

# Respondent, a Key Factor in National-Level Surveys

## An Analysis

K N PATHAK, J S TOMAR

Statistics holds paramount position as a scientific tool to translate relevant data into information for effective policymaking. But the data collector does not receive adequate cooperation from respondents. At times, the respondents are incapable of providing information. The 61st, 66th and 68th rounds of the National Sample Survey have brought these problems to the fore. Lack of clarity marred the information received from about 20% respondents. This analysis presents some ways and means to make a survey/census more productive and useful for development planning.

The views expressed in this article are those of the authors and not of the institution to which they belong. The authors are grateful to an anonymous referee for very detailed comments and suggestions on an earlier version of the article.

K N Pathak ([knpathak@gmail.com](mailto:knpathak@gmail.com)) is the Joint Adviser in the NITI Aayog, Government of India and J S Tomar ([jeet.iasri@gmail.com](mailto:jeet.iasri@gmail.com)) is Deputy Director in the Central Statistics Office, Ministry of Statistics and Programme Implementation, Government of India.

Statistics holds a paramount position in developing planning as a scientific tool to translate relevant data into information for effective policymaking. The development plans and policies of any country are formulated on the basis of data collected from the ground level. They provide information, both for policy formulation and its evaluation, so that needful interventions could be made whenever necessary. Surveys are conducted to collect different kinds of data from various categories of stakeholders. India too has a long history of surveys and censuses. Even in the British era, the government got a survey/census conducted at specific intervals. Much importance was attached to the survey/census by the government after independence. In the early decades of planning in our country, P C Mahalanobis developed and used statistical tools and large-scale sample survey data in various fields of planning and policy.

The quality of statistics plays a very important role in the formulation of development and investment policies. It is also crucial for improving transparency and accountability in policy planning and implementation with better governance and management. It also helps in having greater control on delivery of public services. However, it is observed that the data collected from various stakeholders is not utilised to the extent it is envisaged for. There are a number of issues to be considered from the point of view of conceptualisation of a scheme/project for which the data is collected, the modality which is adopted for collecting the data and finally the way it is utilised.

For the policymakers, it is desirable that the quantum of data to be collected is optimised and the procedure is simplified with

a view to make it less cumbersome, time saving, more useful and result-oriented.

Some of the key issues faced during data collection for different national-level surveys could be listed as follows: (1) Illiteracy/lack of awareness on the part of respondent; (2) no visible benefits for the respondent; (3) too many surveys or successive surveys coming in a row; (4) respondent's fear about sharing the information; (5) absence of appropriate statutory laws for orienting the respondent for sharing the information; (6) unavailability of trained and experienced manpower; (7) lengthy questionnaire; (8) complexity of questions, often being beyond the understanding of respondent; (9) approaching the right person at the right time; (10) wrong information intentionally being provided by the respondent to hide her/his real status (for example, a well-off person claiming herself/himself under the category of below the poverty line (BPL) and vice versa).

### National Sample Surveys

Before we take up the remedies for the above-mentioned issues, it would perhaps be relevant to briefly touch upon recent experiences from the National Sample Survey Office (NSSO) surveys. The intricacies of a national-level survey could be better elaborated through the schedules used. For instance, one of the schedules used for the Employment–Unemployment Survey has a question pertaining to response code. In the response code, there are five options, namely: (i) informant: cooperative and capable; (ii) cooperative but not capable; (iii) busy; (iv) reluctant; and (v) others.

Every respondent covered under the said survey will compulsorily fall under any of the above five options. The level and quality of data/information obtained from a respondent depends upon her/his behaviour (cooperativeness) and capability. Out of these options, only option (i) that is, informant: cooperative and capable, fully meets the objective of the survey. The respondents from all the other options are not able to provide the information of desired level and quality.

## NOTES

Further, it also dilutes the level and quality of overall information gathered from option (i) Thus, the inference drawn from such data would also lead to distortion.

The data obtained through the 61st (2004–05), 66th (2009–10) and 68th (2011–12) rounds for the NSS Employment–Unemployment Survey leads to a number of important observations.

It has been observed that in the above-stated three rounds, an average of 80% respondents fall in category i. They are cooperative and capable, about 17% respondents fall in category ii. They are cooperative but not capable. About 2% respondents fall in category iii. They are busy and cannot respond to the survey. About 1% respondents are in categories iv and v. They are either reluctant or are in the category, “others.” In other

words, the responses of about 20% of the respondents—about a fifth—do not lead to clarity. The data generated through the information from such unwilling respondents often yields intriguing estimates. This also has its long-term repercussions on data collected from successive rounds. The other impact of such data coming below the level of expectation deprives the scholars, policymakers and other stakeholders to raise questions about the reliability and veracity of data.

