On the Necessity to Include Personal Interviewing in Mixed-Mode Surveys

Bart Buelens[1], Jan van den Brakel[1]

[1] Statistics Netherlands

There are several reasons to consider mixed-mode data collection procedures in survey sampling. De Leeuw (2005) emphasized that mixing modes of data collection can improve data quality and response rates, since the disadvantages of one mode can be compensated with the advantages of other modes. A second driving force behind mixed-mode designs is the reduction of administration costs of a survey. Cost is an important limiting factor in survey research and is the main reason that the Dutch Integrated Safety Monitor (ISM) employs a sequential mixed-mode strategy using web interviewing (WI), paper and pencil interviewing (PAPI), computer assisted telephone interviewing (CATI) and computer assisted personal interviewing (CAPI). While CAPI is the most costly mode, it is applied in the ISM to reach subpopulations traditionally known to be difficult to contact through the other modes. The ISM has now been conducted twice. In this paper the importance of including CAPI is assessed. It is investigated how sensitive the representativity of the response and the survey outcomes are to omission of the CAPI mode.

Survey design

The purpose of the ISM is to publish figures about justice and crime victimisation, police performance, and security perception of the Dutch population aged 15 years and over. In its present form, this survey has been conducted in the fourth quarter of 2008 and 2009. A stratified sample of persons is drawn using the 25 police districts in the Netherlands as stratum variable.

The data collection of the ISM follows a sequential mixed-mode design. All sampled persons are approached for WI with an advance letter inviting them to complete a questionnaire on the website of Statistics Netherlands. Alternatively respondents can receive a PAPI-questionnaire on their request. Persons with a listed landline or mobile telephone number who did not respond after two reminders are approached by telephone by an interviewer (CATI). The remaining persons who did not respond after two reminders are visited by an interviewer at home (CAPI). Table 1 shows the sample size, response rate, and the proportion of the response collected through CAPI.

Table 1   Size of ISM sample, response, and proportion of the response collected through CAPI in 2008 and 2009.

	Sample size	Response rate	Total response	Proportion CAPI
2008	28,309	59.9%	16,964	10.5%
2009	32,852	58.4%	19,202	11.2%

While the 2008 sample is smaller than the 2009 sample, the response rate is close to 60 percent in both years, with a proportion collected through CAPI of just over 10 percent.

The weighting procedure of the ISM is based on linear weighting. This implies that estimates of population parameters are obtained from functions of the observations in the sample expanded with the so-called survey weights. The generalised regression (GREG) estimator is used to derive the survey weights from the sampling design and the available auxiliary information about the target population (Särndal, Swensson, and Wretman 1992). The weighting model includes age, gender, ethnicity, household size, urbanisation and police district.

Assessing the effect of personal interviewing

While CAPI is the most expensive mode, it was believed that this mode is necessary to obtain a sufficient representation in the sample of groups that are known to be difficult to reach, such as young people and immigrants. Since the ISM has been conducted twice, the necessity of the CAPI mode can be studied in a relatively straightforward manner. The representativity of the response obtained with WI, PAPI and CATI and the additional value of CAPI is analysed in the next section with respect to some auxiliary variables for which the true population distributions are known. Subsequently the effect of the CAPI mode on the outcomes of the ISM is analysed by comparing estimates based on the complete sample with estimates based on the sample with the CAPI response omitted.

Characterisation of CAPI respondents

Mode dependent selection effects can cause the sample units who responded through CAPI to differ from those who responded through the other modes. To charactarise the CAPI respondents a weighted logistic regression model is fitted, with the weights those obtained through GREG estimation. By including these survey weights, selectivity with respect to the variables included in the weighting model is removed prior to studying their correlation with response mode. The dependent variable is a binary indicator for responding through CAPI.

In a first model, the independent variables are the same as those used in the weighting model, without interactions. The regression coefficient of gender is not significantly different from zero, see Table 2. The remaining variables are significant, police district having a larger p-value than the others. Two more

Table 2

Variable	Model 1	Model 2	Model 3
Age(11)	2.62e-137***	5.05e-135***	1.23e-134***
Gender(2)	0.60	n.a.	n.a
Ethnicity (2)	5.78e-122***	4.33e-120***	9.15e-120***
Household size(5)	3.08e-27***	n.a.	1.38e-26***
Urbanisation (5)	4.47e-8***	7.93e-13***	n.a.
Police district(25)	3.48e-3**	n.a.	n.a.
Ethnicity x household size	n.a.	n.a.	3.36e-6***
Age x household size	n.a.	n.a.	1.68e-4***
Age x ethnicity	n.a.	7.60e-4***	2.15e-3**
Urbanisation x age	n.a.	2.79e-3**	n.a.
Urbanisation x ethnicity	n.a.	1.0	n.a.

