

## USE OF PRIOR INFORMATION IN BLAISE CATI SURVEYS

*Vesa Kuusela, Hannu Virtanen, Markku Heiskanen  
Statistics Finland, Interview Division, Finland*

### 1. Introduction

In Statistics Finland, our experience of Blaise is limited to CATI surveys. The first CAPI survey is planned to take place at the beginning of 1994. The CATI module of Blaise has some interesting features that are useful in connection with the sampling schemes that are applied in Statistics Finland. An essential feature of the sample surveys is that at least the names, addresses and telephone numbers of the respondents are known in advance. Usually some additional characteristics (e.g. gender, age, profession) are known as well. This is not only typical of Finland and the other Scandinavian countries which have population registers. Similar situations arise frequently in panel surveys as well.

In Statistics Finland, the samples are usually drawn from the Central Population Register. In the register there are the names, addresses and social security numbers, with date of birth, gender, occupations and many other characteristics of all the people living permanently in Finland. This means that, in a survey, quite a lot can be known in advance concerning respondents and their households. On the other hand, the Blaise CATI module allows the insertion of appointment information into the system before the survey begins. That gives many options to plan when and at what time the respondents will be called.

In this paper we describe how we have applied the CATI module of Blaise in Statistics Finland. We also discuss what we have found good in the module, and what we think should be added or changed.

## 2. Sampling in Statistics Finland

All Scandinavian countries have maintained population registers for many centuries, and the registers are still maintained carefully. In Finland, the Central Population Register has about one month's lag due to population movement. The lag is the only notable fault in the register. The coverage of the population register in Finland is virtually 100% and the data is well updated, if we consider characteristics that don't change very often, like date of birth and gender. However, some information, such as that concerning occupations, is not very reliable.

The existence of a register where all of the population of the country is listed makes sampling much easier. It provides many possibilities for the surveyors, such as the application of complex sampling schemes. However, in Statistics Finland a *simple random sample* is usually drawn for all kinds of surveys so that no weighting is needed. For a CATI survey a simple random sample is still more appropriate than for other designs because it doesn't lead to any extra costs as it does, for example, in CAPI surveys.

Not all information available in the register is used in surveys. Generally, the records in the sample file comprise at least the name, address and social security number (i.e. personal ID) of the respondent. In addition, information about two reference persons, that is, the names of the two oldest persons living in the same household, are included in the record. This is done because the target person may be a child, housewife, or student living at home, etc., who doesn't have a registered telephone number.

There are no telephone numbers in the population register. Either interviewers (and supervisors) must find them or the local telephone company searchers for them. To find the right number it is necessary to have information about the reference persons.

According to the latest research, the overall telephone coverage in Finland is about 94%, but when the telephone numbers are attached to a random sample only a little over 80% (81%-84%) of respondents can be provided

with one. We do not know exactly what causes this gap. Numbers which are not listed in the telephone directory explain only 3-5 percentage points of it. It is probable that the remainder is due to people living in institutions, students, etc.

### **3. Contents of the dial screen**

The interviewers like to have as much advance information as possible about the respondents and their households. They say that this helps them to familiarize themselves with the respondents before they start. The interviewers think that this knowledge makes it easier to get a personal rapport with the respondent.

The first screen in the Blaise CATI interview, the *Dial screen*, provides rather good facilities for describing the respondent. Half of the screen is reserved for the *Memo window* where the contact information, like phone number, name, address etc. is displayed. However, in Statistics Finland we have faced situations in panel surveys when the available space on the screen is too small.

### **4. Making appointments before the survey begins**

The way appointment information is handled in Blaise makes it possible to utilize some of the available prior information. Blaise stores the appointment information in the *appointment file* (with the extension *.A02*). The designated date and time of the appointment can be stored in that file even before the Blaise file conversion. For the interviewing, only those cases which are scheduled to be interviewed that particular day (or earlier, if still outstanding) are included in the *day batch*.

In Statistics Finland, we have used prior information mainly to plan the time of the first attempt to interview the respondent. We describe below how we have used the facility for inserting appointment information in the

appointment file.

### **5. Cases for which telephone numbers are not known**

Everyone in the sample file has an address. A prevailing practice is to send a letter in advance to the respondents, telling them that an interviewer will contact them in the near future. In a CATI survey the letter is sent even to those respondents whose telephone numbers are not known. However, they receive a slightly different letter, in which they are given the chance to participate in the survey if they want to. All they have to do is to give, by letter or telephone, a telephone number where they can be reached, and their preferred time to be interviewed. Naturally, they can also call directly to the CATI center to be interviewed.