Substitution of households is usual in large-scale surveys. Households are selected scientifically as per the standardised procedures. An emphasis is also laid on canvassing the schedules from originally selected households. However, if respondents are not providing the data due to various reasons, the data collector

has to substitute such originally selected household as per laid-down procedures to cover the requisite number of households. As observed from Table 2, in the said three surveys, an average of 3% substituted households were surveyed and the reason for substituting 20% households (that is, every fifth household) of these households was, that either the informant was busy or non-cooperative. The problem of response beyond the expected framework does not end here, even after substituting non-responsive households.

Thus, results of large-scale national surveys being analysed with such drawbacks would certainly lead to inappropriate policies at the national level. Hence, it is imperative that sincere efforts are made to find out its reasons and solutions thereof. The solutions of problems involved in this exercise require an in-depth analysis of the underlying problems.

For objective assessment of the situation, a sector-wise analysis of respondents was made. Moreover, we get a view of the trend prevalent among different categories of respondents which is illustrated in Table 1. For instance, there is a clear divergence of trend between rural and urban respondents. While the rural respondent is more cooperative, but less capable to respond, the urban respondent is capable, but more busy and reluctant. Such divergence has been observed in the previous nss rounds, more or less with the similar trend.

There could be unanimity on this issue that getting an appropriate reply from both rural/urban respondents is a challenging task. However, there are certainly ways and means to obtain appropriate replies from respondents of both the sectors. Within the sector (whether rural or urban), a clear dichotomy is observed, which is illustrated in Table 3.

From the table, it is observed that high-income group respondents in both the sectors are capable but more busy, reluctant and non-cooperative as compared to middle- and low-income group respondents.

The issues discussed above are commonly prevalent in case of large-scale national surveys in different countries.

**Table 1: Sector-wise Analysis of Respondents' Type** (%)

Response Code	61st Round			66th Round			68th Round		
	Rural	Urban	Total	Rural	Urban	Total	Rural	Urban	Total
Cooperative and capable	75.8	79.2	77.0	79.7	82.0	80.6	83.5	84.4	83.9
Cooperative but not capable	22.1	18.2	20.7	18.2	14.9	16.8	13.9	12.3	13.3
Busy	1.1	1.2	1.2	1.5	2.1	1.7	1.9	2.4	2.1
Reluctant	0.8	1.2	0.9	0.5	1.0	0.7	0.5	0.8	0.7
Others	0.2	0.2	0.2	0.1	0.1	0.1	0.1	0.1	0.1

Source: Unit-level data, NSS 61st round (July 2004–June 2005), 66th round (July 2009–June 2010) and 68th round (July 2011–June 2012): “Employment and Unemployment Situation in India,” National Sample Survey Office, Ministry of Statistics and Programme Implementation, Government of India.

**Table 2: Sector-wise Analysis of Households Surveyed and Reason for Substitution** (%)

Household Surveyed	61st Round			66th Round			68th Round		
	Rural	Urban	Total	Rural	Urban	Total	Rural	Urban	Total
Substitute	2.0	4.0	2.7	1.8	3.5	2.5	1.6	3.6	2.5
Reason for substitution									
Informant busy	0.1	0.1	0.1	0.1	0.2	0.1	0.1	0.1	0.1
Informant non-cooperative	0.2	0.6	0.3	0.2	0.5	0.3	0.2	0.5	0.3
Others	0.1	0.2	0.1	0.1	0.2	0.2	0.1	0.1	0.1
Total	0.4	0.9	0.5	0.4	0.9	0.6	0.4	0.7	0.5

Source: Same as Table 1.

**Table 3: Sector and Income Group-wise Analysis of Respondents' Type** (%)

Response Code	Income Group	61st Round		66th Round		68th Round	
		Rural	Urban	Rural	Urban	Rural	Urban
Cooperative and capable	High	79.2	82.5	81.7	84.4	85.8	87.1
	Middle	75.6	79.1	79.4	81.6	83.1	84.1
	Low	74.3	77.0	78.5	78.1	82.2	80.9
Cooperative but not capable	High	18.2	13.9	16.1	11.8	11.7	9.0
	Middle	22.2	18.3	18.6	15.4	14.3	12.7
	Low	23.8	21.0	19.5	19.2	15.4	16.2
Busy	High	1.2	2.0	1.6	2.5	1.9	2.7
	Middle	1.2	1.2	1.5	1.9	2.0	2.2
	Low	1.0	0.8	1.5	1.8	1.8	2.2
Reluctant	High	1.1	1.4	0.6	1.2	0.5	1.0
	Middle	0.7	1.2	0.5	0.9	0.5	0.8
	Low	0.7	1.0	0.4	0.9	0.6	0.7
Others	High	0.3	0.2	0.1	0.1	0.0	0.1
	Middle	0.3	0.2	0.1	0.1	0.1	0.1
	Low	0.2	0.2	0.1	0.1	0.1	0.1

Source: Same as Table 1.

