

Experiences of Two Capi-Papi -- Comparisons in Finland

Markku Heiskanen, Kirsti Ahlqvist, Statistics Finland

One of the cornerstones of quantitative survey interviewing is the standardized measuring situation. Survey questions are often considered as certain kinds of fixed stimuli to similarly rouse the expected responses (concerning facts, opinions, values) in every population group studied. The survey dogma assumes that the measurement situation can (and must) be firmly controlled. On the other hand, there are many studies indicating that even the slightest alteration of the details may cause dramatic effects in the results. Features like clothing of the interviewer, his/her behaviour and attitudes, voice emphasis, small changes in question wording, interview modes, sampling and so on provide a jungle of sources to spoil the results. Could the portable be one of them ?

The influence of technology and engineering has become a more decisive factor in the interview process with the CAI-system. Routing errors disappear, all relevant questions are asked, coding systems are incorporated, no basic data entry is needed afterwards. These means naturally improve the control of measurement, but they have not been major problems in our surveys so far. The main problem is whether the respondent gives "the right" answer to the questions, or are there obstacles in the interviewing situation that prevent the respondent from revealing him/herself.

When our interviewers first began computer assisted interviewing in 1993, the question of the contents of the Capi data was widely discussed. Two larger Capi/Papi comparisons have been made up to the present. The first one was a comparison of our Survey on Living Conditions (SLC) in 1993, the second one the Household Expenditure Survey (HES) in 1995. Both are large surveys with a face to face interview. Some of the main results of these mode comparisons are reported in this article.

1. The Survey on Living Conditions

The pilot Survey on Living Conditions was carried out in the autumn of 1993 to compare Capi and Papi results. For all interviewers it was their first contact with the respondent using Capi- method and for most interviewers with computer assisted interviewing in general (a couple of Cati studies had been compiled at a small Cati-center). This may account

for the fact that the average duration of the Papi interviews was 65 minutes and that of Capi interviews 75 minutes.

A random sample of Finnish population consisting of 1 000 persons aged 25-54 years was drawn from the Finnish Central Population Register. One half of the sample was interviewed by Papi-method, the other half by Capi-method. Since 140 interviewers participated in the mode comparison, they only had a few interviews of both types. Thus the interviewer effect on personal level was kept small. The type of the starting interview was randomised to each interviewer. The rate of nonresponse was 24% in both samples, 3 percentage units lower than in the actual survey conducted in spring 1994. Most of the background variables, such as gender, age and place of residence, did not show any significant differences in participation rates. People with a higher level of education were, however, slightly overrepresented in the Papi-dataset, a fact that was not corrected with the weighting of the results by gender, age and region. Even so weights are used in the comparison.

The main topics of the SLC were household/family structure, living conditions in childhood, residence and housing conditions, social activity and leisure activities, contacts and relations to relatives and friends, education and training, work and working relations, health and security, and evaluation of own life.

A cursory comparison of most questions was made using chi-square and t-tests. Most variables were categorical. Of the 329 questions tested (the complicated question sets of household structure and sickness/disease lists were omitted from the overall comparison), there was a statistically significant difference ($p < 0.05$) in 31 questions.

Table 1. Differences between Capi and Papi results in the Survey on Living Conditions, 1993.

Topic		

Total		
Background		
Everyday life		
Activity in organisations		
Social support		
Work		
Health		
"Values of life"		

Hence nearly one out of ten questions produced different results. This was quite an surprising result; more particularly since due to small sample sizes the tests do not give significant differences except in cases where the results are also primarily interpreted as significant in contents and meaning; in extreme cases not even a difference of 8 percentage units (a dichotomised variable) did exceed the significance level. In addition, some variable classes were combined because of zero observations. In relative terms, most differences between Capi and Papi results were found in questions under the topics Everyday life and Activity in organisations (20% of the results in each topic differed).

Yet larger differences have been reported in mode comparisons. E.g. in the study of Baker et al. the overall per cent of Capi-Papi differences was

13.5 (Baker et al. 1995, 419; they compared 445 questions in the US National Longitudinal Survey of Labor Market Experience).

