Retrospective questioning with the aid of Blaise

by

Hille Wesseling

Netherlands Central Bureau of Statistics
Kloosterweg 1
6412 CN HEERLEN
tel. 045-707387

8-10-1993

Retrospective interviews
A new method with the aid of Blaise

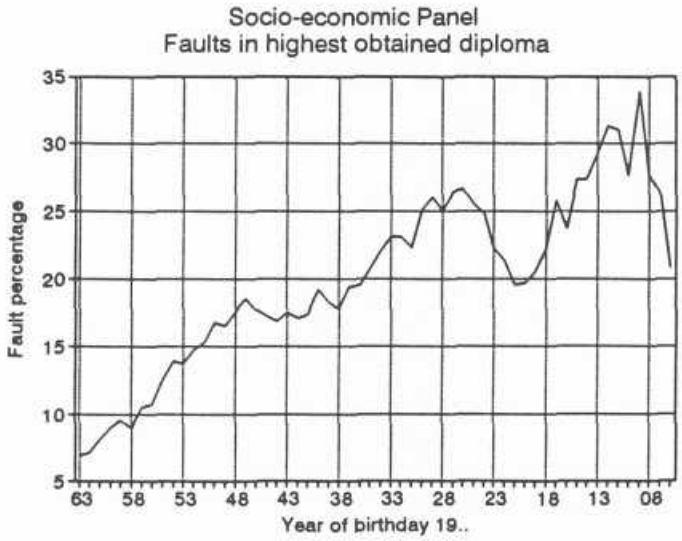
Hille Wesseling

Problem definition

It is a well-known fact that the reliability of retrospective questions can be affected negatively by memory effects. When answering questions such as: has your household composition changed, do you still work with the same company, is your occupation still the same? etc. interviewees forget to report many changes. In panel (or panel-like) surveys, respondents can be asked the complete questionnaire at different points in time, each time giving a full report of the present situation. A drawback is that changes have to be discovered by comparing the questionnaires with those of the previous waves. Moreover, differences do not necessarily imply actual changes: they may be the result of typing errors, which are discovered. Changes in variables such as for example sex and date of birth are extremely unlikely. In panel surveys these can be verified by comparing the information with answers given in previous waves, t-2, t-3. All this makes the 'after-processing' of such a panel survey very labour intensive.

The graph below illustrates how memory effects correlate with, among other things, the age of the respondent. The data come from the Socio-economic Panel and refer to waves 1, 11, and 13.
In these three waves all respondents (1740) had to answer the same standard set of questions about education. The percentage in the graph shows a decrease in the level of the highest obtained diploma. Obviously such a decrease should not be possible. The graph shows how these errors increase as the respondents are older. They start at 7% for people born in 1963 to 34% for those born in 1911. This is an average of 16.4% (285 respondents).

This paper does not intend to give the best solution to these problems. It is only meant to give subject-matter experts a technical possibility to cope better with memory problems in the field. Of course, the survey must be panel-based.



The new approach

Let us illustrate with the aid of a CBS survey how memory effects in the field can be prevented.
The **Family Expenditure Survey** (FES) is partly a CAPI survey. Respondents participate for one year and have to complete various questionnaires.
The CAPI questionnaires cover 'general information', 'meter readings' (gas, electricity, water) and a 'retro' questionnaire. For the 'income' and 'periodic expenditure' questionnaires, CAPI versions have been developed which will be used next year. The survey begins with the 'general information' questionnaire in januari. At the end of the year respondents have to answer a retro questionnaire which reports changes in household composition, company, occupation and social position in the preceding year. To this end a CAPI questionnaire has been constructed which uses an external file containing the data of the previous 'general information' questionnaire. Previously, interviewers had been asked to report changes in the paper 'general information' questionnaire. They then only completed the part of the questionnaire to which the change referred.
We had the impression changes were often forgotten with this method.
In the new - CAPI - approach, respondents are confronted with the old data. The interviewers read out from the screen which answers the respondent gave the last time and then ask whether the situation has changed. Respondents' memories are jogged as it were.

