



Catalogue no. 11-522-XIE

**Statistics Canada International Symposium
Series - Proceedings**

**Symposium 2004: Innovative
Methods for Surveying
Difficult-to-reach Populations**

2004



THE EXPERIENCE OF SENSITIVE SURVEYS IN FRANCE

Benoît Riandey and Marie-Ange Schiltz¹

SUMMARY

Sensitive surveys require considerable know-how both in sampling design and in interviewing and collection. The French regulatory framework has added many constraints to this already difficult undertaking, but since the summer of 2004, the introduction of European regulations has loosened the constraints imposed on statistics in France. Examples of surveys of diverse populations will show how survey technique is inseparable from treatment of sensitive questions. These examples will demonstrate that there is no good survey technique without considering the ethics.

KEYWORDS: France; Legislation; Public Statistics Ethics; Sensitive Surveys.

1. INTRODUCTION

When comparing the treatment of sensitive subjects in different countries, variations in the list of subjects may be more surprising than variations in the techniques used to deal with them. Many survey topics explored in Canada cannot be handled in France or have been just recently. French history and legislation contribute to this situation. Because of conflicts between the transparent work of the statistical society and the respect for individuals, statisticians often find themselves in a contradictory position. Both technical competence and sensitivity to ethical considerations are needed to collect more reliable data with the consent of the survey subjects, but French laws impose procedures that dissuade statistical researchers from moving forward in some areas. Quantitative religious sociology has disappeared from French research, at a time when French society is concerned with the integration of immigrant populations. The lack of investment by French institutes in self-administered, computerized data collection is a technical limitation. Nevertheless, other approaches have been explored such as newspaper surveys or snowball surveys, and their results will be compared to those of more traditional surveys. Fortunately, researchers have seized opportunities for random and anonymous data collection on drug use, violence against women, the health status of inmates or the living conditions of persons with AIDS. With the use of identifier compression techniques, it has become possible to match survey data or administrative data and process them longitudinally, which was formerly strongly prohibited.

2. THE FRENCH CONTEXT

Discussion on sensitive surveys usually focuses on content. What is most interesting about this symposium is that it dissects the methodological or technical investment and the ethical concerns that these surveys entail. In the French context, the especially restrictive regulatory framework of these surveys cannot be ignored. Consider, for example, the survey on religious beliefs in the 25-nation Europe (Tinq, 2004). Its authors reported that they left out France to avoid inextricable negotiations with the controlling authority, the CNIL (Commission Nationale de l'Informatique et des Libertés). This legal aspect thus begins our exploration of the topic before discussing our experience with methodological and ethical concerns.

¹ Benoît Riandey, INED, 133 Boulevard Davout F-75020 Paris, France (riandey@ined.fr) ; Marie-Ange Schiltz, CNRS-CAMS, 54 Bd Raspail F-75006 Paris, France (schiltz@ehess.fr).

2.1 A Restrictive Framework (1978-1995)

The institutional framework for public statistical surveys, established by the 1951 public statistics law, was profoundly affected by the law on information technology and freedoms, passed in 1978. This law grew out of a reaction of fear with regard to the State, which developed when personal identification records were computerized. This highly innovative law was undeniably a legal advance for the protection of individual freedoms. It exerted a profound influence on the 1981 convention of the Council of Europe and, later, on the European directive of October 24, 1995. However, the European authorities had the wisdom to implement its principles much more flexibly. In France, by contrast, the initiation of the original law resulted in a jurisprudence that was sometimes questionable, especially for statistics and research: the law of 1978 does not recognize the purpose and specificity of these sectors, nor does it recognize the guarantee of statistical confidentiality, even though that guarantee had never been found deficient since the adoption of the legislation in 1951. All the new protections introduced regarding the collection of sensitive information make it difficult to conduct statistical surveys in important fields in the social sciences.

Article 31 of the law on information technology and freedoms prohibits the collection, without *express agreement*, of personal information relating to ethnic origins, religion, political or union affiliations, or, starting in 1981, sexuality or health. This article means that to collect personal information on these topics, the interviewer must give the survey subject a double message:

- “Yes, you are protected by statistical confidentiality, and you can answer without risk;
- but sign to indicate that you accept the risk of giving us such sensitive information.”

This procedure is understandable only in the health field, where the public is accustomed to the precautions taken to protect medical confidentiality. Elsewhere, it is completely intolerable. Thus we must conclude that it is impossible to conduct rigorous probabilistic surveys in the sociology of religion, in political science and in studies of immigrant populations.