It is not possible to say which method yielded more reliable results. E.g. the Labour Force Survey of Statistics Finland indicates that in 1993 85% of the people with earned income were wage earners. The corresponding number was 82% in Capi results and 88% in Papi. With most variables there exists no comparison possibilities. Nevertheless, we can look at the trends in the differences: does one or the other method produce more or less certain activities or positive stands.

The topic Everyday life includes questions of different leisure time activities. The Papi results showed in four cases more activity than the Capi results. In one variable (sports activities) there were more active participants in Capi results. The results of the topic Activity in organisations were in line with the leisure time questions above: the Papi results reported more activity than the Capi results.

Unofficial social aid and support (from relatives, friends, neighbours, fellow workers) is received and given more often according to the Capi results. The number of friends was also greater in Capi. But in four out of 12 cases social aid was received or given more often in Papi results. In work-related questions fear connected to working conditions was more common in Papi in three cases, in Capi in two cases.

It is quite common to carry out mode comparisons by comparing differences in the results on total level with certain key variables. The need for details in comparison depends on the purpose the results are used for. Total counts are often an adequate accuracy in statistics production. But in social research the behaviour of different population groups often needs examination. Do female respondents differ from males? Or different age groups, or different social strata ?

We chose two topics of the SLC for a more detailed analysis. The questions concerning social support and health are also used by many other Finnish researchers, not merely by those of Statistics Finland. These questions may to some degree be sensitive or delicate to the respondent. E.g. in social relations, revealing that one does not have friends, can rouse unpleasant feelings in the respondent. On the whole, neither of the topics is extremely delicate (like drug use or sexual behaviour) but represent the middle range sensitivity in social surveys.

Although the health responses on the main distribution level had only one significant mode difference in the 57 questions tested, the detailed analysis of background variables showed that education, which as a variable reflects the social status of the respondent, was a key determinant causing differences in the mode comparisons that were not found in the total distributions (Ahola 1996). The common formula of mode differences was as follows: people with the highest level of education report more health deficiencies in Capi than in Papi results, whereas respondents with a low level of education reported more health deficiencies in Papi interviews. The mode differences were also larger in the low education group. One explanation to this might be that the change in the interview situation had raised the status of the interviewer compared

with that of the interviewee. Or that the interaction situation had become more official and formal, and thus preventing the information mediation process. The differences were smaller in health symptoms when socially desirable questions were compared (e.g. fatigue, over-exertion) than in questions which were socially less desirable (e. g. questions concerning depression).

The result that population groups with a lower level of education give a more positive report of their health in Capi seems somewhat contrary to Baker et al.'s statement that Capi yields greater respondent willingness to disclose sensitive information (Baker et al. 1995, 413).

Certain response structures remained the same in both modes. Both in the Capi and Papi results women reported more health symptoms than men. The structure of commonness of different symptoms was also maintained by men and women.

An analysis by gender, age group and education was also made with certain key variables in the topic Social support. The conclusion drawn of that comparison was that statistically significant, to certain direction consistent, mode differences could not be found, although older respondents seemed to give somewhat more positive picture of their social relations in the Capi results than in the Papi responses (Heiskanen 1995).

As a result of these mode comparisons, the question was raised that the international results, where in general no significant differences between Capi and Papi modes were found (e.g. de Leeuw 1992, Martin et. al., 1993), might possibly not hold in all cultures. Finnish society has been characterised by a small social distance between different social stratas, thus small changes in the interaction situation might influence their responding stand, particularly among people with the lowest level of education.

2. The Household Expenditure Survey

When the second mode comparison study was carried out the interviewers of Statistics Finland had worked with computer assisted interviews for about a year. The HES had formerly been conducted every five years with the Papi method, but from 1994 onwards it was carried out on an annual basis and with a smaller sample size. The one-year sample was divided into two subsamples of equal size in 1995; one half was interviewed by Capi, the other half by Papi.