P-values of Wald-test statistics of independent variables in logistic regression modeling with 'responding through CAPT as the binary dependent variable. Data from 2009 is used. All variables are categorical, with the number of categories given in brackets with the main effects. N.a.=not applicable (term not included in model), *=p<0.05, **=p<0.01, ***=p<0.001.

models are fitted (see Table 2). Gender and police district are left out, but interaction terms with urbanisation and household size are now included. The only interaction that is not significant is that of urbanisation and ethnicity. While Table 2 is based on 2009 data, similar results are obtained for 2008. These findings imply that CAPI is a data collection mode suited to reach particular subpopulations with respect to age, ethnicity, household size and urbanisation.

Comparing the distributions of these variables in the CAPI group and in the non-CAPI group, CAPI is found to be successful in reaching young people, non-western immigrants, small households, and people living in urban areas. As an example, the distribution of age in the 2009 sample is shown in Figure 1. The pattern that emerges is due to selection effects only, as oversampling of specific demographic strata is not part of the sampling design. Among the CAPI respondents, the age groups of up to the age of 39 are overrepresented, while older respondents are underrepresented. For the non-CAPI group the reverse is true, though the deviation from the distribution in the population is not as severe. The distribution in the complete sample is the same as that in the population, since the sample is weighted by age.

If the CAPI respondents are omitted from the sample, the remaining response will have an underrepresentation of the younger age groups. Similar arguments apply to ethnicity, houshold size and urbanisation, and show that CAPI is a mode particularly suited to reach non-western immigrants, small families, and people living in cities.

Applying the GREG weighting model used for the complete sample to the sample excluding the CAPI response results in a new set of weights. Figure 2 shows these new weights together with the weights of the complete sample, for the ISM data of 2009. As expected the distribution of weights for the sample
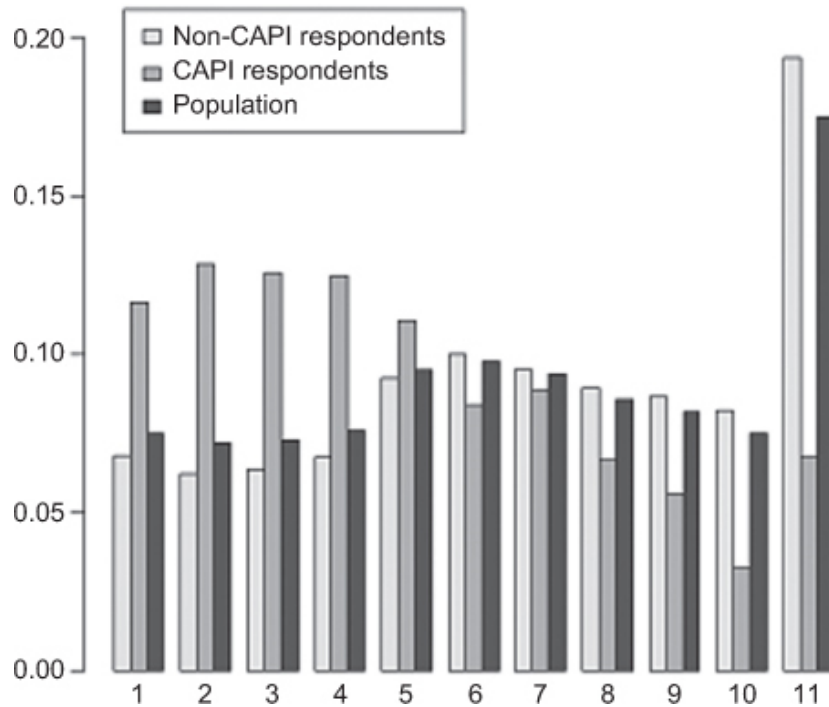
Figure 1   The distribution of CAPI and non-CAPI respondents across the age groups for the 2009 sample.

without CAPI respondents has a larger tail to the right, indicating that the sample is less representative. Figure 3 presents the distribution of weights by age group. In the sample without CAPI, the weights of the younger age groups are larger and more widely spread than those of the complete sample, while the differences in weights are smaller for the older age groups. This confirms that omitting CAPI respondents has an adverse effect on the representation of younger people. Similarly, the representation of non-western immigrants, small households, and people from urban areas is worse in the sample without CAPI than it is in the complete sample.

