Improving quality and containing respondents' burden in socio-demographic surveys: the case of intra-household relationships

Carolina Facioni, Luciano Fanfoni, Martina Lo Conte, Stefania Macchia, Paolo Piergentili, Luciana Quattrociocchi, Marco Scuderi [1]

Abstract

The paper shows the main results of an experimental survey carried out by the Italian National Institute of Statistics - Istat. The aim was testing a new approach for collecting data on socio-demographic variables, particularly on household relationships, managed by a household grid, as proposed by Eurostat.
Cognitive interviews were used to investigate on possible critical issues and on respondents' perception in terms of burden. Following the attempt to modernise social surveys and Population Census, as well as the recent regulation changes introducing same-sex civil unions in Italy (same-sex couples), this new approach meets the dual need to contain the respondents' burden, and to offer higher quality statistical information. The tested approach completely reverses the traditional way of collecting intra-household relationships: this information is no longer asked in relation to a unique Reference Person, but for each family member in relation to all the others. The results of the test confirmed the feasibility of adopting this new approach for all socio-demographic surveys. As a matter of fact, even if some aspects need to be improved, it was showed that it contributes to reduce the ambiguity in identifying family nuclei, without having a relevant effect on the interview length and with no increase of the perceived respondents' burden.

Keywords: Cognitive interviewing, family relationships, household grid, socio-demographic surveys.

1   Carolina Facioni (facioni@istat.it); Luciano Fanfoni (fanfoni@istat.it); Martina Lo Conte (loconte@istat.it); Stefania Macchia (macchia@istat.it); Paolo Piergentili (piergentili@istat.it); Luciana Quattrociocchi (quattroc@istat.it); Marco Scuderi (scuderi@istat.it), Italian National Institute of Statistics - Istat.

1. Introduction[2]

Socio-demographic surveys play a leading role in documenting relevant family transformations taking place in Italy. Policy makers need analysis on family relationships when planning and implementing economic and social policy interventions. The nature of these relationships can be jurisprudential, as well as economic, biological, social, emotional, or simply residential, and can include one or more of these aspects. In order to satisfy this fact-finding need, the Italian National Institute of Statistics – Istat carries out a wide range of socio-demographic surveys, the largest and best known being the Population Census.

Data collection on family composition generally precedes the sections of the questionnaires dealing with the survey subject. These opening questions represent therefore an extremely delicate moment, both for winning the possible fear of invasion of privacy and for guaranteeing the reliability of the information collected.

The current official socio-demographic surveys, being designed at different times and with different needs, are characterised by a lack of harmonisation in the way of collecting data and in the classifications used for the kinship variable, thus creating great difficulty for anyone wanting to use, analyse and compare data on family composition from these surveys.

Furthermore, up to now in socio-demographic surveys, family relationships have been observed with respect to a Reference Person (RP), usually an adult family member or the holder of the family form in the Demographic Registers. Despite the 23 categories of the relationship classification (RP23), identifying all bilateral relationships between family members is not always possible when families have more than one nucleus. Some family structures may therefore be incorrectly registered. In Italy, for Eurostat regulated surveys, this has brought about a need to use additional questions in some surveys, such as the Labour Force Survey (LFS), to retrace all family types required on a supra-national level.

Following the attempt of modernise social surveys and Population Census, as well as the recent regulation changes introducing same-sex civil unions in Italy (same-sex couples), Istat has given a high priority to the standardisation of variables concerning kinship and family composition for all surveys integrated in the Master Sample[3]. In this context, it was thought to experiment the introduction of the Household Grid (HHG) proposed by Eurostat (European Commission, 2009). Such approach for collecting data on intra-household relationships is very different from the traditional one: information for all family members is no longer asked in relation to a RP, but for each single component in relation to all the others. This allows to use a more agile classification (with only 16 categories instead of 23) and also to simplify item descriptions, as reported below, in order to eliminate possible ambiguities impacting on the family structure identification. The Household Grid approach can respond, therefore, to the needs not only to facilitate survey harmonisation, but also to contain respondents' burden and to produce a high quality statistical information.