The response rate obtained in a survey is a critical determinant of survey's data quality. Hence, survey organisations focus on assessing and improving response in continuing surveys by applying different ways and means which depend on the stage of development of that country and the existing political set-up. For instance, in China every family selected for a household survey has to sign an agreement to guarantee their accuracy in recording their daily income, expenditure, etc.

### Remedies

The remedy which flows from the diagnosis prescribes that on the one hand, efforts are needed to cultivate an inclination among all such urban respondents who are either reluctant or non-cooperative. On the other hand, for the rural respondents, who are cooperative but less capable, steps have to be taken for enhancing their capability.

The above-mentioned problems/issues could be addressed through the following measures:

**(1) Adequate Publicity/Awareness Campaign:** To enhance effectiveness of national-level surveys/censuses, at least 1% or 2% of the budgeted amount provided for the surveys/censuses should be spent on publicity/awareness campaigns. The use of electronic as well as print media along with social media would certainly yield desired results. Further, as local bodies/panchayati raj institutions have a direct/indirect influence at different levels in urban pockets/rural areas, their involvement would be very helpful in eliciting the requisite information from the respondents.

**(2) Changing Mindsets:** For making a survey/census a success, it is necessary to change the mindset of the respondents by convincing them that, may be there is no direct benefit to them by a survey, but the information to be provided by them is an essential input for formulating the national policies for development, which is ultimately for their welfare/benefit. Since every respondent spares her/his valuable time to patiently listen to each and every question of the investigator for providing desired

information, as a token of appreciation, the surveyor agency may provide a brief certificate to the respondent, duly acknowledging her/his contribution in making the survey a success.

**(3) Motivating the Interviewer/Data Collector:** For extracting desired results smoothly and conveniently, interviewers/data collectors should be provided appropriate training for a specific period along with some incentives and a time-bound promotion for their hard work. Similarly, the data collector/investigator engaged on an ad hoc basis may be assured of getting priority in future surveys/data collection work. To address the problem of reluctance among highly qualified and experienced youth to take the survey/data collection work, involving local educated persons such as a small shop owner or seasonal/part-time agricultural worker will also bring desired results as they are familiar with ground reality.

**(4) Deploying a Regular Interviewer/Data Collector Instead of Ad Hoc Appointments:** Filling up the vacant positions of supervisors/investigators instead of using ad hoc/temporary enumerators could be a better step towards addressing the problems of surveys/censuses. The job of data collection through surveys may not appear to be lucrative but once a data collector is insured of career progression, she/he would certainly take up the job sincerely. The scope for a data collector/investigator to rise to the level of a supervisor and above would certainly be a strong motivating factor.

**(5) Comprehensive Questionnaire/Schedule:** Intricacies involved in questionnaire/schedule and length also need to be looked into. Table 4 analyses the said issues:

**Table 4: Analysis of Time Taken in Canvassing the Schedule**

Descriptive Statistics	Time to Canvass Schedule (in minutes)		
	61st Round	66th Round	68th Round
Number of observations	1,24,536	1,00,869	1,01,510
Minimum	10	20	20
Maximum	300	260	210
Average	106.47	96.09	97.44
Std deviation	45.05	37.65	33.81

Source: Same as Table 1.

From the table it is observed that the average time taken in canvassing the schedule in all the said three rounds was about more than one and a half hours. Moreover, the deviation from the average was also found to be much wider. This could be taken care of by ensuring that utmost care is taken while preparing the questionnaire/schedule. It should also be ensured that the questions are framed in a comprehensive manner which could elicit more information through a short question. Further, the questions should have a logical flow so that the respondents could follow it easily and could answer promptly and conveniently. They should be so simple that even an illiterate or a neo-literate could easily follow the question and answer it conveniently.

**(6) Rigorous Implementation of the Collection of Statistics Act, 2008:** This act envisages the collection of statistics on economics, demographics, social, scientific and environmental aspects, and for matters connected therewith or incidental thereto. The Collection of Statistics Act, 2008 needs to be implemented in true letter and spirit. Once it is made binding on every respondent approached by the investigator/data collector, people would certainly come forward to provide the solicited information and thus the requisite data could be gathered without much difficulty. Adequate publicity measures should be taken for disseminating the details of this act as well as its implications for the common people. The respondents should also be appropriately assured of the confidentiality of the information provided by them.

**(7) Uniformity in Approach, Mechanism and Periodicity of Data Collection through a Central Agency:** It is observed that data from various sectors is collected through different agencies. There is no uniformity in the approach or mechanism adopted by each agency and periodicity of the data collection by these agencies also differs from each other. Hence, it is desirable have a uniform approach and mechanism and also have a specific periodicity of data collection. These could be ensured by assigning this task

to a central agency which could be the nodal guiding agency for the different agencies involved in the data collection and also for eliciting a desired response from the different surveys. Such a measure taken by the government would also address the problem of identical data being collected by different agencies from the same set of respondents causing respondents' fatigue, additional cost to the government and wastage of manpower.