The retro questions

The CBS uses a standard household questionnaire in which the composition of the household is established. For each member of the household we ask: position in household (husband, wife, child, etc.), sex, date of birth and marital status. In the new approach it is not necessary to type in all the data again. It is only necessary to check whether the household has any new members or members who have left the household, or whether the position in the household or the marital status have changed. This meant the household block had to be adapted. The interviewer is now asked to code whether a person is still present, or has left. Also, of course, new members can be reported.
For the other parts of the questionnaire (social position, company and occupation) the blocks are preceded by questions describing the previous situation. All retro questions are asked in one block. Annex III contains the retro block in Blaise-language.
For social position there are two questions, as this part also includes a block on 'main occupation'.

An example of a 'retro' question is:

Last time
- you considered yourself to be a pupil/student
- you worked for less than 12 hours a week

Is this still the case?

1. **Yes**
2. **No**
3. **Previous information is incorrect.**

If the respondent answers 'yes', the whole block of the part concerned can be skipped. If the answer is 'no' or 'previous information is incorrect', the whole block is asked again.

Two different techniques are used to impute texts in retro questions The first is basing a text on a code which is read from the external file. The second is reading a text directly from the external file. The former technique is used for text imputations of social position. As mentioned above, this retro part consists of two questions and each question consists of two, code-based, text imputations. The following options are code-based.

Last time
- *you considered yourself as...*
  - *...a salaried employee*
  - *...self-employed*
  - *...assistant to self-employed*
  - *...a conscript*
  - *...housewife/husband*
  - *...volunteer*
  - *...a pupil/student*
  - *...a pensioner*
- *you*
  - *...worked for 12 hours or more a week*
  - *...worked for less than 12 hours a week*
  - *...had no paid work*

The second question concerns main occupation i.e. what the respondent spent most time on. The interviewer continues reading directly after the answer to the first question.

- *you spent most time on...*
  - *...paid work*
  - *...housework*
  - *...studying*
  - *...voluntary work*
- *and in addition you consider yourself as*
  - *...unemployed*
  - *...disabled*
  - *...unemployed and disabled*

Depending on the situation, text imputations are left out. The text imputations of the other parts 'company' and 'occupation' need not be derived but can be read directly as a string from the external file.

An example of a retro question for the company block:

You worked at:... Schiphol

Do you still work there?

1. **Yes**
2. **No**
3. **Previous information is incorrect.**

And an example for the occupation block:

Your occupation was: ... electrical engineer

Is this still your occupation?

1. **Yes**
2. **No**
3. **Previous information is incorrect.**

For the FES it is important to know when a certain change has taken place. Indeed, the Blaise questionnaire asks for dates of changes. With the aid of these dates it is possible to follow all persons in a household for a whole year. Exceptions are changes which may occur due to seasonal labour. If the surveys are conducted only in January and December, a change in the variable 'social position' may be missed as the respondent only worked in the summer months. The method as used in the FES reports only the starting and the end situation. Naturally, variations are possible. It is possible to ask historical questions from a certain date. However, this involves the retro problem again. Another method is to do an extra survey halfway through the year.
From a methodological point of view there is room for improvement in question formulation and answer categories. The two retro questions on social position are especially eligible for this. Next year - 1994 - each text imputation will be incorporated in a separate question and no longer combined in one question, as is now the case (one question actually contains two questions). This is a methodological problem which we shall not go into here. The analysis of the retro questionnaire at the CBS has not yet been concluded.

The external file

We used Manipula 1.5 to make the external file.
With the aid of the specifications of the 'general information' questionnaire Blaise generates a Manipula set-up which is suitable to read from Blaise files. In the set-up the variables are derived which ensure that the correct imputations appear on the screen of the retro-questionnaire.
We adapted the set-up by making use of record type. This makes the set-up considerably shorter and thus more convenient to use (See annex II).
By running the Manipula set-up on the Blaise files of the 'general information' questionnaire, we get an ASCII file. With the Blaise tools option, this ASCII file can be made a Blaise external file (see annex III, external file in Blaise).
The main variables in this file are:

identification.....	text	[14]
sex.....	complete	[1]
initials.....	text	[5]
date of birth.....	text	[8]
codes for social position.....	6x complete	[1]
description of company.....	text	[40]
description of occupation.....	text	[40]

The variables sex, initials and date of birth are repeated in a loop of 1-10. This is the maximum number of persons to be included in a household roster. The variables social position, description of company and description of occupation are repeated in a loop from 1-6 (the maximum number of personal questionnaires). As mentioned above, the codes for social position in the Blaise questionnaire are converted to text which is suitable for imputation in the questions. The remaining variables are used directly from the external file.

