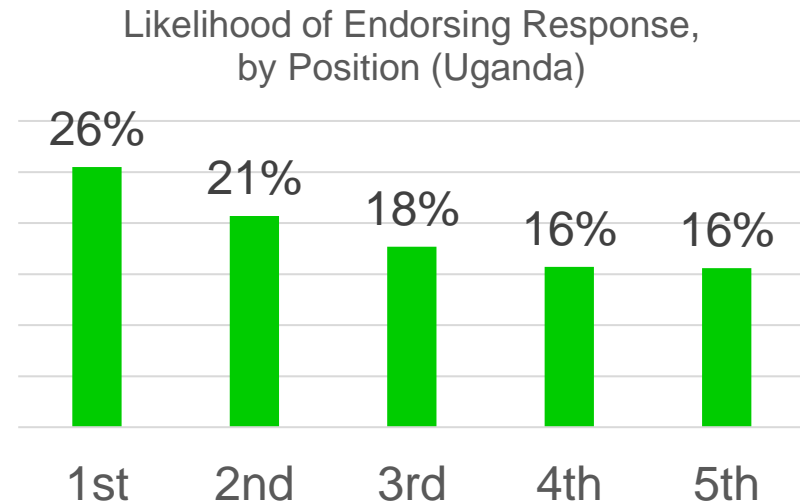
Experiments in multi-country African IVR surveys show primacy effects for:

- Urban versus rural
- Zone (Nigeria)
- Satisfaction

For satisfaction, authors moved from 4-category to 2-category responses

Source: [Leo, Kalow, Moss, 2018](#)

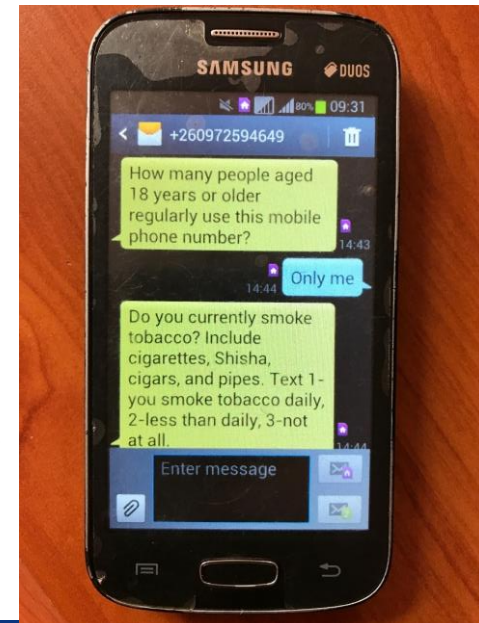
Experiments in 4 African SMS surveys show consistent evidence of primary effects



Source: [Lau, Sanders, and Lombaard, 2019](#)

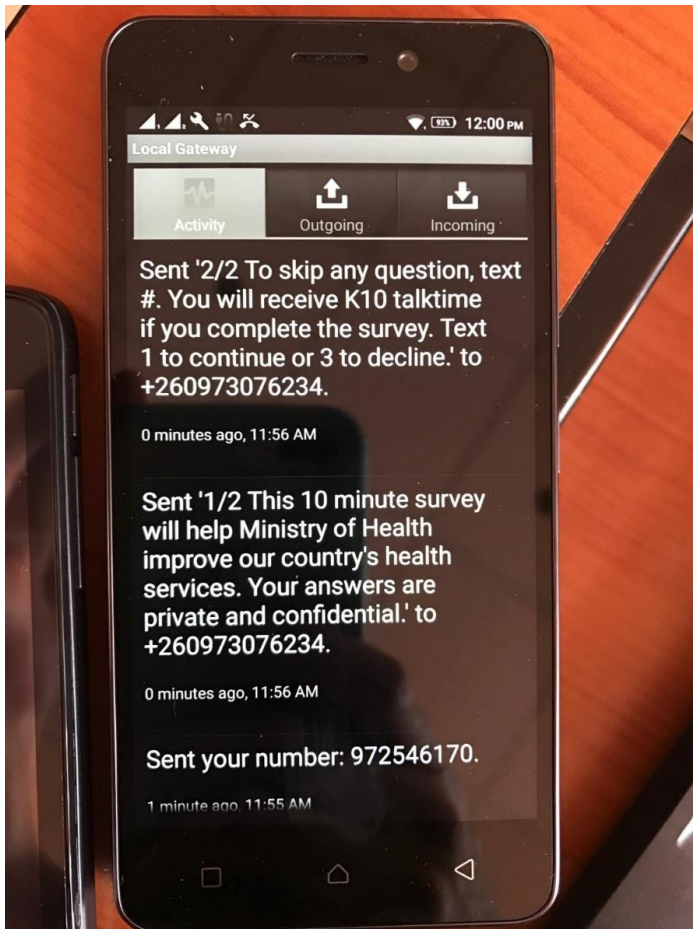
4.5 Reducing Measurement Error: Questionnaire Design

1. Test **locally, and in-person** ([Firchow and Mac Ginty, 2017](#))
2. Avoid multi-select options ([Lau, Sanders, and Lombaard, 2019](#))
3. Avoid embedding response options in question ([L'Engle et al., 2018](#))
4. Use categories instead of exact numeric responses ([L'Engle et al., 2018](#))
5. Consider a modular design for longer surveys ([West, Ghimire, Axinn 2015](#))
6. Be careful when asking about shared/multiple phones
7. Consider requesting confirmation for high value survey items
8. Don't forgot basic instructions, especially for numeric responses



