

## Multimode surveys from the perspective of total survey error

Gaël de Peretti, Tiaray Razafindranovona<sup>1</sup>

### Abstract

National statistical offices are subject to two requirements that are difficult to reconcile. On the one hand, they must provide increasingly precise information on specific subjects and hard-to-reach or minority populations, using innovative methods that make the measurement more objective or ensure its confidentiality, and so on. On the other hand, they must deal with budget restrictions in a context where households are increasingly difficult to contact. This twofold demand has an impact on survey quality in the broad sense, that is, not only in terms of precision, but also in terms of relevance, comparability, coherence, clarity and timeliness. Because the cost of Internet collection is low and a large proportion of the population has an Internet connection, statistical offices see this modern collection mode as a solution to their problems. Consequently, the development of Internet collection and, more generally, of multimode collection is supposedly the solution for maximizing survey quality, particularly in terms of total survey error, because it addresses the problems of coverage, sampling, non-response or measurement while respecting budget constraints. However, while Internet collection is an inexpensive mode, it presents serious methodological problems: coverage, self-selection or selection bias, non-response and non-response adjustment difficulties, 'satisficing,' and so on. As a result, before developing or generalizing the use of multimode collection, the National Institute of Statistics and Economic Studies (INSEE) launched a wide-ranging set of experiments to study the various methodological issues, and the initial results show that multimode collection is a source of both solutions and new methodological problems.

Keywords: Multimode survey, total survey error, government statistics

### 1. Introduction

#### 1.1 The French and European context

Over the last 20 years, national statistical offices (NSOs) have had to cope with requirements that are difficult to reconcile. On the one hand, they are being asked to provide statistics that are of high quality, in the broad sense of the term. For example, for European NSOs, which are supposed to apply the code of practice established by Eurostat, the statistical arm of the European Commission, quality relates not only to the precision of the data produced, but also to their relevance, timeliness, accessibility, comparability and coherence (Desrosières 2003). Moreover, in the case of survey statistics, NSOs have to work with concepts that are increasingly tricky to measure with questionnaires (e.g., literacy, numeracy, and well-being); survey hard-to-reach or minority populations (e.g., homeless people, people who have disabilities or are dependent, and immigrants or their descendants); and employ innovative methods to produce more objective measurements (e.g., GPS for transportation surveys, physical tests for health surveys, and biological samples) or to guarantee the confidentiality of responses and thus ensure truthful answers from respondents (e.g., use of a headset to collect responses to questions about domestic violence in a survey on victimization). On the other hand, NSOs are dealing with budget restrictions in terms of both financial and human resources, in a context of falling response rates, due in part to growing difficulty in reaching households, but also to excessive canvassing, particularly by players such as private firms conducting opinion, marketing and other polls.

---

1. Gaël de Peretti, INSEE, Département des méthodes statistiques, 18 boulevard Adolphe Pinard, timbre L120, 75675 Paris cedex 14, [gael.de-peretti@insee.fr](mailto:gael.de-peretti@insee.fr); Tiaray Razafindranovona, INSEE, Département des méthodes statistiques, 18 boulevard Adolphe Pinard, timbre L120, 75675 Paris cedex 14, [tiaray.razafindranovona@insee.fr](mailto:tiaray.razafindranovona@insee.fr).

To address these challenges, many NSOs have implemented or are considering implementing multimode surveys to control or reduce collection costs, while maintaining survey quality (Lyberg 2012). For example, Eurostat decided to launch a project on multimode collection for social surveys, consisting of two components:

- the development of Internet collection tools
- the consequences of implementing such tools.

This project is based on the Labour Force Survey (the employment survey in France) and is aimed at harmonizing and standardizing the survey's collection processes and sharing common practices. The goal is to mitigate cost and time constraints that are potential sources of quality degradation.

Similarly, INSEE has set up a plan to experiment with Internet collection for household surveys to determine the extent to which this new mode might help to resolve the above-mentioned difficulties (de Peretti and Razafindranovona 2013; Sautory et al. 2014).