Similarly, criminologists are still prohibited from asking questions about offences and sentences. In epidemiology, cancer registries remained illegal until the passage of the bioethics law in 1994 and, indeed, so were surveys on tax revenues conducted by INSEE (Institut national de la statistique et des études économiques) from 1978 to 1986. It was then that an amendment to the law of 1951 re-established the right of INSEE and statistical units within government ministries to receive personal information for exclusively statistical purposes. But, access to survey frames has remained difficult for private sector statisticians and research institutes, which were not covered by this law.

Administrative statistics have also been subject to very severe limitations, with the prohibition against collecting “excessive” information for administrative purposes: nationality does not appear in health insurance files, since this piece of information is not useful but is instead “excessive” for purposes of reimbursing expenditures on health care. In fact, public statistical bodies are supposed to evaluate administrative discrimination (such as in access to care), but to protect individuals, they are often prohibited from collecting data on the very characteristics against which the discrimination exists.

For the same reason, ethnic minorities are not entitled to count their numbers. Thus the Armenians in France have not been able to count the number of surnames ending in “ian” in the telephone directory. This situation reflects the weight of French history and some less-than-glorious moments of the French government such as dealing with the Jewish community at the time of the Dreyfus Case in the nineteenth century (then in collaboration with Nazism), or more recently in the abandonment of the “harkis” (Algerian soldiers loyal to the French) after the Algerian War.

2.2 An Evolving Situation (1995-2004)

The creation of a single European market has necessitated the circulation of personal data and forced an agreement between free trade countries (the United Kingdom, the Netherlands) and countries more sensitive (tense because of their history) regarding the protection of personal data (Germany, Austria and France). The European directive that

was passed on October 25, 1995 set out conditions for circulating European personal data, including provisions for circulating such data outside the national territory. In 2001 the directive was becoming an international standard when the events of September 11, 2001 imposed a more security-driven order on the world. Nevertheless, this law was first imposed on 15 countries and then on the 25 countries of the European Union including France... but only after the vote of the French parliament on August 6, 2004.

For the French, the first benefit of this was a simplification of the procedures hanging over the collection or use of personal data. As a result of the patient efforts of the Société française de la statistique and the Conseil national de l'information statistique (CNIS, 2000), the right to reuse personal data for statistical or research purposes was recognized. In terms of the nuanced language of the lawmakers, "*The reuse of personal data for statistical or research purposes is not presumed to be incompatible with the purpose for which they were collected.*" In simpler terms, private statistical institutes and research institutes, like public statistical bodies, are now entitled to access personal data from other institutions, notably for use in sampling frames.²

What about sensitive surveys? The European directive provides for the *derogation from the express agreement requirement on grounds of important public interest*, but the French government has always opposed this provision, which it saw as reducing freedoms. The Société française de statistique had been obliged to acknowledge its failure to change this position. Imagine its surprise to discover, in the fall of 2004, that the parliament had passed this derogation, seen as indispensable to monitoring the funding of health insurance. INSEE and ministerial statistical units will henceforth be exempted—by the CNIL—from the requirement of this express agreement on sensitive data.

2.3 Public Statistics and Sensitive Surveys

On that ground, are we going to see a proliferation of surveys that were lacking in the past?³

There are still limitations: the derogation applies only to INSEE and ministerial units; private polling institutes can therefore benefit from it only as sub-contractors of ministerial statistical units, and research institutes only by sub-contracting data collection from public statistical bodies. As well, the derogation does not allow for the European survey of religions, which is foreign to this framework.

Traditionally, INSEE has always avoided doing opinion surveys⁴ or surveys of political or religious behaviour. Its cautious attitude probably has to do with its special place among the National Statistical Institute of Macro-Economic Studies, and with the tension that arose in the past over the computerization of the registry of individuals. French statisticians are all the more admiring of Statistics Canada's involvement in its surveys of violence against women, surveys of aboriginal peoples, and surveys of persons with disabilities. In France, the initiative for these subjects comes from research institutes such as INED (Institut national d'études démographiques) or, more rarely, from the universities.⁵ However, there has been some movement: the HID Survey of Disabilities, which for a very long time was unthinkable in France, has been conducted by INSEE since 1999 (Ravaud et al., 2002); then INSEE

² To be more precise, consider the example of records on electrical meter readings by Electricité de France (EDF): private polling institutes would have always had access to these records for surveys conducted for the EDF; since 1986, the EDF is entitled to transmit this file to INSEE and ministerial statistical units for purposes of statistics (on residential mobility), surveys (on rents) or the updating of its survey frame; since August 6, 2004, and subject to the agreement with the CNIL, the EDF would be entitled to transmit this file to survey institutes or INED for the same purposes.