An economical method

A great advantage of this approach is the reduction in interviewing time. In most cases (nearly 90%) the blocks can be skipped. Asking the whole questionnaire would have taken twice as much time. The 'general information' questionnaire took about 10 minutes, the retro questionnaire 5 minutes. The data processing after the fieldwork also is less labour intensive. There is only a small portion of company and occupation blocks to be coded. For each respondent unchanged blocks can be copied from the previous survey to a new file. Changed blocks thou can be added to this file. The following table shows the number of completed and skipped blocks in the retro questionnaire of the FES 1992.

	Total	no changes	changes
Are there any changes in the total questionnaire	1994	1073 (53.3%)	921 (46.7%)
Per block			
Household box	1994	1761 (88.3%)	233 (11.7%)
Social position	3670	3283 (89.5%)	387 (10.5%)
Company	1747	1561 (89.4%)	186 (10.6%)
Occupation	1596	1413 (88.5%)	183 (11.5%)

Changes are incl. category 'previous information incorrect' (approx. 0.4% per block).

Distribution to interviewers

Fortunately we have a good CAPI infrastructure at our disposal at the CBS. The LIPS-SPIL system is an important part of this. LIPS-SPIL is a data communication system which enables interviewers to read the address data (municipality, street, house number etc.) into their handheld computers via a modem. Besides it's possible to send external information like previous data linked to the addresses.
In the FES 2236 addresses and the same number of records for the external file (text imputations) are distributed to 275 interviewers.

Evaluation

As the first results of skipped and completed blocks were revealed, it turned out to be that the questionnaire had not been tested sufficiently. Due to an error, an incorrect text imputation appeared on the screen in the second retro question concerning the 'social position'. Instead of 'student' the screen showed 'volunteer' and vice versa. In 30% of the cases the interviewer typed in 'unchanged'. The interviewer probably thought that in an unchanged situation the retro question should be answered with 'yes', regardless of what is shown on the screen. In principle, obviously, the answer should have been 'previous information incorrect'. We recommend adjusted instructions for the interviewers because some of them are not used to the changing texts in the type of retro questions involved.

From a Blaise-technical point of view this experiment has been a success. Although the analysis of the retro questionnaire has not yet been concluded, we have the impression that the data are plausible. The field work has not given rise to any insurmountable problems either. This was enough to make the FES-group decide to use an improved retro questionnaire again for the FES next year.