## **1.2 Advantages and disadvantages of Internet collection**

The interest generated by Internet collection is obviously not unwarranted, and there is already ample literature on the ergonomics and design of questionnaires (Couper 2008). Indeed, Internet collection has many advantages:

- rapid collection that is more efficient than that of a self-administered paper questionnaire, because survey administrators can better manage filters and use embedded edits
- almost zero marginal cost for the collection of a questionnaire, once the e-questionnaire has been developed
- greater flexibility for the respondent in choosing when to complete the questionnaire
- varied content, such as images, sounds and videos, which may make questions or modules easier to understand and reduce the cognitive effort required of respondents.

However, this form of collection has shortcomings that suggest NSOs should exercise caution before they generalize it or use it systematically in household surveys without prior testing to ensure that the data collected are robust and reliable. Response rates are still low (see below for specific examples). In France, coverage is not perfect, since only 80% of the population has a home Internet connection. NSOs need to address technical problems, such as e-questionnaire compatibility with different browsers. E-questionnaires are still self-administered questionnaires, with the attendant measurement errors, such as greater 'satisficing' (Krosnick 1991).

## **2. A brief overview of Internet collection at INSEE**

### **2.1 Business surveys**

In France today, more than half of all businesses respond to official statistical surveys online, but the online response rate varies very widely from survey to survey, from 12% to 95% in 2013.

By and large, there are two collection approaches. Usually, INSEE sends an introductory letter to businesses containing a password, an identifier and the address of the survey website. However, businesses can obtain a paper questionnaire on request. This is the sequential multimode approach. In other cases, the multimode approach is concurrent: businesses receive both a paper questionnaire and information on how to respond online (website address, password and identifier). After three consecutive online responses for repeated surveys (for example, business outlook surveys), INSEE omits the paper questionnaire and mails out only the online response information.

To ensure some consistency in online business surveys, INSEE is currently developing a common platform that will be open to all parties involved in official business statistics programs.

### **2.2 Household surveys**

Internet collection for household surveys is less advanced. INSEE has been conducting annual census surveys since 2004. The initial experiments with Internet collection began in 2011. In 2015, INSEE will offer the online response option to all census respondents, about 9 million individuals, for the first time.

Few traditional surveys use the Internet as a collection mode integrated into the production process. Three surveys use it concurrently with another collection mode:

- the survey on entry into the labour force, which covers a cohort of children that started *sixième* (the first grade of secondary school in France) in 2007, uses the Internet as a concurrent mode in a simultaneous multimode collection approach that also includes telephone and face-to-face interviewing
- the employment survey (the European Union Labour Force Survey) uses the Internet concurrently with a self-administered paper questionnaire for the survey of non-respondents
- the survey of income and living conditions (European Union Statistics on Income and Living Conditions [EU-SILC]) uses the Internet concurrently with a self-administered paper questionnaire on well-being.

In addition, the telephone survey on information and communications technology (ICT) uses the Internet as a supplementary mode to correct a coverage gap in the telephone frame used for collection (see below for more details on the production process; see also Gombault and Duée 2012).

### **3. Internet collection experiments for household surveys**

#### **3.1 Can multimode collection help reduce total survey error?**

While the term ‘total survey error’ was coined in the late 1970s (Andersen et al. 1979), the fact that there are many sources of survey error has been known for much longer, at least since the early 20th century (Groves and Lyberg 2010). The components of total survey error are generally well known and include sampling error, non-response error, coverage error, measurement or observation error (including social desirability and ‘satisficing’), specification error, capture error, coding error, and imputation error.

Not all these sources of error are necessarily quantifiable in the same way as sampling and non-response errors, which can be approximated or corrected with statistical processes. However, the value of this conceptual framework is that it encourages statisticians designing a survey to consider all possible sources of error and devise solutions that minimize total survey error within the limits of the budget.