4.5 Reducing Measurement Error: Multiple SMS

Multiple messages may come out of order



Respondents may have difficulty scrolling





5. Ethics

Ethical Issues for Consideration

[Mwaka et al. \(2019\)](#) qualitative study on consent in Uganda

- Balance between comprehensiveness and burden
- Distinction between surveys for **research** versus **surveillance**
- Suggestions:
 - Provide prenotification (SMS) with ways of getting more info
 - Provide opportunities for callbacks
 - Be clear about confidentiality (phone numbers are linked to national ID)

[Rodriguez-Patarroyo et al. \(2020\)](#) Informed consent in Colombia

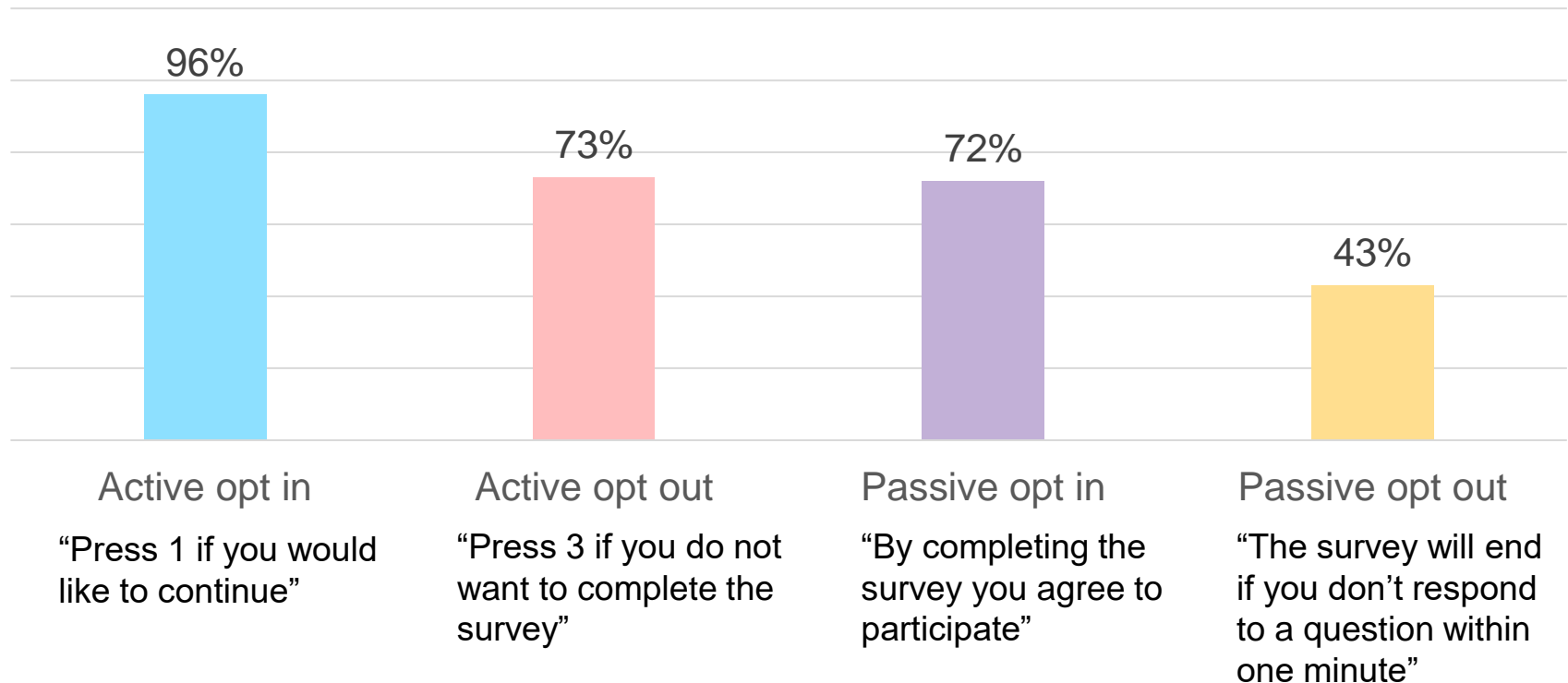
- Debate among experts about role of consent in research/policy
- Importance of community sensitization, trust building

Ethical Issues for Consideration

[Ali et al. \(2019\)](#) global survey of 114 mHealth stakeholders

- “current ethics and regulatory requirements associated with the conduct of MPS [mobile phone surveys] are clear:” 73% disagreed!

Acceptability of Different Consent Methods for Research





6. Discussion

Other Resources

- Innovations for Poverty Action [RECOVR Initiative](#) ([handbook](#), [research synthesis](#), [research hub](#), [COVID-19 questionnaire repository](#))
- World Bank Mobile Phone Panel Surveys [handbook](#)
- World Bank COVID-19 Surveys [technical note](#)
- [JPAL's guidance](#) (and [example protocol](#), [South Africa learnings](#))
- WAPOR [list of public opinion surveys](#) on COVID-19



Thank you

Contact: Charles Lau | email: clau@rti.org