The aim of the test presented here was to assess the feasibility of this new approach for the Italian socio-demographic surveys. This paper describes the methods adopted (Section 2) and the main results of the experiment, evaluated, as to the respondents' burden, in terms of interview length and of easiness of compilation, and, from the quality perspective, in terms of the correct identification of household structures (Section 3). Finally, Section 4 summarises some conclusions and perspectives concerning the feasibility of the adoption of this new approach.

---

3   Master Sample integrates a set of balanced and coordinated sample surveys (Continuous census, Labour force survey, Aspects of daily life survey, Eu-Silc) in the context of the Population Census and Social Surveys Integrated System (CSSIS).

2. Methods

The test was based on a small experimental survey aimed at collecting, through the new approach, the main socio-demographic information of each household member and all the intra-household kinship relations (variables usually collected in social surveys within the 'General Household Information section').

The interviews were followed by retrospective cognitive questions to investigate on possible critical issues and on respondents' perception in terms of burden. Cognitive interviewing has emerged as one of the most prominent methods for identifying and correcting problems with survey questions. This method is used to investigate the response process, so as to determine whether the questions are generating the information that the authors intend (Sudman *et al.,* 1996). Such interviews could consist of respondents' elaborations regarding how they constructed their answers, explanations on what they interpret the question to mean, reports on any difficulties they had answering and anything else that sheds light on the circumstances that their answers were based upon (Beatty and Willis, 2007). Interviewers are generally involved: they can have a minimal intervening role in a 'thinking aloud process', or can interact asking direct questions according to a scripted protocol. This second type of practice was applied, as it is detailed in Paragraph 2.2.

Three data collection techniques were used in the experiment: PAPI (Paper and Pencil Interviewing) administered by an interviewer, CATI (Computer Assisted Telephone Interviewing) and CAWI (Computer Assisted Web Interviewing), in order to replicate the strategies adopted for the main Italian socio-demographic surveys.

Also for the self-administered interviews (CAWI), the presence of a non-participating observer, documenting uncertainties or problems in answering, was considered useful. Therefore, CAWI respondents were asked to fill in the questionnaire in a centralised location, the Istat CAI (Computer Assisted Interviewing) laboratory at the presence of an interviewer.

Generally, cognitive interviewing samples are not designed to be representative of any population. However, it is recommended to take into account demographic variety of respondents and to include people relevant to the topic of the questionnaire being tested (Willis, 1994, 2005). In other

words, whatever topic the questions focus on, the sample should cover a variety of situations relevant to that topic (Beatty and Willis, 2007).

For the sample selection, therefore, two constraints were taken into account: firstly, to cover different types of households mainly in terms of household size; secondly to contain costs in terms of time and resources.

Given these restrictions, the test was run completely in-house, interviewing Istat colleagues characterised by different household types. To avoid bias due to the working experience on this topic, only employees working in the General Directorate and in the IT Department – and therefore not having experience on socio-demographic surveys – were considered.

The sample selection was based on household size, gender, age and education: 129 individuals were extracted from the total employees considered (338), with the aim to obtain about 100 interviews, homogeneously divided among the three techniques.

2.1 The questionnaire design

The questionnaire consisted of three sections: the first one, about the number of household members, was followed by a section collecting some demographic information on each member, such as gender, date of birth, marital status, *etc.* (General Household Information, GHI). These variables are a subset of those considered in the GHI section of most socio-demographic surveys. The choice of such variables was made on the assumption that moving some questions (for instance, place of birth, citizenship or labour status) in the individual questionnaires could be more efficient.

The third and final section concerned the intra-household relationships, for each single member in relation to all the others, collected through the household grid (HHG), which was the core of the test.

Figures 1 and 2 show the paper version of the GHI and the HHG sections.

Figure 1 - General Household Information Section (GHI)

Order number of household members	A First Name *Write in capital letters*	B Surname *Write in capital letters*	D Gender 1=M 2=F	E Date of birth dd/mm/yyyy	F Age in completed years *(to be filled in if you do not remember the date of birth)*	G Marital status *(only if age >14 years)*	H Year of marriage or of civil union *(only if married or in civil union)*	I Marital status before last marriage /civil union *(only if married or in civil union)*
********	****** ***** ********	*** ****** *******	*** ********	**** ****** ********	*****************	******** ****	*** ******* ********	******* ****** ******
1								
2								
3								
4								
5								
6								
7								
8								
9								
10								