³ It was possible to maintain political polls on quotas owing to an agreement with the CNIL on a simplified consenting arrangement. Thus, INED was also able, on the bases of the census and in collaboration with INSEE, to conduct the MGIS survey on the integration of immigrants (Tribalat, 1995).

⁴ Other than surveys of opinions on economic conditions; at most, it has collected information on voter registration and electoral participation as well as on religious observance, but without specifying the religion.

⁵ At INSEE's request, the World Fertility Survey was conducted in France by INED (Leridon, 1979). The ACSF survey (Spira, 1993) on sexual behaviour and the ENVEFF survey (Schiltz, 2001) on violence against women were conducted by telephone, under the responsibility of INSERM and the University of Paris respectively, without the interviewers handling the calls. The CNIL recognized the anonymity of this procedure and considered that these random and anonymous surveys did not come within the purview of the law. French private polling institutes have never used this method to conduct random political surveys.

conducted a national survey of homeless persons (Brousse, 2002) on the model of the regional survey conducted by INED (Marpsat, Firdion, 2000). The Family History Survey (Cassan, 2001), associated with the 1999 census, includes questions on common-law unions and adoptions and also on mother tongue, an indicator of “ethnicity.” DREES, the new statistical and research unit of the Ministry of Social Affairs, conducts surveys on populations at risk and such sensitive topics as violence. Despite this trend, sensitive surveys tend to be designed by a research institute, although they may then be implemented in collaboration with INSEE.

3. EXPERIENCES

When conducting surveys of populations that are difficult to reach because of their minority nature, their marginality, their living conditions or the opprobrium associated with them, it is necessary to define the criteria for inclusion in or access to this population and not to neglect covering the shadow, namely the “hidden population”.

3.1 Defining the Population

Determining the scope of the population depends on the choice of criteria for inclusion in it. These criteria are often neither clear nor visible. We will consider two examples: homosexuality and physical disability.

Homosexuality may be defined on the basis of various factors: one’s fantasies or tendencies, an attraction toward persons of the same sex or persons with minority sexual behaviours, self-identification as a homosexual, a “gay” lifestyle, or identification with the social group. None of these definitions solves the problem of determining the scope of the “target” population because of the many ways of characterizing who is “homosexual” and the vagueness surrounding the concept of sexual relations. For example, voluntary surveys define as “homosexual” those who include themselves in a survey with the explicit title “Homosexuals and AIDS” (Pollak, Schiltz, 1991). By contrast, the classification procedure followed in the ACSF survey on sexual behaviour in France is based on random recruitment within the general population: researchers devised in advance an unequivocal and behaviouristic definition of homosexuality, and they classified as homosexual those individuals who answered “yes” to the question “have had at least one sexual experience with a person of the same sex during your life.” According to this survey, 4.1% of males and 2.6% of females sexually active during their life reported that they had had at least one same-sex partner (Spira, Bajos, 1993).

Another choice was made in the INSEE Survey of Handicaps, Disabilities and Dependency. Because there are so many ways of identifying the disabled population, a resolutely, multi-dimensional definition of the target population was adopted. It reflects five possible approaches to disability, each authorizing a different interpretation of the disabled population. These groups are ranked in descending order of prevalence⁶: 1) functional disabilities identified by seven factual questions without explicit reference to the designations “handicap,” “disability” or “dependency” (21.1%); 2) limitation on the type or amount of activity, as determined by a question on limitations lasting six months or more (11.7%); 3) self-reporting of a handicap (9.5%); 4) use of some form of support, whether it be human assistance, a prosthesis, a device, technical support, adaptation of the dwelling (7.7%); 5) the seeking of official social recognition of the handicap (6.2%). These five groups are of different sizes and do not completely cover the target population (Ravaud *et al.*, 2002).

3.2 Surveying the population

In practice, difficult-to-reach populations can be accessed in a variety of ways. For this purpose, it is possible to use

- a social designation, by recruiting persons receiving some form of social assistance;
- a relational identity; in this case, inclusion takes place gradually, through contacts maintained by the target group’s members;
- the identity asserted as a result of an individual’s decision to take part in a study;

⁶ The figures used here concern the population aged 16 and over.

- a classification decided in advance by the researcher to include individuals who have reported one or more behaviours.

In the case of the Vespa Survey of Persons with HIV (Obadia, Lert, 2004), the objective is to establish a random sample of persons who know themselves to have been infected with HIV for at least six months. The survey is based on the medical services provided to this population. In France, it is estimated that 93% of persons with HIV have received some form of in-patient or out-patient hospital care. In the absence of a sampling frame of HIV-positive individuals, researchers prepared a complete list of the hospital units in charge of these patients in French territory, and representativeness was obtained by drawing these departments at random. Similarly, for the study of homeless populations, the French survey accessed the target group through records of social housing benefits and free meals (Marpsat, Firdion, 2000).