**(8) Use of Latest Information and Communication Tools/Techniques:** To minimise human error in data feeding and processing (non-sampling errors), the use of latest information technology tools such as web portals, hand-held devices/palmtops should be introduced. Such measures would certainly be time saving as well as cost-effective. However, it should also be ensured that the hand-held devices are user-friendly and also the persons handling them are provided with the requisite training for using them.

**(9) Data Collector's Skill:** The data collector should apply her/his skill and intelligence to extract facts when a respondent intends to hide her/his real status. For instance, a well-off person could claim herself/himself under BPL category for illegal benefits or a BPL category person can claim a non-BPL category to avail privileges not entitled to her/him.

Besides the above-mentioned remedies, following actions also may be considered for ensuring effectiveness of a survey/census: (a) A data collector has to ensure that she/he arranges the interview schedule in such a way that the timing of interview/data collection does not obstruct the work of the respondent. For example, approaching a doctor during her/his hospital/clinic hours or a farmer during her/his fieldwork should be avoided. (b) To deal with the problem of a non-responsive respondent, the data collector should get the selection altered. She/he should instead try to replace this non-responsive respondent by one who gives clear responses. (c) For ensuring that data collection is as per the prescribed methodology and laid-down procedures, rigorous supervision of data

collection is necessary. Meticulous scrutiny of filled-in schedules is also necessary to ensure that no block is left blank and there are no intra- or inter-block inconsistencies. (d) For generating authenticity in the data collected through any survey/census or to build a confidence in the mind of its users, it is desirable that the data collected is cross-checked with any existing data on same account. Concurrent evaluation should preferably be conducted to find out that the data collection is going on as per the requisite format and desired pattern. (e) Many issues highlighting the relevance of the data/information come to the fore when it is put to use for planning, policymaking and development. At that stage, we also come to know about the intricacies, if any, involved in the process of data collection. That should be kept in view for taking corrective measures for ensuring successful conduct of future surveys.

### Conclusions

Based on the above elaborations, it is observed that, optimising the survey questionnaires/schedules is very necessary. For giving new directions to various development programmes and also to take up corrective measures, it is necessary that concrete information/suggestion is elicited through different tools used in a survey. As it is known, about Rs 5,55,000 crore is spent every year by the Government of India on plan schemes, most of such schemes need modifications with the changing dynamics of economy as well as social

mobility. However, if the desired quality of data is not generated through different surveys, the basic purpose of survey itself remains unserved. Therefore, it is desirable that the issues/intricacies involved in different surveys conducted in our country are addressed with a view to find amicable solutions. That will not only make surveys easier, but would also help in eliciting appropriate information with optimum cost within a prescribed time period.

To ensure the success of a survey, it is necessary that it is planned appropriately, the schedule is framed by taking into account the ground-level situation and the same is pre-tested before it is administered. For the collection of data of various kinds required by the government, there should be one nodal agency which could provide expert advice/guidance for planning and execution of a survey. Necessary steps should also be taken to make all the stakeholders fully aware about the survey. Steps should be taken to properly educate the respondent about her/his responsibility/obligation for providing the requisite information. There should be a proper monitoring of the survey work, ensuring that the procedure is being adhered to. A survey should not suffer for want of necessary funding. Hence, optimum allocation should be ensured to make the survey technically sound and thereby ensuring its timely completion. Adopting above ways and means would make a survey/census more productive and useful for appropriate development planning.

## Economic & Political WEEKLY

### EPW 5-Year CD-ROM 2004-08 on a Single Disk

The digital versions of *Economic and Political Weekly* for **2004, 2005, 2006, 2007** and **2008** are now available on a single disk. The **CD-ROM** contains the complete text of 261 issues published from 2004 to 2008 and comes equipped with a powerful search, tools to help organise research and utilities to make your browsing experience productive. The contents of the **CD-ROM** are organised as in the print edition, with articles laid out in individual sections in each issue.

With its easy-to-use features, the **CD-ROM** will be a convenient resource for social scientists, researchers and executives in government and non-government organisations, social and political activists, students, corporate and public sector executives and journalists.

Price for 5 year **CD-ROM** (in INDIA)

Individuals - Rs 1500

Institutions - Rs 2500

To order the **CD-ROM** send a bank draft payable at Mumbai in favour of *Economic and Political Weekly*.

Any queries please email: [circulation@epw.in](mailto:circulation@epw.in)

Circulation Manager,

**Economic and Political Weekly**

320-321, A to Z Industrial Estate, Ganpatrao Kadam Marg, Lower Parel, Mumbai 400 013, India