Source: Istat, Questionnaire of the 2018 Experimental survey on household grid

Figure 2 - Household grid with intra-household relationships (HHG)

Order number of household members	Name or initials *Write in capital letters*	Name or initials									
**********	******************	1	2	3	4	5	6	7	8	9	10
1			I_I_I	I_I_I	I_I_I	I_I_I	I_I_I	I_I_I	I_I_I	I_I_I	I_I_I
2				I_I_I	I_I_I	I_I_I	I_I_I	I_I_I	I_I_I	I_I_I	I_I_I
3					I_I_I	I_I_I	I_I_I	I_I_I	I_I_I	I_I_I	I_I_I
4						I_I_I	I_I_I	I_I_I	I_I_I	I_I_I	I_I_I
5							I_I_I	I_I_I	I_I_I	I_I_I	I_I_I
6								I_I_I	I_I_I	I_I_I	I_I_I
7									I_I_I	I_I_I	I_I_I
8										I_I_I	I_I_I
9											I_I_I
10											
		2									

Source: Istat, Questionnaire of the 2018 Experimental survey on household grid

In practice, in the household grid only the colored information needs to be asked (Figure 2); once the components' names are known, the responses have to be written row by row: Who is *'component 1'* in relation to *'component 2'*?, Who is *'component 1'* in relation to *'component 3*?', and so on.

In this way, even though the number of questions is higher compared to the traditional approach, the respondent's burden is supposed to be lower, since the number of items is smaller respect to the previous RP23 classification and the corresponding descriptions are simpler (Figure 3).

Figure 3 - Intra-household relationships items used in the test

INTRA-HOUSEHOLD RELATIONSHIPS	
01 Husband/Wife	09 Stepbrother/Sister (with both different parents)
02 Partner in civil union	10 Son/Daughter-in-law
03 Partner/Cohabitée	11 Brother/Sister-in-law
04 Son/Daughter	12 Father/Mother-in-law
05 Stepson/daughter	13 Grandparent
06 Parent	14 Grandchild
07 Stepparent	15 Other relative (not included in the list)
08 Brother/Sister	16 Other non-relative

Source: Istat, Questionnaire of the 2018 Experimental survey on household grid

Using the CATI and CAWI electronic questionnaire, the interview becomes even easier, since each information requested corresponds to a question and the grid is transformed into a series of questions customised with each member's name (Figure 4).

Figure 4 - Series of questions for intra-household relationships in electronic questionnaire

Source: Istat, Questionnaire of the 2018 Experimental survey on household grid

2.1.1 The electronic questionnaire

The IT tool used for the implementation of the electronic questionnaire is LimeSurvey, an open source software. The same electronic questionnaire was used for all the techniques (CAWI self-administered, CATI for telephone interviews and CADE, Computer Assisted Data Entry, for the PAPI questionnaires). It was structured in three sections:

- the first one asks for the number of 'household members', after providing the necessary definitions;

- the second one regards the loop of questions to collect the information of the GHI section;

- the last one contains the series of questions reproducing the household grid. All the questions are customised with the components' names already collected in the previous section.

A set of rules have been implemented to prevent from non-response and consistency errors, similar to those used in socio-demographic surveys that adopt computer-assisted techniques. In details, consistency rules have been managed with two different approaches:

- for the consistency between HHG relationships and the information collected in the GHI section, the list of intra-household relationships displayed for each component was subject to the answers given in the GHI section. The following examples may help clarify:

  ✓ if the age difference between the two persons involved in the relationship is less than 14, the item 'mother/father' was not displayed;

  ✓ if two persons have not declared to be married in GHI Section, the item 'wife/husband' was not displayed.

- for the consistency among other data regarding different components, error messages were displayed after giving the 'inconsistent' response (for example, in case two members declared to be married with the same person, or when members with the same gender (same-sex couples) declared to be 'mother/father' of the same person).

2.2 The cognitive test

After the interview, a set of cognitive questions were asked by interviewers. A probing-based paradigm was used, instead of the think-aloud one.