Different approaches to homosexuals in France illustrate the other possibilities for reaching a population for which no sample frame exists. The "Gay Report" (1983) combined snowball recruitment and self-inclusion using questionnaires disseminated at gathering places and associations. The result was a sample of men identifying themselves as "homosexual." These were primarily urban males from the Paris region, the most visible and the most active, who fell within a relatively limited age range (25 to 40 years of age).

From 1985 to 2004, various "gay newspaper" surveys have been conducted, based on a questionnaire explicitly targeting homosexuals through the press. These surveys sought to determine gay lifestyles and homosexuals' reactions to AIDS (Pollak, Schiltz, 1991). Recently this approach has been supplemented by disseminating the survey via Web sites. With this approach, a large number of volunteers can be recruited rapidly and inexpensively. Owing to the wide circulation of the media selected, the sample draws recruits from well beyond the Parisian scene: homosexuals and bisexuals, self-identified to be sure, at the provincial level including younger as well as older. With this easy-to-replicate operation, it is possible to track changes and the differentiated adaptations of sub-groups (younger compared to older; the most disadvantaged persons compared to workers or executives). In 1986 and 1987, two surveys of non-readers supplemented the observation focusing on readers: their recruitment, using the snowball method, took place in gathering places in Paris and in the provinces according to quotas defined on the basis of the characteristics of unmarried adult males in the census. With this procedure, it was possible to recruit more homosexuals from disadvantaged social strata. On the other hand, this procedure favours the response of homosexuals who identify the most with the "community" by displaying their sexual preference in gathering places and agreeing to be interviewed face-to-face regarding their sexuality.

A comparison of the characteristics of these differently constituted populations tells us that snowball recruitment, even when accompanied by strict guidelines on the quotas to be met, is not always perfect. Indeed, in subcultures, the least marginalized groups are also the least visible. A procedure designed to reach the members of a group through a relational network is therefore likely to reach only the most visible and self-affirmed part of the target population. This selection bias is reinforced by the interaction with the interviewer; the lack of anonymity in the face-to-face encounter may drive away persons in denial, persons who are not "out" or those who normally refuse to acknowledge their stigmatized behaviour to an "outsider." Unless exhaustive techniques are used, this approach provides no representativeness. In addition to being time-consuming, it brings out "scenarios" rather than general patterns.

As to voluntary surveys, they are possible only if there is an organized community and if the sense of a shared destiny is expressed through associations or specialized media in which the persons concerned see themselves reflected. It should be kept in mind that entry into a deviant or stigmatized group is often the last stage on a marginal course. Thus, the population of those who include themselves voluntarily accordingly to a negative attribute is not a microcosm of a larger grouping. On the contrary, it is very likely to give a more hardened image of the group, and the more marginalized the group, the more hardened the image. The narrower the definition that is used, the more specific will be the channels for disseminating the survey and the more distinctive the characteristics of the spontaneous sample will be in relation to those of the reference population.

Another approach to homosexuality is the one used by the ACSF survey on sexual behaviour in France. This large-scale probabilistic survey, focusing on the general population and conducted by telephone, operates in two stages. The first step is to identify, among the 20,000 persons contacted, those individuals whose behaviours expose them

more than others to the risk of HIV, and then they are questioned in detail. Persons are selected on the basis of a neutral filter question that is “random” in the sense that, in the midst of the risk behaviours,⁷ it includes an item about being born on the 7th, 17th or 20th day of a month. Thus a “yes” to the overall question does not necessarily refer to stigmatized behaviours. The purpose of including this random element is to recruit a control sample and make the filter question more neutral. This procedure, the most unassailable from a statistical standpoint, recruits far more than just the ranks of self-identified homosexuals and gives the least specific image of the set of men who have engaged in a homosexual activity during their life. But compared to the surveys described above, it is also the most costly and the one that yields the lowest number of homosexual or bisexual males (n=121), and, for the women concerned, a number that is nearly too small for statistical processing (n=41) (Messiah, Mouret-Fourne, 1993).

For these multi-stage surveys, the success of the filtering is key. The strategy of including a question neutralized by a random element is not always sufficient. Thus, in the ACSF survey, this filtering, which occurred relatively early in the interview process, was not enough to prevent concealment of practices that are not always easy to acknowledge or talk about, even by telephone. As proof, a more than 30% larger control sample suggests probable denial of homosexual practices during the interview, even though the word “homosexual” was never used (Riandey, Firdion, 1993).