Thus, the Internet might be an alternative collection mode for limiting non-response, improving coverage rates and reducing measurement error for sensitive questions (less social desirability or greater confidentiality), and it might be a quick, inexpensive means of running a pilot survey to explore a new subject. With this in mind, INSEE developed an experiment plan.

#### **3.2 Principles and protocol of the experiment plan**

Because of the difficulty of producing general results for multimode surveys (Couper 2011), INSEE decided to carry out experiments in parallel with telephone or face-to-face surveys. These experiments must not have any impact on the core survey, particularly in terms of timeframe and workloads, but also in terms of sampling; therefore, we used disjoint samples. In addition, since this is a process of knowledge accumulation, the plan is to test new questions in each experiment.

We developed the following standard protocol for these experiments:

- sampling: the sample is selected from residence tax files, which contain the addresses of individuals, so that we can carry out mass mailings
- questionnaire: we used a selection of questions from the core survey, with completion time averaging 15 to 20 minutes and not exceeding 25 minutes
- contact and follow-up: we sent an introductory letter to respondents containing the online response information, and we had the option of sending two postal reminders three weeks apart, with a paper questionnaire enclosed with the first reminder.

#### **3.3 Results of the initial experiments**

### **3.3.1 The 2010 housing and residential mobility survey**

This was INSEE's first online household survey experiment (except for the use of paper and online multimode collection for non-respondents of the survey of employment starting in 2008, and for an ICT methodology survey in 2010). The experiment had three main objectives: test the administration of an online survey, measure respondent sensitivity to questionnaire length (a 15-minute version compared with a 25-minute version), and measure the mode effects on questions about home comfort by comparing with the face-to-face housing survey of 2006 (Amiel and Denoyelle 2012).

The key results of this survey are as follows:

- The response rate is low: 19% (24% including respondents who used the paper questionnaire). The protocol for this experiment was different from the standard protocol in that the selected individuals received a reply form with the first reminder that they could mail in to request a paper questionnaire, but there was no prepaid envelope.
- Well-educated and well-off people are overrepresented.
- Answers to the income questions were of poor quality (possible comparison with tax data).
- More people were dissatisfied with their housing, even after we controlled for conventional socio-demographic characteristics (age, sex, income, education, employment status, number of occupants, and urban unit size). However, it was impossible to determine whether the difference was a selection effect associated with unobservable variables (the people who chose to respond have a greater propensity to live in poor housing conditions) or the opposite, a nullification of social desirability (it is more difficult to complain about one's housing to an interviewer).

### **3.3.2 The 2011 survey on quality of life**

In 2011, following the report of the Commission on the Measurement of Economic Performance and Social Progress, known as the Stiglitz-Sen-Fitoussi report, INSEE decided to test some indicators of quality of life using an online survey. Some of the questions in this experiment are already included in the EU-SILC, a face-to-face survey that we will use as a benchmark.

The key results of this survey are as follows (Amiel et al. 2012):

- The overall response rate (online and paper) is higher than in the first experiment, but this is largely because it was easier to respond on paper (we mailed out the questionnaire). The response rate is 38% (46% of which are online responses, for an online response rate of 16%, down slightly from the previous experiment).
- The results for overall life satisfaction are similar to the results of the EU-SILC.
- The relationships between the various dimensions of well-being (housing, financial constraint, physical safety, economic security, health, emotional well-being, environmental conditions, social connections, confidence in society) and overall satisfaction are convergent between the two surveys.
- The overall satisfaction of postsecondary graduates tends to be lower in the EU-SILC and higher in the survey on quality of life.