In general, both paradigms aim at generating verbal information that is usually unseen in a survey interview, in order to evaluate how well the questions are meeting their objectives (Beatty and Willis, 2007). Literature shows that there are advantages and disadvantages for both approaches. On one hand, an interviewer probing could introduce bias into the data collection process and can create artificiality (Conrad, Blair and Tracy, 2000). On the other hand, probing does not interfere with the actual process of responding (since it intervenes after the questionnaire but still capturing information stored in short-term memory), while thinking aloud might, since participants have to provide verbal information during the response process. This can also increase the effort spent on creating a response which has an unknown impact on the real answer (Willis 1994). However, the probing, when used, should involve only a few questions per interview (Oksenberg, Cannell and Kalton, 1991).

Another important factor to be decided in a cognitive test is whether it should be standardised or determined by the interviewer judgement, and to what extent (Willis, 2005; Presser *et al.*, 2004).

The cognitive questionnaire used for the test was standardised and administered by the interviewer. It only had a few questions, starting with the easiness/difficulty in identifying the pertaining intra-household kinships and, in case of difficulty, for which relationship this had been encountered. Moreover, it was asked whether it had been necessary to read again or ask the interviewer to repeat some concepts and, if so, which ones. A deepening was carried out to check whether the options 'Other relatives' and 'Other non-relatives' had been selected correctly. Finally, the respondents were asked to assign a score from 1 to 10 to judge easiness/directness in giving the answer, with the aim of getting an overall feedback on the household grid approach. Lastly, respondents' suggestions/proposals were also recorded.

3. Results

3.1 The survey results

The test was run between May 29 and June 5, 2018. A first letter was sent by email to 338 Istat employees of the General Directorate and the IT Department in order to illustrate the purpose of the experiment and to inform that a sample of them would be selected. Then, the 129 sampled colleagues were contacted by telephone to fix an appointment for the interview and to let them know the data collection mode they would be interviewed with.

Participation rates were very high, due to a good cooperation of the colleagues. The total response rate was 86%: CAWI interviews registered the most positive result with almost 91% of complete interviews, while the lowest rate was found for the PAPI mode (77.5 %) (Table 1).

**Table 1 - Contacts results by data collection technique**

CONTACTS RESULTS	Data collection technique			Total
	CATI	CAWI	PAPI	
Complete	41	39	31	111
Refusal	2	2	3	7
No answer	3	1	5	9
Appointment	0	1	1	2
Total	46	43	40	129
RATES				
Response	89.1	90.7	77.5	86.0
Refusal	4.3	4.7	7.5	5.4
Other no answer	6.5	4.7	15.0	8.5

Source: Istat, Experimental survey on household grid - 2018

Respondents were equally distributed by gender, with a slight prevalence of males (47% of women and 53% of men). Almost half of respondents (45%) were aged between 40 and 49, about one out of five was over 60 and only 7% under 40 (Figures 5 and 6).
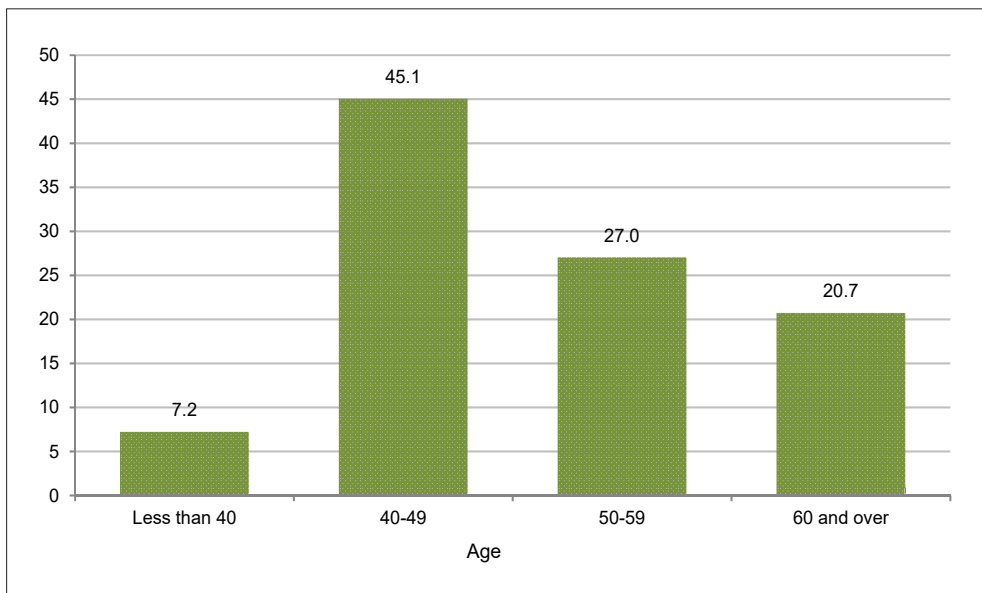
As for marital status, the prevalence of respondents was married (with 64%), followed by 14.4% singles (Figure 7). None of the interviewees was 'in civil union' or 'previously in civil union'.