From past experiences, a few principles can be retained. Filtering, which always has the potential to definitively exclude false negatives, must be introduced as late as possible in the interview. The idea is, firstly, to minimize the opportunities for denial, such as by proposing to the respondent a set of response categories that contain a random-type answer mixed with others that are worded in a nuanced and inclusive manner: “partner of the same sex as you” or “consumption of soft or hard drugs”; if the question is too definite, it may give the respondent an excuse not to answer. This is what appears from a comparison of the results of two probabilistic surveys conducted almost simultaneously in the general population to estimate various types of violence, especially toward women. These two surveys, very similar in their purpose and close in time, provide quite dissimilar estimates, often by a ratio of 1 to 4. Large disparities are observed for all indicators that lend themselves to comparison. These differences are largely attributable to the construction of the interview and the filtering strategy: the filtering in the survey with the lowest estimates is based on a single question on the existence of an assault. The second survey is organized around situating the respondent in the different spheres of her life: public, professional, family and conjugal. For each of these spheres, a list of facts is repeated without ever using the words “assault” or “violence” (Schiltz et al., 2001). This strategy increases the angles of observation without filtering. It serves to minimize the negative rankings observed in many other surveys and thus to obtain prevalences that are higher and probably closer to reality.

Whatever approach is followed, one finding keeps coming up: the existence of a “hidden population” that could not be reached because of the choice of the criterion for inclusion, the procedure for accessing the population, the survey instrument chosen and the effect of filtering. In all cases, the research team must question itself about the “grey” areas. However sophisticated the procedure chosen in a survey of social assistance recipients, consideration of the hidden population is essential. This is fundamental to the discussion on the limitations of the results produced. Individuals systematically excluded from this social assistance are precisely those who, although entitled to it, stay away because they feel ashamed about seeking help or because they wish to assert their freedom to live on the margins of society, or those who, being too marginalized, do not meet the administrative criteria (undocumented, no address, refusal to participate in reintegration or detox programs) or can no longer go through the procedures (because they are illiterate, too alcoholic, etc.).

In the Vespa survey, an investigation of persons aware of their HIV-positive status but not receiving hospital care showed the diversity of situations of “out of scope” persons and the impossibility of reaching them through the standardized interview system: senior executives who choose a physician in town to ensure discretion and take advantage of clinic hours convenient for them; individuals who reject the idea of treatment or prefer alternative medicine; etc. Along similar lines, the collection of data on homeless persons via shelters and food distribution centres was supplemented in 2002 by a study that sought to learn more about the more independent homeless persons, those who did not use direct assistance services (Quaglia, Razafindratsima, 2004).

⁷ Sexual relations with a person of the same sex, multiple partners or the use of a hard or soft drug during the year.

Furthermore, often in this type of investigation of difficult-to-reach populations, the biases are accentuated by the topics: survey subjects find it difficult to tell about marginal sexual behaviours, consumption of illegal products or violence that they have suffered, or to talk about a deadly disease related to stigmatized behaviours. Thus we must pursue ways to survey subjects in confidence.

3.3 Surveying In Confidence

“We cannot foresee how survey subjects will react to a survey. This is illustrated by the INED Fertility Survey, which aroused so much apprehension in those in charge of collection and which was so well received by interviewees and interviewers alike.” said Jacques Desabie in a public meeting.

This comment applies to many surveys that deal with sensitive topics. Witness all those men and women who broke the silence in the first surveys of contraception, divorce, sexual behaviour, violence—practices or situations that were painful or stigmatized and, at the time, may have been unacceptable or on the boundaries of the acceptable. Depending on the tolerance or opprobrium that weighs on the conduct surveyed, the statistical results may be incredibly accurate or, on the contrary, underestimated. Thus, at the end of the 1970s, the contraceptive practices documented by surveys of women perfectly matched statistics on the sale of pills and IUDs. On the other hand, whatever the method used or the period, the results of surveys on abortion substantially underestimate this practice (on the order of 40% to 65%) in relation to medical records, which are themselves deficient (Moreau et al., 2004).

The situation is not always one of obstruction. Between two surveys, the normative order may evolve. Take, for example, newspaper surveys of homosexuals on their sexual practices. In this period, the AIDS epidemic, highly concentrated in marginal and stigmatized groups, was a lethal menace. In 1985, the first survey broke the silence. Despite many open-ended questions, the men who answered found it difficult to confine themselves to standardized answers; half of the 1,000 questionnaires returned were accompanied by letters in which subjects described the traumatic experience of discovering their homosexuality, which was lived in solitude. The following year, the publication of the results in specialized journals sent back a commonplace image of the homosexual experience and lifestyles, and the 1986 survey, repeated according to the same procedure, collected 2,600 responses unaccompanied by personal commentaries.