### **3.3.3 The new protocol of the information and communications technology survey**

Since 2007, INSEE has conducted an annual telephone survey on information and communications technology (ICT). The telephone survey was based on a sample selected from an incomplete telephone-number frame covering about 50% of the population. Initially, INSEE used methodology surveys to identify coverage biases and devise a new protocol to correct them. Ultimately, INSEE selected a telephone, online and paper multimode collection approach, and made specific choices for sampling and adjustment (or calibration), as follows:

- first stage of sampling: selection of 40,000 individuals from tax files (including 19,000 listed in the telephone directory)
- second stage of sampling: selection of 7,000 individuals from the 19,000 for the telephone survey (with possible online or paper reminder in the event of non-response), referred to as the telephone sample; selection of 15,000 individuals from the 21,000 not listed in the telephone directory for a sequential online

and paper survey; and selection of 5,000 individuals from the 19,000 for a sequential online and paper survey, referred to as the methodology sample

- comparison of the telephone and methodology samples: higher rates of Internet equipment ownership and Internet use, even after calibration on socio-demographic variables, in the methodology sample
- taking into account the results of the various collection modes, by using finer calibration margins on the socio-demographic characteristics, but also taking into account the differences in the variables of interest between the telephone and methodology samples.

It is also important to bear in mind that the Internet collection rates remain low, far below the telephone collection rates, as follows:

- The online and paper response rate for the 2013 ICT survey is 34% (compared with 38% in 2012), with 18% online (compared with 17% in 2012).
- The telephone response rate is 64% (compared with 73% in 2012) and 67% with online and paper reminders (compared with 79% in 2012).

### **3.3.4 The 2013 survey on quality of life at work**

This survey, the online counterpart of the 2012 survey on working conditions, had several objectives (Razafindranovona, de Peretti and Barrau 2013):

- test a pure Internet protocol with no paper response option
- test sensitivity to the introductory letter with two versions, one focusing on working conditions and the other on psychosocial risks
- compare responses with those of the face-to-face survey
- test new questions on psychosocial risks, both for wording and by introducing open questions that will be used to prepare closed questions for the 2015 survey of psychosocial risks.

In view of the subject matter, the target population was employed people aged 18 to 65. This accounts *a priori* for the higher response rates, as follows:

- The response rate for the pure Internet protocol is 44%; for the Internet and paper protocol, it is 52% (with 32% online).
- Executives, members of intermediate occupations and high-income earners are overrepresented among online respondents; employees, women and low-income earners are overrepresented among paper respondents.
- The response rates are equivalent for the two types of introductory letter.

Testing the questions on psychosocial risks enabled us to introduce closed questions in the 2015 survey of psychosocial risks after studying the results of the open questions, and to add response choices to some closed questions through the use of “Specify” when respondents select the “Other” response choice. In addition, we made preliminary comparisons of online and paper collection. After applying controls using matching-type techniques (Heckman, Ichimura and Todd 1998), we observe a smaller ‘satisficing’ effect in response-scale questions and lower well-being (as defined by the World Health Organization) in online collection than in paper collection.

### **3.3.5 The survey of robbery, violence and security**

The main purpose of this survey was to compare its results with the results of the victimization census in the face-to-face survey of the living environment and security (SLES) (Razafindranovona et al. 2014). Another objective was to study the self-selection of individuals using the following protocol:

- We gave two individuals the opportunity to respond (from here on, the individual to be surveyed, i.e., selected by the sampling procedure, will be called the Kish individual, since he or she is chosen by the Kish method).
- We included a household questionnaire at the beginning of the survey.
- We selected a subsample of teenagers (ages 14 to 19) whose identity is unknown in the frame.

Of the 40,000 introductory letters we mailed out, 4,500 did not reach the intended recipient (returned as “Not at this address”). Of the 35,500 people who received the letter, only 14,500 responded (53% on paper and 47% online):

13,000 questionnaires were usable, and 1,500 were screened out, either because the respondent did not complete enough of the questionnaire or because the respondent was not the Kish individual. This yields a response rate of 32% (37% if we exclude the “Not at this address” returns).

Since the SLES (administered annually since 2004) has produced stable results over time, has a high response rate of about 75% and is consistent with administrative data on complaints laid with police services, it was selected as the ‘gold standard.’