Figure 5 - Respondents by gender

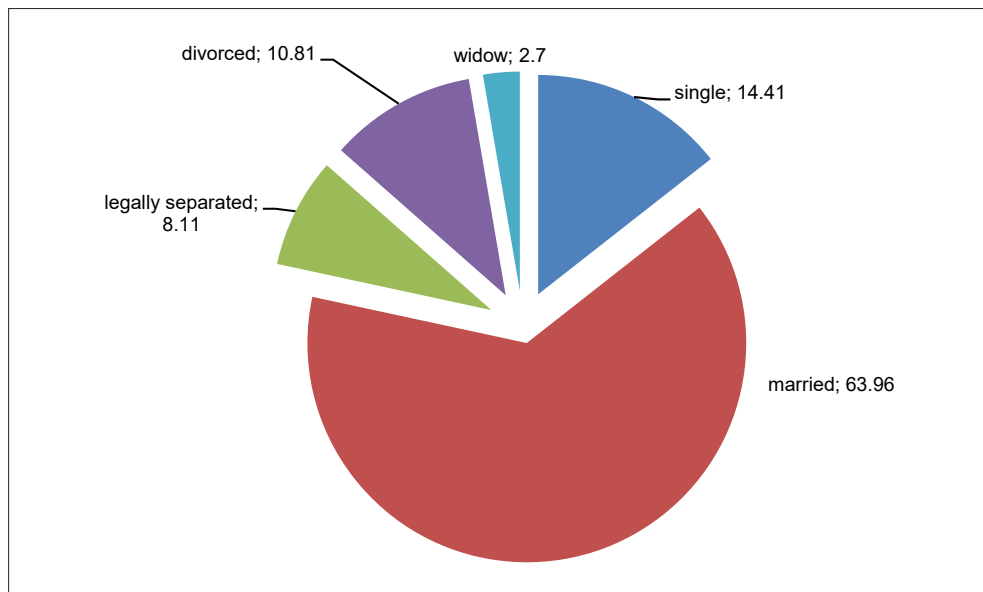Females; 46.9

Males; 53.1
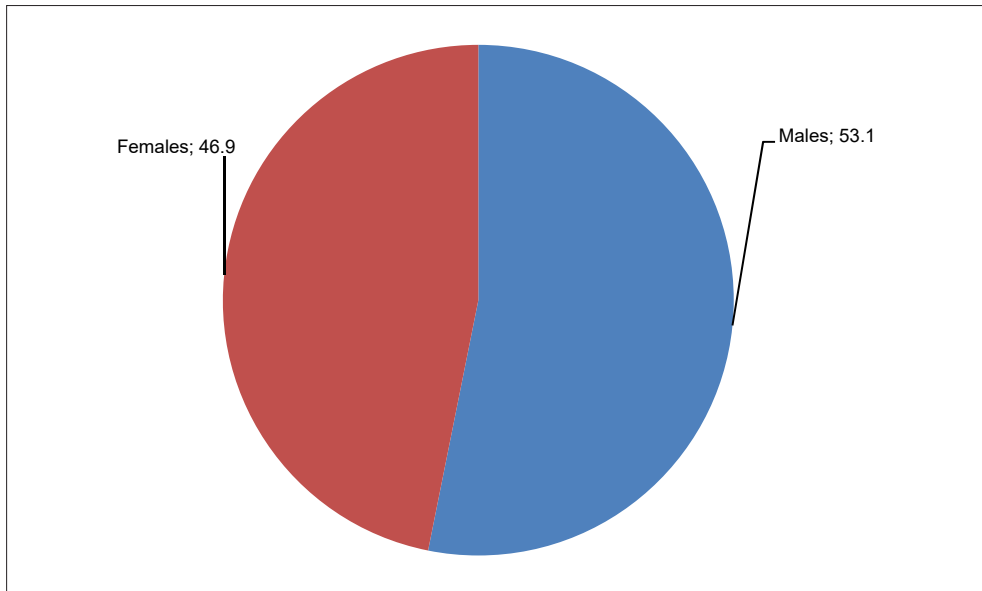
Source: Istat, Experimental survey on household grid - 2018