Supervisors then enter the telephone numbers and the appointment information in the appointment file by using the *Treat Form* option in the *CATI management* menu. The respondents whose telephone numbers are not known are kept out of the *day batches* (until their numbers are announced) by making appointments for them, in advance, for the last day of the survey. On the last day, there is no interviewing.

### **6. Survey on the usage of farming fields**

The purpose of the survey on the usage of farming fields was to get as early as possible an estimate of the field areas sowed for different grains.

In Finland, it takes a couple of weeks for summer to spread from the south to the north. This means that when farmers in the southern parts of the country have already sown their fields, the farmers in the northern parts are still waiting for the ice in their fields to melt. If respondents were selected completely randomly, some of the respondents would not even have started sowing by the time that they were telephoned. The client, the Ministry of Farming, was not interested in farmers' plans, but in how much they actu-

ally sowed of the different grains. Plans can change for many reasons. The easiest way to handle this had been to let them be 'too early' in the *day batch*, and make new appointments for them. However, this was not considered to be an appropriate manner to handle a respondent.

The solution was to include farms in waves in the *day batches*, starting from the coastal areas in south and southwest Finland, and advancing gradually to the north. In practice, this was carried out by making an "appointment" for each farm according to the whereabouts of the farm in the country. The first wave was interviewed in mid-May, and the fifth wave in the most northern part of Finland was interviewed just before mid-June.

## **7. Call scheduling on the basis of respondents' characteristics**

In some surveys we have tried to allocate the time of the *dials* according to the personal characteristics of the respondents. The idea is that it is more probable, e.g. that a pensioner is home during the day time than a person of working age. In practice, this means that only elderly people were called during the day time, and only people of working age in the evenings. It is an interesting subject for research, to find out which kinds of people have the highest chance of being found at home during the day, and which in the evenings.

## **8. Discussion**

The interviewers in Statistics Finland have found it good that they have information about the respondents and their households in the *Dial screen*. They feel that they know something of the household in advance and they can adapt better to the respondent's world, even in CATI surveys. This is thought to reduce the number of refusals.

The costs of a survey depend partly on the total number of dials needed for an interview. In the ideal case, all respondents would answer on the first

## *Use of prior information in Blaise CATI surveys*

call and give answers quickly and precisely. However, in Finland only about one-third of the sample is reached by the first attempt. On the other hand, the great majority of cases are reached by the third dial. The problem is that the cases which are hard to reach cause quite a lot of extra work. We have found it very difficult to decide the exact time at which to "cut off the tail" of the sample.

In a CATI survey, it is not the interviewer but the system which decides when to dial a specific number. Decisions about controlling the appointment patterns have to be made before the interviewing begins. Suppose that, in the survey on the usage of farming fields, summer arrives faster or more slowly than expected. In that case, it would be reasonable to change the appointments of all respondents who have not yet been interviewed. However, it is not very easy to access the Blaise appointment file for that purpose, except one record at a time. A batch access method would be far more convenient.

In general, the *Treat Form* option is too "powerful" and a bit cumbersome for dealing with the telephone numbers alone. There should be a separate procedure for the input and maintenance of telephone numbers.

For *busy telephone* there is quite a good scheduling procedure. A busy telephone will be dialled again after 5 minutes and again after intervals of 5, 10, 10 and 15 minutes. Users can define the time intervals, but we have found the default values satisfactory. It would be useful to have a similar procedure for the *No answer* cases as well.

Suppose the respondent is having a two weeks holiday trip. It is possible for the case to have been in the *day batch* for the maximum number of times before the respondent gets back. If one could define that, after a first No Answer, the telephone number is called, say, at intervals of 1 day, 2 days, 2 days, 3 days, 3 days and 5 days, the respondent could be reached after returning from the trip.

In version 2.4, the *Dial screen* tells the interviewer the result of the last dial

and who made it. However, it does not give the total number of dials and calls. Knowledge of the total dialling history would make it easier to evaluate if the sampled person might answer, or is away on a holiday trip, or if the telephone number is wrong. It would be easier than at present to decide whether or not the sampled person should be left as a non-respondent.

The layout of the *Dial screen* serves pretty well as it is now. However, we think that it could serve even better if the layout was slightly different. The size of the *Memo window* has been a problem in our surveys. It is rather small if one wants to know more than the basic facts about the respondent. Moreover, the shape of the window is impractical. If it were wider it would be more legible.