MANIPULA SETUP for external file retro questionnaire (EXAMPLE RECORDTYPE)

```
INPUTFILE "GENERAL" BLAISE
  FIELDS
  EA_Ident      5    8 string    0
  Spil_ID      13   14 string    0
     -          -    -   -       -
     -          -    -   -       -
     -          -    -   -       -
  RespoHHK    102    1 integer   0

  Pers 5 array[1..8] of
      {5 is starting position}
    record
      EVP2          integer[ 1]
      Kind          integer[ 1]
      Ander         integer[ 2]
      PLHH          integer[ 2]
      M_V           integer[ 1]
      Geboren       string [11]
      Leeftij       integer[ 3]
      MissDg0       integer[ 1]
      Sch           integer[ 3]
      Akk           integer[ 1]
      BurgSt        integer[ 1]
    endrecord (1,2,3,4,5,6,7,8)
              {Intervalnumbers}

  CtrlPers    103    1 integer   0
  Kostw       104    1 integer   0
  HuwJr       105    2 integer   0
     -          -    -   -       -
     -          -    -   -       -
  WlkZiek7    437   40 string    0

  Arb 5 array[1..6] of
    record
      Bo_nr1        string [ 6]
      OP0           integer[ 1]
      Select0       integer[ 1]
        -               -
        -               -
      Diplom05      integer[ 1]
      Klas306       integer[ 1]
      ExamUn05      integer[ 1]
    endrecord ( 9,10,11,12,13,14)

  Huishoud    477    1 string    0
  EigenZaa    478    1 integer   0
  ProdEigZ    479    1 integer   0
     -          -    -   -       -
     -          -    -   -       -
  EINDEVrg    854    1 string    0
  StatInt     855    1 integer   0
  DummCOBS    856  100 string    0
```

```
OUTPUTFILE "LIPS.TXT"
  SETTINGS
    InitRecord=Yes
  FIELDS
  EA_Ident      1    8 string
  Spil_ID       9   14 string
  ENQNr        23    5 integer
  GemCode      28    4 integer
  Stuur        32    1 integer
  BO_nr0       33    6 string
  AantalPP     39    2 integer

  Pers array[1..8] of
    record
      Pers_Id       string [14]
      M_V           integer[ 1]
      Init          string [ 3]
      Gebdat        string [ 8]
    endrecord;

  Bzh array[1..6] of
    record
      DplMil        integer[ 1]
      WrklArb0      integer[ 1]
      WrklArb1      integer[ 1]
        .
        .
        .
      Bedrijf       string [40]
      Beroep        string [40]
    endrecord


VAR
  i,Nr : integer
```

```
MANIPULATE
  compute stuur:=3

  for i:=1 to 8 do
    if i<=AantalPP then
      compute Pers[i].Pers_Id:=str(i);
      {*** birthday from 11 to 8 positions ***}
      compute Pers[i].Gebdat:=substring(Pers[i].Geboren,1,2)+
                              substring(Pers[i].Geboren,4,2)+
                              substring(Pers[i].Geboren,7,4)
    endif
  enddo

  for i:=1 to 6 do
    if i<=AantalPP then
       .
       .
       .
      {Deriving social position}
       .
       .
      {Company}
      if Arb[i].SoortB_W>'' then
        compute Bzh[i].Bedrijf:=Arb[i].SoortB_W
      else
        compute Bzh[i].Bedrijf:=Arb[i].SoortB_Z
      endif

      {Profession}
      if Arb[i].beroepW>'' then
        compute Bzh[i].Beroep:=Arb[i].BeroepW
      else
        compute Bzh[i].Beroep:=Arb[i].BeroepO
      endif
       .
       .
       .
    endif
  enddo

  SELECT

ENDMANIPULATE
```