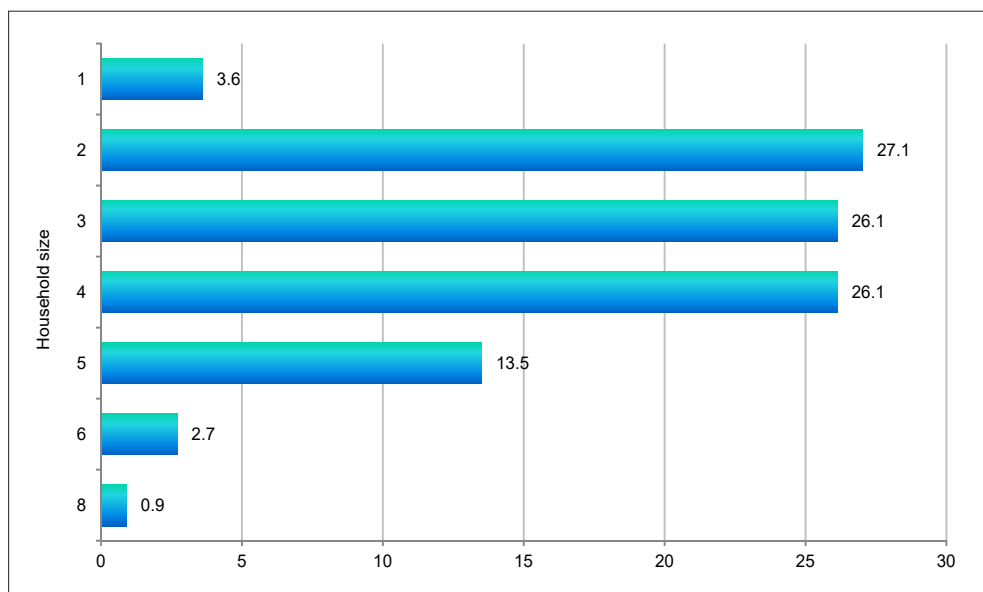**Figure 6 - Respondents by age**

Source: Istat, Experimental survey on household grid - 2018

Figure 7 - Respondents by marital status

Source: Istat, Experimental survey on household grid - 2018

Figure 8 - Respondents by household size

Source: Istat, Experimental survey on household grid - 2018

Table 2 - Intra-household relationships (number of times that items were selected)

Intra-household relationship	N.	%
01 Husband/Wife	75	14.6
02 Partner in civil union	0	0.0
03 Partner/Cohabitée	20	3.9
04 Son/Daughter	18	3.5
05 Stepson/daughter	8	1.6
06 Parent	254	49.3
07 Stepparent	8	1.6
08 Brother/Sister	92	17.9
09 Stepbrother/Sister (with both different parents)	0	0.0
10 Son/Daughter-in-law	10	1.9
11 Brother/Sister-in-law	5	1.0
12 Father/Mother-in-law	5	1.0
13 Grandparent	0	0.0
14 Grandchild	6	1.2
15 Other relative (not included in the list)	8	1.6
16 Other non-relative	6	1.2
TOTAL	**515**	**100.0**

Source: Istat, Experimental survey on household grid - 2018

The interview length represents a direct measure of response burden (Bradburn, 1978; Frankel and Sharp, 1981; Sharp and Frankel, 1983) and it is therefore a good indicator of the applicability of the HHG approach.

Table 3 shows the average interview length, distinguishing for the 3 sections only for CATI and CAWI (recorded by Limesurvey, which allows automatic registration of the compilation times).