Within the interview itself, some precautions are required in order to gain individuals’ confidence. Attention to the wording of the questions can prevent some blockage effects. Thus it is desirable to

- conduct a qualitative survey to study the vocabulary: an analysis of conversations with women recently faced with an unplanned pregnancy shows that women use very different terms to describe an elective abortion (Bajos, Ferrand, 2002);
- diversify the angles of approach and wordings; thus, as a result of an analysis of vocabulary relating to unplanned pregnancies, the Cocon (cohort on contraception) survey offered four ways of identifying elective abortions in the questionnaire;
- avoid naming traumatic or stigmatized behaviours (in the survey of violence against women, the words “violence” or “assault” are never used);
- limit the prescriptiveness of questions concerning stigmatized or culpable behaviours (homosexuality, drug addiction, etc.);
- pay attention to the vocabulary used. Because of their education, their morals, their culture or simply their lack of experience, some people are not prepared to hear, let alone use, the descriptive language employed in social surveys, which they may perceive as crude, immodest or inappropriate. For example, to spare minors interviewed face-to-face from having to use the language of sexuality, the survey on teenage sexual behaviour presents a written and numbered list of sexual behaviours; the teen interviewed enters the numbers corresponding to the activities with which he/she has experimented (Lagrange, Lhomond, 1997). In this type of situation, it is also common practice to use yes/no answers;

- avoid assigning a person a status with which he/she refused to be classified. For example, some persons of foreign origin do not identify with immigrant status. In France, a person from Portugal, which is now part of the European Union, has trouble accepting such a classification: “Me, an immigrant?”

The use of introductory questions can make the question less threatening and signal the research team’s tolerance of certain practices:

“Does polygamy exist in your country? Does your husband have more than one wife?”

“Did your former husband do this? And yourself?”

Lastly, as in the case of the first homosexuals surveyed in 1985, think of the liberating effect of the open-ended questions: they enable respondents to share their feelings about a unique experience that cannot be conveyed by standardized answers.

Confidence between interviewer and interviewee develops gradually, and interviewees can, at any point, halt the interview, especially if it is being conducted by telephone. They do so, but only rarely. This shows the importance of both the technical and ethical qualities of the questionnaire. The sincerity of the interviewee is obtained only on the existence of both conditions. Only then can full agreement from the interviewee be observed with the completion of the interview.

In the survey on sexuality in France, many interviewees would never have agreed in advance to the entire content of the survey but said that they were satisfied after co-operating in this regard to prevent AIDS among young people. Survey statisticians find it hard to avoid this ethical dilemma, but they draw strength from the social benefit around which their work revolves: an image of homosexuality as a diverse and commonplace reality; the prevention of AIDS; the national campaign against violence against women, called “Break the Silence”; the “New Beginning” campaign to reduce the 200,000 abortions in France despite contraception. The interviewees’ opinions reinforce these survey statisticians in their work, since on the whole, interviewees are proud to have contributed to these actions or to have shed light on a social problem—be it abortion, violence or sexuality—or helped to gain recognition for their group (homosexuals, drug addicts) or their demands or needs (persons with AIDS, inmates).

We cannot ignore the difficult ethical tradeoffs between the representativeness requirements of the survey and transparency or indeed loyalty toward the survey subjects. Thus it is difficult to state in advance everything that is going to be covered in a sensitive questionnaire. To avoid selective consent of the sample to such surveys, don’t we quite often reveal the topic of the survey only gradually? Isn’t the name of the survey often allusive? the introductory letter vague? the target population unstated?

When we are interested in individuals who see themselves as belonging to a fairly structured community of shared destiny, we must also obtain the confidence of the group. The definition, its consequence, and the delineation of the population, cannot ignore the struggle of individuals classified on the basis of unconventional behaviours that are often condemned. The statistical undertaking is then faced with protective or demand-driven strategies conveyed and relayed by the organized community. This may range from maintaining privacy—by omission, negation or, on the contrary, exaggeration of a positive self-image—to a demand that the classification be abandoned or redefined, or even a demand for redress. Obviously, the most powerless groups are not in a position to make such demands. Their members share only a “shameful differentness”; being too fragmented as individuals, they cannot join together to act in order to escape from their negative designation or turn it to their advantage.