As usual, online respondents are younger (a median age of 44, compared with 48 in the initial sample), wealthier (a median annual household income of €36,500, compared with €27,700) and better educated (45% postsecondary graduates, compared with 24%). We also observe that respondents to the self-administered questionnaire (online and paper) are more likely to report being victims than are respondents to the SLES, despite numerous adjustments (or calibrations) taking into account the variables of interest of the survey (see Table 3.3.5-1). We attempted four successive calibrations (all of them taking non-response adjustment into account): calibration on socio-demographic margins (CSM), CSM and complaint rates (CSM + CR), CSM + CR and the feeling of insecurity (CSM + CR + FI), and CSM + CR + FI and Internet access.

**Table 3.3.5-1**

**Prevalence of victimization in the survey of robbery, violence and security and the survey of the living environment and security**

%	Robbery with violence	Robbery without violence	Physical violence	Threats
After non-response adjustment	2.6	6.4	2.7	7.3
CSM	2.9	6.4	2.9	7.7
CSM + CR	2.8	5.9	2.7	7.2
CSM + CR + FI	1.9	4.4	1.8	4.9
Total calibration	1.8	4.2	1.7	4.8
<b>SLES benchmark</b>	<b>1.0</b>	<b>2.8</b>	<b>2.2</b>	<b>4.5</b>

**Note:** CSM stands for calibration on socio-demographic margins. CR stands for complaint rates. FI stands for the feeling of insecurity.

**Sources:** 2013 survey of robbery, violence and security (SRVS), 2013 survey of the living environment and security (SLES).

**Scope:** Individuals aged 14 and over living in an ordinary dwelling in metropolitan France.

**Interpretation:** After we calibrate on socio-demographic margins and the complaint rate, 5.9% of individuals report having been victims of robbery without violence in the SRVS, compared with 2.8% in the SLES.

Introducing variables of interest to try to control for selection effects is obviously not always possible. For example, we observe that complaints to police services are underestimated in the survey of robbery, violence and security (SRVS). This might suggest that the individuals selected for the SRVS were more likely to respond if they were victims who did not file a complaint. The adjustment associated with the variable for the feeling of insecurity in the neighbourhood is the most influential adjustment. It is important to note that there is a very large discrepancy between the rate in the SLES, 12% (fairly stable since 2004), and the rate in the SRVS, 30%. Overall, however, the SRVS produces higher estimates for the prevalence of victimization.

Since the SLES serves as a benchmark, the hypothesis that there is an uncontrolled-for selection effect associated with unobservable variables appears to be valid. To test this hypothesis, we compared the victimization results for those who responded online after receiving the introductory letter and for those who responded online after the first reminder. The hypothesis to be tested is as follows: people motivated by the subject because they have been victims are more likely to respond to the survey. Comparison of the victimization rates for those who responded before the reminder and for those who responded after the reminder validates this hypothesis, since the rates for the former are invariably higher than the rates for the latter (see Table 3.3.5-2). To verify the robustness of this result, we tested logistic models for each type of victimization with socio-demographic variables (sex, age, family status, education, employment status, and type of urban unit) and a response-date variable (before or after reminder). Once again, the hypothesis appears to be validated, since we observe a negative effect on reporting victimization for people who responded after the reminder.

**Table 3.3.5-2**  
**Controlling for the respondent motivation effect**

	Robbery with violence	Robbery without violence	Physical violence	Threats
<b>Before</b> reminder (%)	2.9	8.2	2.8	10.3
<b>After</b> reminder (%)	2.4	6.4	2.7	7.2
Odds ratio After reminder compared with before reminder	0.80	0.78	0.92	0.65

**Source:** 2013 survey of robbery, violence and security (SRVS).

**Scope:** Online respondents aged 14 and over living in an ordinary dwelling in metropolitan France.

**Interpretation:** Of those who responded online to the SRVS, 10.3% of those who responded before the reminder reported being victims of threats, compared with 7.2% of those who responded after the reminder.

