

Centre for Rapid Insights: Our Methodology

By Priya Vedavalli and Vibhav Mariwala

Introduction

This note captures the key learnings from the expert interviews and pilot studies we conducted to further strengthen our methodology. We consulted sectoral experts to identify important questions based on existing data gaps in the Indian landscape as well to understand how to overcome some of the limitations of IVR survey.

Experts we spoke to include:

- *Sonalde Desai*, Professor, University of Maryland College Park and Professor and Director, NCAER-National Data Innovation Centre (NDIC)
- *Sharan Sharma*, Assistant Research Professor, University of Maryland
- *Dr. Lancelot Pinto*, Consultant Respiriologist and P.D. Hinduja Hospital, Mumbai
- *Anand Deshpande*, Founder of Persistent Systems; Founder, deAsra Centre of Excellence in Nano Entrepreneurship, Gokhale Institute of Politics and Economics.
- *Rukmini S.*, Independent data journalist; CASI Fall 2022 Visiting Fellow

Based on the conversations and a secondary literature review we have captured the strengths and limitations of IVR surveys, sources of measurement error and identified certain areas of improvement for the tool we use as a part of CRI.

Strengths

1. **Spread and scale:** Suitable for large scale studies that need to be conducted across various regions in India since it is easier to translate and programme the survey questions to be run across states
2. **Reach:** Can also be run in regions with low smartphone penetration but have access to feature phones. The penetration of feature phones (vs. smartphones) is higher in India which makes it possible to reach vast sections of the population, particularly those belonging to low-income households
3. **Simple close ended questions:** Suitable for close ended questions with a small set of response choices, such as binary or select-one option type questions
4. **Time:** The survey responses can be collected with a matter of few hours which is a fraction of the time taken by in-person surveys
5. **Cost:** Costs compared to traditional in-person surveys or phone surveys are limited since there is no human intervention involved

Limitations

1. **High break-off rates:** If the survey is long, it becomes hard for the respondents to stick through the entire survey, without a human being present to motivate them to respond. The length of the survey should be less than four questions.

2. **Cannot be used for questions with over five response categories or open-ended questions:** The respondent might not be able to recall the corresponding response categories if there are more than five options.
3. **Hard to gauge if the respondent understands the questions:** It is hard to check if respondents understood the question accurately as there are no means to clarify/ask queries or repeat the question.

Sources of survey error

1. Measurement error

- a. **Interviewer bias:** The voice can be a key factor that influences the respondent's interest to complete the survey. Several studies also looked at how the gender of the automated voice, personalisation of a question can influence the responses. These are important considerations while designing the survey.
- b. **Questionnaire bias/error:** Questions in IVR based surveys are shorter, entailing yes/no questions, or select one option with a maximum of four options. This could lead to errors as the choice of IVR and question design may not be ideal for complex questions.
- c. **Respondent Bias/error:** There could be respondent bias/error due to several factors such as unfamiliarity with the technology, voices may be hard to follow (they may take some time to get used to the format and voice). Respondents may also find it difficult to hear as they may have to keep the headset away and respond to a question every time.
- d. **Bias Resulting from mode of administration:** Respondents may be more willing to respond to sensitive questions if they feel that the responses will be anonymous/confidential.

2. Nonresponse error

- a. **Unit nonresponse:** Certain members of the panel may be unavailable/unwilling to participate in the survey. This will affect the survey statistics. One needs to look at the panel and check for patterns in unit-level nonresponse so that the responses can be weighted accordingly.
- b. **Item nonresponse:** Item nonresponse refers to not answering a particular question in a survey. This will also affect survey statistics. It is important to keep track of those who picked up the call but did not answer all the questions so that the subsequent responses can be reweighted.
- c. **Break-off rates compared to other modes:** Break-offs rates are high in IVR-based surveys as the respondent may not be engaged. Asking a few non-sensitive questions might be one way to keep them engaged.

3. Coverage error

- a. **Phone coverage:** Since access to phones are critical, there is coverage error in terms of the people we will be able to reach using IVR surveys

- b. **Geographic coverage:** It is critical to understand the states and districts covered and not adequately covered by the sample.

Ways to strengthen the CRI tool

1. **Tracking nonresponses:** Tracking both unit and item level non responses is critical to ensure a representative sample.
 - a. Unit-level: For example, if women of a particular age group are unlikely to respond between 9-10 am, it is critical to reweight/impute the responses of those who respond from that demographic group.
 - b. Item level: For example, if women of age 20-30 did not answer the second question in a two-question survey but answered the first question, it is critical to reweight/impute the responses to ensure accurate results.
2. **Having a repeat the question option:** We currently do not give the respondents the choice of listening to the question again. Having this feature will help reduce respondent bias. It could also help to design longer surveys with more options. Other areas of improvement to reduce respondent bias include:
 - a. Allowing sufficient time (5 seconds) for a respondent to enter their response
 - b. When asking personal questions, one should give the option of not responding "If you choose not to respond, press '0.'"
3. **Tracking time taken to respond:** Certain respondents might be responding before listening to the question. To ensure those responses are removed, it is important to track the time taken for each respondent. Additionally, one can also enable the acceptance of response only after the question is played and does not accept responses for the next questions unless heard.

Note: This is a work in progress and will be updated after further inputs from relevant stakeholders.

Priya Vedavalli and Vibhav Mariwala are, respectively, Principal and Consultant at Artha Global.

References

de Leeuw, E. D., Hox, J. Dillman, D. A. (2008). IVR: Interactive Voice Response. In the International Handbook of Survey Methodology (pp. 285–298). New York: Routledge.