Survey outcomes

Some statistical outcomes of the ISM are shown in Table 3. This table contains estimates for eight key variables, for 2008 and 2009, calculated using the complete sample and the sample without CAPI. There is little difference between the point estimates based on the two samples. Ordinary tests for significance of the observed differences cannot be applied since the samples are not independent, due to the sample overlap. The point estimates based on the sample without CAPI are mostly within two standard errors of the estimates based on the full sample. There are three variables in 2009 for which this is not the case, as indicated in Table 3. This does not occur in 2008. These differences are caused by selection effects and mode effects, which are strongly confounded. Jäckle, Roberts, and Lynn (2010) discuss the issue of comparability of microdata collected using different modes.
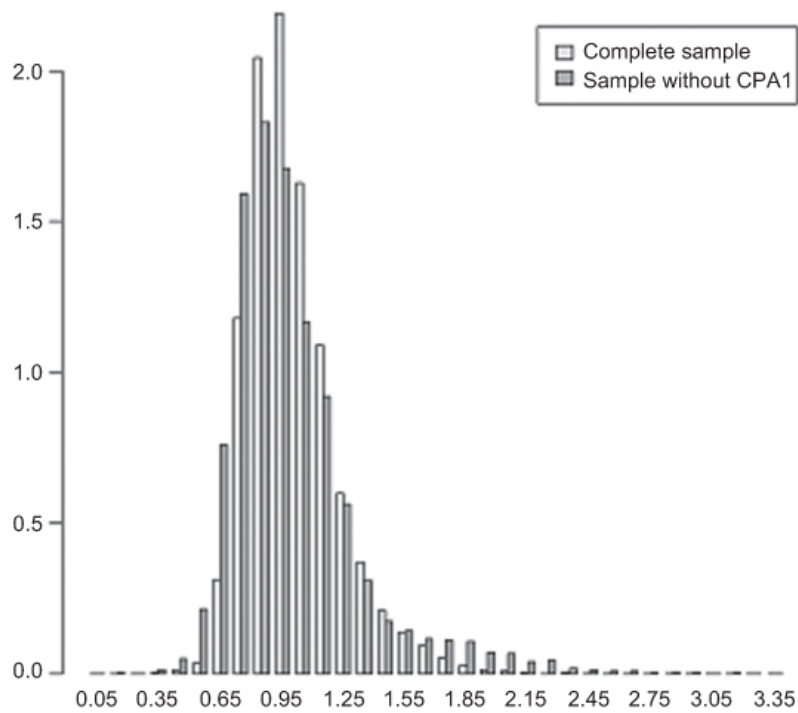
Figure 2   Relative distribution of all correction weights for the complete sample and the sample without CAPI, for 2009.
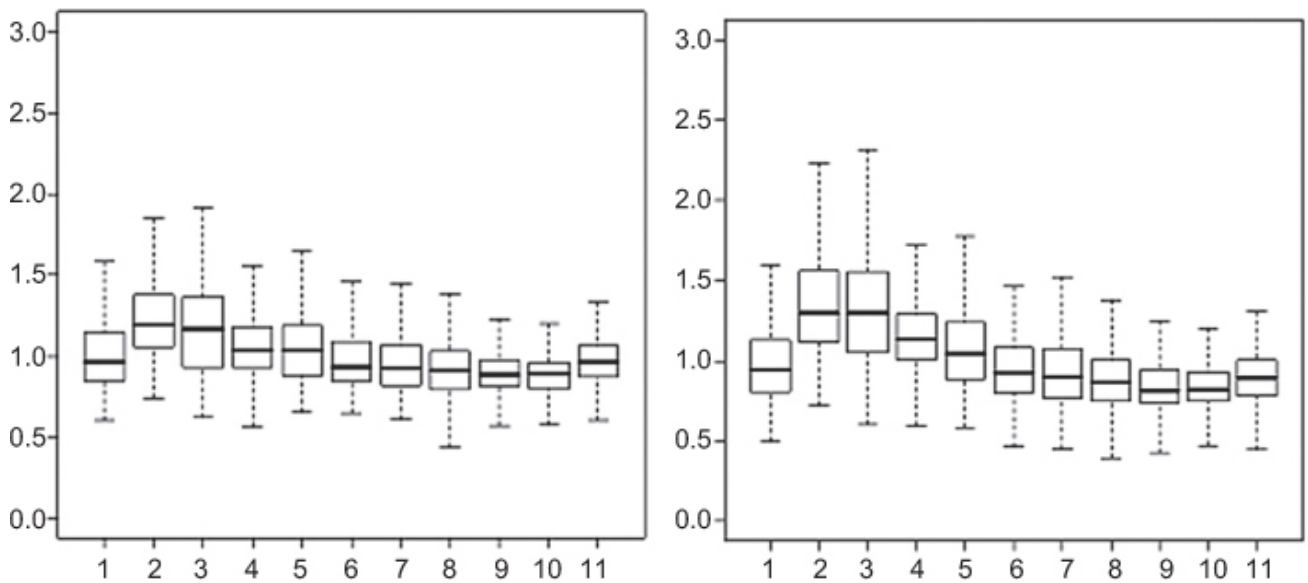