Table 3 - Interview length per technique and per section (minutes and seconds)

CATI					
Interview lenght	N.	Mean	Dev std	Min	Max
TOTAL	41	4.10	2.17	0.50	10.23
Section N. of components	41	0.35	0.36	0.8	3.37
GHI Section	41	2.34	1.10	0.37	5.31
HHG Section	41	1.1	1.6	0.0	5.42
CAWI					
Interview length	N.	Mean	Dev std	Min	Max
TOTAL	**39**	**5.11**	**2.56**	**1.32**	**15.59**
Section N. of components	39	0.39	0.26	0.11	1.58
GHI Section	39	3.8	1.52	0.50	9.43
HHG Section	39	1.25	1.8	0.12	5.34
PAPI					
Interview length	N	Mean	Dev std	Min	Max
TOTAL	31	5.16	1.40	2.0	9.0
Section N. of components	31	-	-	-	-
GHI Section	31	-	-	-	-
HHG Section	31	-	-	-	-

Source: Istat, Experimental survey on household grid - 2018

Overall, the compilation took about 4-5 minutes. The shortest interview was obtained through the CATI mode (on average 4'10''), followed by CAWI (5' 11'') and PAPI (5' 16'').

Using CATI, the first section (number of members) lasted on average 35'', the GHI Section 2'34''and completing the household grid involved about 1 minute. On the other hand, the CAWI self-compilation took on average 3'8'' for the GHI section and 1'25'' for the household grid.

Note that the compilation time for the first section depends mainly on the understanding of the explanations provided for concepts such as the household and who to include as members.

A factor that greatly affected the interview length was the household size. Clearly, the compilation time grew along with the number of members, due to the increase of the information requested.

Table 4 - Interview length (minutes and seconds) per technique, section and number of components (in brackets the number of households interviewed)

Data collection technique	Number of components	Interview Lenght			
		TOTAL	Number of compontents Section	GHI Section	HHG Section
CATI	1	1.19 (3)	0.25 (3)	0.54 (3)	0.0 (3)
	2	2.40 (11)	0.40 (11)	1.41 (11)	0.19 (11)
	3	3.23 (11)	0.24 (11)	2.17 (11)	0.42 (11)
	4	6.1 (12)	0.44 (12)	3.35 (12)	1.42 (12)
	5	6.54 (4)	0.32 (4)	3.55 (4)	2.26 (4)
CAWI	2	3.27 (11)	0.49 (11)	2.7 (11)	0.32 (11)
	3	4.3 (9)	0.40 (9)	2.26 (9)	0.56 (9)
	4	5.26 (10)	0.28 (10)	3.23 (10)	1.35 (10)
	5	7.11 (7)	0.37 (7)	4.16 (7)	2.17 (7)
	6	7.32 (1)	0.56 (1)	3.32 (1)	3.43 (1)
	8	15.59 (1)	0.43 (1)	9.43 (1)	5.34 (1)
PAPI	1	5.0 (1)	-	-	-
	2	4.8 (8)	-	-	-
	3	4.40 (9)	-	-	-
	4	6.0 (7)	-	-	-
	5	6.0 (4)	-	-	-
	6	8,3 (2)	-	-	-

Source: Istat, Experimental survey on household grid - 2018

In addition, results show that the largest households (2-8 members) were assigned to CAWI[4], while those interviewed through CATI or PAPI had a maximum size of 5 and 6 respectively (Table 4).

On account of this, in order to control for a possible household size effect on the interview length, we compared households with the same size. For instance, focussing on 4 people families, it can be observed that CAWI reported the lowest total length (5'26''), while CATI and PAPI lasted on average 6'. Compilation times for the section on members and their relationships are quite similar both for CAWI and CATI (slightly lower for the former).

3.2 The identification of 'nuclei' in households

As already mentioned, when families have more than one nucleus, it is very important to identify all the bilateral relationships between family members (couple and parent-children relationships). The analysis of data showed that the HHG allows to solve possible ambiguities, which would remain with the old approach, due to the higher complexity of asking kinships in relation to one Reference Person or to the lack of some information. For example: if in a household there are more than two adults and children with similar age, the edit/imputation phase bears the risk of assigning the parent-children relation to the wrong people. This would not cause important problems from a statistical point of view, since households with more than a nucleus are not frequent in Italy, but it would give a picture not corresponding to the real situation. The test demonstrated that the availability of information on the relationship of each member with respect to all others guarantees the correct identification of all family nuclei.