## 4. Conclusion

The transition to multimode collection and, in particular, the introduction of Internet collection appear to be necessary adaptations to technological change. However, Internet collection cannot be introduced without a careful assessment of the effects on survey quality, which often has to be done on a case-by-case basis because of the difficulty of generalizing results from any particular survey. Consequently, INSEE implemented an experiment plan to progressively study the possible effects of introducing Internet collection in household surveys. In addition to the experiments described, INSEE has a number of ongoing or forthcoming projects:

- testing the capability to collect data about rent and occupancy expenses for a housing survey without pressure from field interviewers to check documents
- testing the capability to collect financial data for an assets survey
- testing the capability to collect data on several household members with a long, technical questionnaire in a short collection timeframe for an employment survey
- studying the existence of a coverage bias associated with the telephone survey frame for a household economic conditions survey (and for the ICT survey).

Clearly, the experiment plan to help us reconfigure our surveys is a medium-term or even a long-term investment. In addition, ergonomic and contact issues will also require investment, in particular through the field of cognitive sciences, to improve response rates and response quality. In the near term, the Internet appears to be more of a complementary collection mode than a principal mode for household surveys.

## References

Amiel, M.-H., and T. Denoyelle. 2012. "Enquêtes en ligne : comparaison de modes de questionnement sur le thème du logement." In *Journées de Méthodologie Statistique*. Paris: INSEE.

Amiel, M.-H., P. Godefroy, and S. Lollivier. 2012. "Les personnes modestes en milieu urbain sont celles qui cumulent le plus de difficultés en matière de qualité de vie." In *France Portrait Social Insee Références*, p. 89–105.

Andersen R., J. Kasper, M.R. Frankel, and associates. 1979. *Total Survey Error*. San Francisco: Jossey-Bass Publisher.

Couper, M. 2008. *Designing Effective Web Surveys*. Cambridge: Cambridge University Press.

Couper, M. 2011. "The future of modes of data collection." *Public Opinion Quarterly* 75 (5): 889–908.

Desrosières, A. 2003. "La qualité de quantités." *Courrier des statistiques* 105–106: 51–63.

Gombault, V., and M. Duée. 2012. *Un exemple d'enquête multimode à l'Insee : l'enquête TIC*. Paper presented at the 7th francophone conference on surveys of the Société française de statistique, Rennes.

Groves, R., and L. Lyberg. 2010. "Total survey error: Past, present and future." *Public Opinion Quarterly* 74 (5): 849–879.

Heckman, J., H. Ichimura, and P. Todd. 1998. "Matching as an econometric evaluation estimator." *Review of Economic Studies* 65 (2): 261–294.

Krosnick, J. A. 1991. "Response strategies for coping with the cognitive demands of attitude measure in surveys." *Applied Cognitive Psychology* 5 (3): 213–236.

Lyberg, L. 2012. "Survey quality." *Survey Methodology* 38 (2): 107–130.

de Peretti, G., and T. Razafindranovona. 2013. *Les enquêtes multimode : multi-problème ou multi-solution ?* Paper presented at the 45ème Journées de Statistique de la SFdS, Toulouse.

Razafindranovona, T., B. Dietsch, C. Burricand, and G. de Peretti. 2014. *Le multimode pour mesurer la victimation : est-on dans la zone de sécurité ?* Paper presented at the 8th francophone conference on surveys of the Société française de statistique, Dijon.

Razafindranovona, T., G. de Peretti, and A. Barrau. 2013. *The Philosophy of French Experiments on Internet and Mixed-mode Data Collection*. Paper presented at the UNECE Seminar on Statistical Data Collection, Geneva, Switzerland.

Sautory, O., G. de Peretti, and T. Razafindranovona. 2014. *Experiments on Internet and Mixed-mode Data Collection at Insee*. Paper presented at the Fenstats and Eurostat Conference of European Statistics Stakeholders, Rome.