Figure 3   Left panel: Distribution of correction weights of the complete sample over 11 age groups for 2009. Right panel: Distribution of the weights for the sample excluding CAPI respondents.

The standard errors of the estimates based on the sample excluding CAPI are larger than those based on the complete sample. The reason for this is twofold: there is more variation between the weights, as shown before, and the sample size becomes smaller when omitting part of the response.

Table 3

Variable	2008				2009			
	Complete sample	Std. err.	No CAPI	Std. err.	Compile sample	Std. err.	No CAPI	Std. err.
Victim	25.80	0.43	25.82	0.52	26.04	0.34	25.90	0.38
Crimes	0.43	0.01	0.43	0.01	0.45	0.01	0.44	0.01
Nuisance	1.65	0.02	1.66	0.03	1.70	0.02	1.72	0.02
Degrad	3.62	0.03	3.67	0.03	3.63	0.02	3.70**	0.02
Unsafe	24.33	0.42	24.86	0.51	26.05	0.34	26.77*	0.38
Polcont	31.13	0.45	30.63	0.53	31.88	0.36	30.94**	0.40
Satisfied	59.74	0.91	59.95	1.09	58.04	0.71	59.16	0.78
Judgem	43.56	0.48	43.16	0.57	41.27	0.39	40.89	0.42

Estimates for eight key ISM variables: the percentage of people that have been victim to at least one criminal offence in the last year (victim), the number of crimes per person (crimes), an indicator on a scale 1–10 for the degree to which people perceive nuisance from youngsters, drug addicts, noisy neighbours (nuisance), and indicator (1–10) how people perceive degradation of their neighbourhood by grafitti, litter etc. (degrad), the percentage of people feeling unsafe at times (unsafe), the percentage of people who had contact with the police in the last year (polcont), the percentage of people who said to be satisfied with the police at that contact (satisfied), and the percentage of people satisfied with police performance in general (judgem). Significance of differences between outcomes based on the complete sample and the sample without CAPI are indicated by * and ** for p-values of <0.05 and <0.01 respectively (see main text for details).

Table 4 Annual mutation in the estimates from 2008 to 2009 for the eight key variables.

Variable	Complete sample	No CAPI
Victim	0.91%	0.34%
Crimes	3.11%	3.75%
Nuisance	3.17%	3.65%
Degrad	0.19%	0.65%
Unsafe	7.07%	7.68%
Polcont	2.43%	1.00%
Satisfied	-2.84%	-1.33%
Judgem	-5.26%	-5.28%

Table 4 shows the annual mutation in the estimates of the eight survey variables. While there are some differences between estimates based on the sample including CAPI and that without CAPI, there are no survey variables for which the two versions report conflicting developments.

Conclusions

The CAPI data collection mode is important to reach certain subpopulations that are otherwise underrepresented in the response. In particular, including CAPI is beneficial for the representation of young people, non-western immigrants, and people living in urban areas. Not conducting CAPI interviews causes the ISM response to be less representative, resulting in more extreme survey weights. Generally, this will have a negative effect on the variance estimates of the survey variables. Some of the eight survey variables considered are estimated differently based on the sample without CAPI compared to the

complete sample, although in most cases there is no significant difference. The substantial reduction in survey cost that can be achieved by omitting the CAPI mode would come at the expense of a decrease in accuracy

REFERENCES

De Leeuw, E. 2005. "To Mix or Not to Mix Data Collection Modes in Surveys." *Journal of Official Statistics* 21: 233–55.

Jäckle, A., C. Roberts, and P. Lynn. 2010. "Assessing the Effect of Data Collection Mode on Measurement." *International Statistical Review* 78: 3–20.

Särndal, C.-E., B. Swensson, and J. Wretman. 1992. *Model Assisted Survey Sampling*. New York: Springer-Verlag.