In this situation, the co-operation of the spokespersons of the movement—if they exist—is indispensable at all stages of the survey, from its design through to the dissemination of the results. By the same token, if members of the research team are involved in the movement, this should not be considered as an obstacle to objectivity. On the contrary, such persons may be in a position to contribute a tone lacking the sentimentality and moralism so often encountered among intellectuals who defend minority groups to which they do not belong. They may also facilitate access to the field and reduce the mistrust that marginalized persons feel toward the questionnaire and its possible use by the police. The dissemination of the results to the persons concerned is then part of the researcher’s work. Objective information on stigmatized practices and lifestyles may even contribute to the process of liberating the group.

3.4 Attention to Statistical Confidentiality

The effort devoted to formulating questions and organizing the interview is often not enough to ensure that sensitive behaviours are measured without bias. To establish confidence, renewed attention must be paid to statistical confidentiality when a survey is conducted on disapproved behaviours or stigmatized groups. Such a situation only heightens the tensions generated by the double paradox of using intimate and anonymous information to create public data and to address a sub-population that is both exacting about “the truth” and highly suspicious of politics.

While there are well-developed techniques for protecting computer files, confidentiality during the interview is more difficult to implement, especially in a face-to-face interview. Interviewers and survey managers need more training in this area. One should, for example, avoid announcing the subject of the survey to the entire household; have the spouse leave the room during the interview; neutralize neighbours who come to celebrate the interviewer; balance the length of a self-administered questionnaire to avoid drawing the attention of obstreperous youngsters; refrain from testing a sensitive questionnaire on a female colleague, etc.

Unfortunately, in France, the tools of confidentiality are too often rejected: computer-assisted self-interviewing is not used; it is either face-to-face or by telephone. Private or public interviewer systems are not equipped with tactile screens to be used by interviewees, and the latest analog telephone systems are supposed to prevent self-administered telephone surveys. And yet the works of Turner and the entire RTI team have shown the extent to which sensitive behaviours can be under-reported unless these techniques are used: the reported prevalence of hard-to-discuss behaviours is sometimes three times greater when the survey has been self-administered by telephone or on a portable computer (Turner, 1998 or Rogers *et al.*, 1999).

Nevertheless, other techniques have been implemented. The simplest of these are not the least effective. The Escapad survey (Beck, 2004) is probably the most reliable French survey of drug addiction and alcohol and tobacco consumption: a group of young people simultaneously complete a self-administered paper questionnaire, then put it in a box in the room in full view—this is the secrecy of the ballot box applied to survey questionnaires. In contrast, epidemiologists use techniques involving the compression (hashing) of identifiers to deduplicate or match questionnaires (Quantin *et al.*, 2004). More simply, in the survey on inmate health (Désesquelles, 2002), the face-to-face questionnaire and the medical questionnaire completed by the physician are matched using commercial bar codes. In the Vespa survey (Obadia *et al.*, 2004) of persons with AIDS, a third party does the interface between the polling institute and the hospital to match face-to-face and medical questionnaires or refused questionnaires and medical questionnaires. As in the inmate survey, the objective of the system is to control or rectify the medical representativeness of the respondents.

4. CONCLUSION

Currently, in France, the legislative framework for collecting sensitive data from fragile or stigmatized populations is evolving substantially under the impetus of European laws and the need to produce statistics that are harmonized between countries, despite a history and cultural traditions that are extremely varied if not in contradiction. Related to this easing of the legislative framework is a slow expansion of the areas of enquiry. Questions on sexual behaviours or religion, which were taboo until recently, are no longer objectionable. The problems encountered in approaching difficult-to-reach populations are now more due to difficulties inherent in the subject to be studied than to legal or moral restrictions. As thinking continues to develop and as growing financial and human resources are committed, it is becoming possible to make improvements in a number of areas: techniques of approach, questionnaire and question design, the representativeness of the population contacted, and as a result, the statistical quality of the measurements produced. However, it is regrettable that practical experience and the theoretical output of sociology on the behaviours and paths of stigmatized or deviant persons are not better known and taken into account in the increasingly plentiful production of official statistics on fragile populations, and that French public and private polling institutes have not yet invested in methods of self-administered collection by telephone or on portable computers, since these are the best guarantees that the answers to sensitive questions will be sincere.