The interviewers are trained in the subject matter more carefully than in many other surveys of Statistics Finland. The survey consists of a face to face interview and a two weeks' diary of household expenditures. The datasets were collected during the whole year 1995. Sample sizes were 1721 households in Capi and 1729 in Papi. Response rates were nearly the same (68% in Capi, 67% in Papi).

The length of the interview was 72 minutes in Capi and 70 in Papi. Of the 7 interviews that lasted over three hours only one was made by Papi.

Here the two research questions were: 1) Are there significant differences between the Capi and Papi results and 2) Does the social status of the household (education of the head of the household) affect the responses in a similar way as in the SLC? For the question 1) we chose preliminarily 13 commodities or groups of commodities (or payments) for the mode comparison. Three commodities from the diary datasets were also included in the comparison.

The 13 chosen expenditure items cover 50 per cent of the consumer expenditures obtained from the interview, and 20 per cent of all consumer expenditures. When questions concerning acquisition of durable commodities, such as furniture, car, household or entertainment equipment, pc-computer or sports and hiking equipment, were asked response cards were used. The interviewee could see the commodity names in these cards.

In most of the questions the answering proceeded in two phases. In the first phase the interviewee was asked whether the household had acquired certain commodities. If this question was answered with yes-alternative, then the price of the commodity (or payment) was inquired. The comparisons are presented in a similar way: Table 2 shows first the percentage of households who had had the questioned consumption activity during the reference period, then yearly mean values of expenditures calculated on those who had the activity, and standard errors of the means. The results were calculated considering the two stage sampling frame.

The results in Tables 2 and 3 have been weighted to correspond with the Finnish household structure. Thus they differ slightly from the nonweighted ones, e.g. the proportion of one person households increases. Using weighted results may be somewhat problematic in mode comparisons, as mode differences can change when weights are used. If we look at the results from a statistical standpoint, then weighting gives a more accurate description of the population studied, and is also the "final truth" published in reports.

Table 2. Differences between Capi and Papi results in the Household Expenditure Survey, 1995¹¹.

	Capi	Papi
Goods (variable code, reference time in months)		
Source: interviews		
Charges for owner-occupied dwelling (300,1)		
Rent of dwelling (301,1)		
Electricity (31000,12)		
Furniture (4000-,12)		
Household		

¹¹We would like to thank Pertti Kangassalo for his help with Sudaan programming.

appliances and
equipment (420,12)

Radio, television,
tape recorder etc.
(700,12)

Personal
computers and
peripherals
(70600,12)

Sports equipment,
hiking gear (7170,
12)

Purchase of car
(600,12)

Physicians' fees
(5100, 3)

Interest payments
on household loans
(302,12)

Amount of loan for
dwelling (asuntvel,
interview)

Amount of
consumption loan
(kulutvel, interview)

Source: diaries

Meat (101, ½)

Alcohol beverages
(121,½)

Outer
garments/outdoor
clothes (200,½)

The interview results show no statistically significant differences between the modes in the main distributions (% of sample). Neither do the means reveal surprising results. Some smaller differences in the results however exist. In six Papi responses the prevalence rate was slightly higher (the widest difference was 3 percentage units), in three questions the Capi mode yielded more commodity purchases. The same holds for the consumption means; in nine cases they were somewhat higher in the Papi results. These results are in line with the results of the British consumer expenditure mode comparison (Manners 1995). Differences of same magnitude also exist in the diary comparison, with the exception of the only statistically significant difference in Outer garments/outdoor clothes.

The second research question was whether the different social strata react differently to the mode change. The respondents were classified into three groups according to the education of the head person of the household. This variable is supposed to describe roughly the habitus of the family. The classification has been made according to the length of education of the head of family (low=education less than 10 years, middle=10-12 years, high=13 years and more). Age of the head person is expected to describe the phase of life of the household.

Table 3, in which the consumption means of the mode comparison variables are presented as an index ratio by education (Capi/Papi*100), shows one significant difference ($p < 0,05$) both among the interview variables and the diary variables. On the total consumption level, differences are in practice non-existent. By looking at the trends in mode differences - including those that did not exceed the significance level - we find however that in the interview variables by both lower and middle education the means are slightly higher in Capi in 5 variables, and as to Papi in 8 variables. In families with a higher level of education, higher consumption was measured in 9 Capi variables, and in 3 Papi variables. Mode differences in the highest education group are quite small except in physicians' fees ($p = 0,11$), but show nevertheless a slight difference compared to the lower education groups.