BLAISE PARAGRAPH 'EXTERN' OF THE RETRO QUESTIONNAIRE

```
EXTERN

  TYPE
    LipsRec = RECORD (length 800)
                ID_nr    = string [14];
                Aant_reg = integer[ 2];
                Pers     = ARRAY[1..8] of
                           RECORD
                             Pers_Id  = string [14];
                             Gesl     = string [ 1];
                             init     = string [ 3];
                             GebDag   = string [ 2];
                             GebMnd   = string [ 2];
                             GebJr    = string [ 4];
                           ENDRECORD;
                Arb      = ARRAY[1..6] of
                           RECORD
                             DplMil   = integer[ 1];
                             WrklArb0 = integer[ 1];
                             WrklArb1 = integer[ 1];
                             TydbestO = integer[ 1];
                             TotUr12  = integer[ 1];
                             BdrfInst = integer[ 1];
                             BedrPrak = integer[ 1];
                             BinnenHH = integer[ 1];
                             Personel = integer[ 1];
                             GrRoute  = integer[ 1];
                             Bezig    = integer[ 1];
                             Bedrijf  = string [40];
                             Beroep   = string [40];
                           ENDRECORD;
                Dummy[30];
              ENDRECORD;


Var
  LipsDat: DataFile  "Lips.dat  " of Lipsrec;
  LipsR  : LipsRec;
  LipsIdx: IndexFile "Lips.idx  " [14];

EndExtern;
```

```
BLOCK RetroB;
VAR
  m            : integer;
  t1,t2,t3,t4 : string[80];
  ASocPos      : array[1..6] of string[50];
  AWorkTm      : array[1..3] of string[50];
  ATSpend      : array[1..4] of string[50];
  Company, Profes : String[50];

QUESTIONS
  PosThen1 "/ - /
  @/ The last time:
  @/ $t1$t2
@/@/ Is that still the case?
  @/" : (Yes, No, onjuist "data is wrong");

  DdThen1  "/ -/
  @/ When has this situation changed?
  @/" : DateType;

  PosThen2 "/ - /
  @/
  @/ $t3$t4
@/@/ Is that still the case?
  @/" : (Yes, No, onjuist "data is wrong");

  DdThen2  "/ -/
  @/ When has this situation changed?
  @/" : DateType;

  BDThen "/ - /
  @/ You worked at: $Company
@/@/ are you still working there?
  @/" : (Yes, No, onjuist "data is wrong");

  BRThen "/ - /
  @/ Your profession was: $Profes
@/@/ Do you still have this profession?
  @/" : (Yes, No, onjuist "data is wrong");

ROUTE
  compute ASocPos[1]:='working';
  compute ASocPos[2]:='liable to military service';
  compute ASocPos[3]:='house';
  compute ASocPos[4]:='volunteer';
  compute ASocPos[5]:='pupil/student';
  compute ASocPos[6]:='pensioned';

  compute AWorkTm[1]:='You worked 12 hours or more in the week';
  compute AWorkTm[2]:='You worked less then 12 hours in de week';
  compute AWorkTm[3]:='You did not have a paid job';
```

```
compute ATSpend[1]:='a paid job';
compute ATSpend[2]:='housekeeping';
compute ATSpend[3]:='studying';
compute ATSpend[4]:='volunteers work';

compute t1:='';
compute t2:='';
compute t3:='';
compute t4:='';

ReadFile (LipsIdx, ID.Ident, LipsDat, LipsR, Result);
if LipsR.Arb[OPInPV].Bezig<>0 then
  if LipsR.Arb[OPInPV].Bezig=1 then
    if LipsR.Arb[OPInPV].GrRoute=1 then
      compute t1:='an employed person';
    elseif LipsR.Arb[OPInPV].GrRoute=2 then
      compute t1:='a self employed person';
    elseif LipsR.Arb[OPInPV].GrRoute=3 then
      compute t1:='someone helping in the company of the partner';
    endif;
  else
    compute t1:=ASocPos[LipsR.Arb[OPInPV].Bezig];
    if LipsR.Arb[OPInPV].Bezig=3 then
      if GslPers[OPInPV]=1 then compute t1:=t1+'man'
      else compute t1:=t1+'wife' endif;
    endif;
  endif;
endif;
if LipsR.Arb[OPInPV].TotUrl2=0 then
  compute t2:='@/- '+AWorkTm[3];
else
  compute t2:='@/- '+AWorkTm[LipsR.Arb[OPInPV].TotUrl2];
endif;
if (LipsR.Arb[OPInPV].Tydbest0<>0) en
    (LipsR.Arb[OPInPV].Tydbest0<>LipsR.Arb[OPInPV].Bezig) then
  compute t3:=ATSpend[LipsR.Arb[OPInPV].Tydbest0]
endif;

if LipsR.Arb[OPInPV].WrklArb0=1 then
  compute t4:='unemployed';
  if LipsR.Arb[OPInPV].WrklArb1=2 then
    compute t4:=t4+'and unfit for work'
  endif;
elseif LipsR.Arb[OPInPV].WrklArb0=2 then
  compute t4:='unfit for work';
  if LipsR.Arb[OPInPV].WrklArb1=2 then
    compute t4:=t4+'and unemployed'
  endif;
endif;
if t1<>'' then
  compute t1:='@/- you consider yourself as '+t1;
endif;
```

```
    if t3<>'' then
      compute t3:='@/- most of the time you spend on '+t3;
    endif;
    if t4<>'' then
      compute t4:='@/- and beside that you consider yourself as '+t4;
    endif;

    PosThen1;
    if PosThen1=No then
      DdThen1
    endif;
    if not((t3='') en (t4='')) then
      PosThen2;
      if (PosThen1=Yes) en (PosThen2=No) then
        DdThen2
      endif;
    endif;
    if LipsR.Arb[OPInPV].TotUr12=1 then
      if LipsR.Arb[OPInPV].Company<>'' then
        compute Company:=LipsR.Arb[OPInPV].Company;
        BDThen;
      endif;
      if LipsR.Arb[OPInPV].Profes <>'' then
        compute Profes:=LipsR.Arb[OPInPV].Profes;
        BRThen
      endif;
    endif;

  ENDBLOCK;
  VRAGEN
    RetroV : RetroB;
```