REFERENCES

- Bajos, N., Ferrand, M. and the GINE team (2002). "De la contraception à l'avortement. Sociologie des grossesses non-prévues", Paris: Inserm.
- Beck, F., Legleye, S. and Spilka, S. (2004), "Cannabis, alcool, tabac et autres drogues à la fin de l'adolescence : usages et évolutions récentes - ESCAPAD 2003", *Lettre Tendances*, No. 39, OFDT, 4p.
- Brousse, C., de la Rochère, B. and Massé, E. (2002), "Hébergement et distribution de repas chauds - Qui sont les sans-domicile usagers de ces services?", *INSEE première*, No. 824, 4p.
- Cassan, F. Mazuy, M. and Toulemon, L. (2001), "Douze millions et demi de grands-parents", *INSEE première*, No. 824, 4p.
- Cavailhes, J., Dutey, P. and Bach-Ignasse, G. (1984), *Rapport Gai. Enquête sur les modes de vie homosexuels*, Paris: Editions Persona.
- CNIS (2000), "Transposition en droit français de la directive européenne N°95/46/CE du 24 octobre 1995. Les attentes des statisticiens", Rapport No. 55, Paris: CNIS report.
- Desabie, J. (1990), Réunion du CNIS, formation Démographie-conditions de vie.
- Désesquelles, A. (2002). "Le handicap est plus fréquent en prison qu'à l'extérieur", *INSEE première*, 854, 4p.
- Lagrange H. and Lhomond B. (1997), "L'entrée dans la sexualité. Le comportement des jeunes dans le contexte du sida", Paris: La Découverte.
- Leridon H. (1979), "Les pratiques contraceptives en France en 1978", *Population et sociétés*, No. 120, 4p.
- Marpsat, M. and Firdion, J.-M. (2000), "La rue et le foyer. Une recherche sur les sans domicile et les mal-logés dans les années 1990", Paris, Ined (Travaux et documents, cahier 144).
- Messiah, A. and Mouret-Fourme E. (1993). "Homosexualité, bisexualité : éléments de socio-biographie sexuelle", *Population-F*, 5, pp. 1353-1379.
- Moreau, C., Bajos, N., Bouyer, J. and the COCON team (2004), "De la confusion de langage à l'oubli : la déclaration des IVG dans les enquêtes en population générale", *Population*, 3-4, pp. 503-518.
- Obadia, Y, Lert, F. and Le groupe Vespa (2004), "Comment vit-on en France avec le VIH/sida ?", *Population et sociétés*, No. 406, 4p.
- Pollak, M. and Schiltz, M.-A. (1991), "Questions de méthode. Six années d'enquête sur les homo- et bisexuels masculins face au sida", *Bulletin de Méthodologie Sociologique*, 31, pp. 32-48.
- Quaglia, M. and Razafindratsima, N. (2004), "INED's 2002 Survey of Homeless People ontacted by Itinerant Services". *Proceedings of Statistics Canada Symposium 2004 - Innovative Methods for Surveying Difficult-to-reach Populations*, Gatineau (Quebec) : November 3-5.
- Quantin C. et al. (2004), "Estimation de la valeur discriminante des traits d'identification utilisés pour le rapprochement des données d'un patient", *Revue d'Epidémiologie et de Santé Publique*, 52, pp. 431-440.
- Rogers S. M., Gribble J. N., Turner C. F. and Miller H. G. (1999), "Entretiens auto-administrés sur ordinateur et mesure des comportements sensibles", *Population-F*, 2, pp. 231-250.

- Ravaud J.-F., Letourmy A. and Ville I. (2002), "Les méthodes de délimitation de la population handicapée : l'approche de l'enquête Insee Vie quotidienne et santé", *Population-F*, 3, pp. 541-566.
- Riandey, B. and Firdion, J.M. (1993), "Vie personnelle et enquête téléphonique. L'exemple de l'enquête ACSF", *Population-F*, 5, pp. 1257-1280.
- Schiltz, M.-A., Jaspard, M., Brown, E., Firdion, J.-M. and the Enveff team (2001), "Enquête nationale sur les violences envers les femmes en France (Enveff) : mise en place d'une enquête sur un sujet sensible qui peut mettre en danger les femmes interrogées", In Dreesbeke J-J, Lebart L. (eds.) *Enquêtes, modèles et applications*, Paris: Dunod, pp. 219-230.
- Spira, A., Bajos, N. and le groupe ACSF (1993), "*Les comportements sexuels en France*", Paris: La Documentation française.
- Tincq, H. (2004), "Une enquête sur la religion en Europe confirme la montée de l'islamophobie et de l'antisémitisme", *Le Monde*, 15 décembre, p. 2 et GFK, *Wall Street Journal*, 10 décembre.
- Tribalat, M. (1995). *Faire France*, Paris : la Découverte.
- Turner, C.F., KU, L., Rogers, S.m., Lindberg, L. D., Plek, J. and Sonenstein, J. L., (1998), "Adolescent sexual behavior, drug use, and violence: Increased reporting with computer survey technology", *Science*, 280 (May 8), pp. 867-873.