Families, where the head person is aged 60 years or over, seem to have somewhat higher total consumption in Papi than in Capi ($p = 0,09$). In loan and sports equipment variables, considerable differences were also found.

This may also indicate an incoherent consumption structure within the group (a couple of outliers, but only on the total level, were omitted from the analysis).

There existed also some statistically significant differences in the age group 40-49 years. In age groups over 30 years the interview variables in the Papi results give somewhat more often higher consumption means, while in the group under 30 years the consumption is higher in the Capi results.

The consumption level and structure differ when education increases. The attitude to consumption probably differs as well. In Finland a sparing consumption style has been traditionally appreciated. At the end of the 1980s a consuming lifestyle and living involving debts entered the traditional consuming patterns, and mostly in younger and higher educated families. The recession in the 1990s ended the "cheerful consumption celebration" and consumers were seen guilty of the recession. This might stand at the background of prevailing consuming attitudes, so that the people having consumed more in the 1990s, might be a sensitive group regarding the mode comparison (Ahlqvist 1996).

3. Discussion

Dufour et al write in their report of the conversion of the Canadian Labour Force Survey to CAI: "The first lesson learned was that any major change in a complicated process requires time to stabilize". In this respect the timing of the Finnish SLC mode comparison as the first Capi project of Statistics Finland was not the best possible. The longer interviewing average in the Capi mode already indicates in this direction.

The interviewers had a very positive attitude to the CAI from the beginning (90% were for, 9% don't know; an inquiry immediately after the SLC). Some drawbacks were nevertheless experienced. 4% of the interviewers reported that the interviewee was shy of the computer, 16% had found that the computer made the interview situation rigid. 21% of the interviewers evaluated that the computer hampered the interaction, and 25% that the computer program restricted the conversation. These results show that the interviewers (and the researchers/programmers) were not yet sufficiently acquainted with the CAI-system. In the actual SLC compiled in spring 1994 the interviewers did not complain about interaction troubles any longer, but on the contrary felt that the computer facilitated the interview situation (a discussion with the interviewers of the Helsinki region, Manderbacka et al. 1994).

The mode differences in the SLC were quite numerous. There might have been more of them in case the samples had had a wider age distribution. If the assumption of different reactions to computer in different educational positions would hold good, it might have become stronger if the young and the old ones had been contrasted.

It might be argued that differences caused by inexperience will balance with time. Some comparisons with the mode study results and the main

SLC 1994 results were made. In some questions in some population groups the results in the main SLC in 1994 did get closer to the Papi results. But in some questions they did not, which reminds that each survey measures situations that are inconstant and subject to variation (here seasonal change in interviewing).

The Houshold Expenditure Survey, where the questions are more fact-like, seems to produce fairly comparable results with different modes. The same structure as in the SLC seems to be at the background though: in the group of people with a high level of education the consumption is slightly higher in many variables by Capi, while people having a lower level of education report higher consumption in Papi.

Table 3. Differences between Capi and Papi results in the Household Expenditure Survey by education of the head of the family, 1995.

Goods (variable code, reference time in months)	Education / Mode index: Capi/Papi * 100
---	--

Source: interviews

Charges for owner-occupied dwelling (300,1)

Rent of dwelling (301,1)

Electricity (31000,12)

Furniture (4000-, 12)

Household appliances and equipment (420,12)

Radio, television, tape recorder etc. (700,12)

Personel computers and peripherals (70600,12)

Sports equipment, hiking gear (7170, 12)

Purchase of car (600,12)

Physicians' fees (5100, 3)

Interest payments on household loans (302,12)

Amount of loan for dwelling (asuntvel, interview)

Amount of
consumption loan
(kulutvel, interview)

Source: diaries

Meat (101, ½)

Alcohol beverages
(121, ½)

Outer
garments/outdoor
clothes (200, ½)

Total consumption

* after the number p
<= 0,05

Goods

Lo
w
edu
cati
on

Hig
h
edu
cati
on

Source: interviews

Charges for owner-
occupied dwelling
(300,1)

Rent of dwelling
(301,1)

Electricity
(31000,12)

Furniture (4000-,12)

Household
appliances and
equipment (420,12)

Radio, television,
tape recorder etc.
(700,12)

Personel computers
and peripherals
(70600,12)

Sports equipment,
hiking gear (7170,
12)

Purchase of car
(600,12)

Physicians' fees
(5100, 3)

Interest payments
on household loans
(302,12)

Amount of loan for
dwelling (asuntvel,
interview)

Amount of
consumption loan
(kulutvel, interview)

Source: diaries

Meat (101, ½)

Alcohol beverages
(121,½)

Outer
garments/outdoor

clothes (200,½)

Table 4. Differences between Capi and Papi results in the Household Expenditure Survey by age of the head of the family, 1995.

Age / Mode index: Capi/Papi * 100

Goods (variable code,
reference time in
months)

Source: interviews

Charges for owner-
occupied dwelling
(300,1)

Rent of dwelling (301,1)

Electricity (31000,12)

Furniture (4000-,12)

Household appliances
and equipment (420,12)

Radio, television, tape
recorder etc. (700,12)

Personel computers
and peripherals
(70600,12)

Sports equipment,
hiking gear (7170, 12)

Purchase of car
(600,12)

Physicians' fees (5100,

3)

Interest payments on
household loans
(302,12)

Amount of loan for
dwelling (asuntvel,
interview)

Amount of
consumption loan
(kulutvel, interview)

Source: diaries

Meat (101, ½)

Alcohol beverages
(121,½)

Outer garments/outdoor
clothes (200,½)

Total consumption

* in front of the number:
n < 20 ; - n < 10

*, ** or *** after the number: T-Test significance $p \leq$
0,05, 0,01 or 0,001

References

Ahlqvist Kirsti, 1996, Kulutustutkimus 1995. Menetelmävertailun (capi-papi) esitutkimuksen raportointia (manuscript, in Finnish).

Ahola Anja, 1996, Responses to health questions with a new interview mode. In "The Role of Statisticians in Today's Information Society". The 20th Conference of Nordic Statisticians. Statistical Reports of the Nordic Countries 66. Copenhagen, 255-270.

Baker Reginald P, Bradburn Norman M and Johnson Robert A, 1995, Computer-assisted Personal Interviewing: An Experimental Evaluation of Data Quality and Cost. Journal of Official Statistics. Vol 11, No. 4, 413-431.

Dufour Johane, Simard Michelle and Mayda Frank, 1996, First year of computer-assisted interviewing for the Canadian labour force survey: an update. In Laaksonen Seppo (ed.): International perspectives on nonresponse. Proceedings of the sixth international workshop on household survey nonresponse. Research reports 219. Statistics Finland: Helsinki, 53-71.

Heiskanen Markku, 1995, Tietokoneavusteisella ja perinteisellä käyntihaastattelulla saatujen tulosten vertailu (Comparing Capi and Papi Results). In "Elinolotutkimus 1994. Aineiston keruu." Tilastokeskus Muistio 1995:2, 58-68 (in Finnish).

Laaksonen Seppo (ed.), 1996, International perspectives on nonresponse. Proceedings of the sixth international workshop on household survey nonresponse. Research reports 219. Statistics Finland: Helsinki.

de Leeuw E D, 1992, Data Quality in Mail, Telephone and Face to Face Surveys. TT-Publications: Amsterdam.

Manderbacka K, Huuhka M, Lahelma E and Rahkonen K.
1994, Haastattelijoita haastattelemassa, Hyvinvointikatsaus
4/1994, 52-54 (in Finnish).

Martin Jean, O'Muircheartaigh Colm and Curtice John,
1993, The Use of CAPI for Attitude Surveys; An
Experimental Comparison With Traditional Methods. Journal
of Official Statistics, Vol 9, No.3, 641-661.