3.3 The cognitive test results

Cognitive interviewers were managed through a structured questionnaire containing the questions described in paragraph 2.2. The analysis of responses[5] showed an overall positive judgement in terms of easiness/directness in giving the answers for all data collection techniques.

---

4   It is worth mentioning that the long time needed for the 8-member household to complete the CAWI questionnaire (16 minutes) was also due to the software (Limesurvey), which doesn't perform efficiently when complex rules have to be checked for many members.

5   Respondents were asked to assign a score from 1 to 10 to indicate the easiness/directness in giving the answers.

Table 5 - Judgement on the easiness/directness in giving the answers (absolute numbers in brackets)

Score	Data collection technique			Total
	PAPI	CATI	CAWI	
6	3.2 (1)	0.0	2.6 (1)	1.8 (2)
7	0.0	0.0	7.7 (3)	2.7 (3)
8	3.2 (1)	0.0	10.3 (4)	4.5 (5)
9	12.9 (4)	4.9 (2)	10.3 (4)	9.0 (10)
10	80.7 (25)	95.1 (39)	69.1 (27)	82.0 (91)
Total	100 (31)	100 (41)	100 (39)	100 (111)

Source: Istat, Experimental survey on household grid - 2018

Table 5 shows that 82% of respondents assigned the maximum score, while the lowest score assigned (6) was provided by only 2 cases. Considering these 2 respondents, one of them simply pointed out that the interview was boring, while the other expressed difficulty in understanding the logical link between 'Marital status' and intra-household relationship. In fact, he had a partner but was not married, so he expected to find a corresponding item in 'Marital status', while in Italy de facto families are not recognised by law. This problem emerged also for other respondents, but it does not regard the household grid approach.

Among the 3 respondents who assigned the score 7, only one provided comments, complaining that the software system responded slowly. His family was the largest one, so this was caused by Limesurvey's already mentioned limits.

Concerning the easiness/difficulty in identifying the pertaining intra-household relationship (see Table 6), 90% of respondents said that the task was 'Very easy for all the relationships'. Some difficulties were reported for:

- adopted/temporarily assigned children;
- 'foster parent': the respondent said it would be easier to reverse the perspective, declaring that the boy was the child of his partner;
- the already mentioned logical link between 'Marital status' and intra-household relationship.

Respondents who declared 'Quite easy for all relationships', suggested to specify that item '09 - brother/sister' also includes siblings with only one parent in common.

In sum, problems encountered by respondents do not depend on the household grid approach, but on comprehension and completeness of definitions.

Table 6 - Easiness/Difficulty in identifying the pertaining intra-household relationship

In general, identifying the intra-household relationship was:	Data collection technique			Total
	PAPI	CATI	CAWI	
1 – Very easy for all relationships	93.5 (29)	92.7 (38)	84.6 (33)	90.1 (100)
2 – Quite easy for all relationships	0	4.9 (2)	10.3 (4)	5.4 (6)
3 – Easy for some relationships, difficult for others	6.5 (2)	2.4 (1)	5.1 (2)	4.5 (5)
4 – Quite difficult for all relationships	0	0	0	0
5 – Very difficult for all relationships	0	0	0	0
Total	**100.0 (31)**	**100.0 (41)**	**100.0 (39)**	**100.0 (111)**

Source: Istat, Experimental survey on household grid - 2018

The need to read again or to ask the interviewer to repeat some concepts emerged in only 11 cases for the following aspects:

- 'brother/sister in law' definition;

- relationships among siblings;

- the inability to find a suitable response item: in a few cases the pertaining relationship category was not selected because it was not displayed as a result of a wrong answer given in the GHI section (see Paragraph 2.1.1).

In perspective, le latter problem could be solved by showing on the upper section of each screenshot a synthetic family scheme, with name, gender and marital status of each household member. In this way, respondents can easily see whether they gave some wrong answers in the GHI section that may have caused the errors.

Finally, cognitive questions regarding the use of the options 'Other relatives' and 'Other non-relatives' showed that they were always selected correctly, except from one case when 'grandchild' was associated to the item '15 - Other parents' instead of the correct one.

3.4 The role of the interviewers

The interviewers played a key role in the success of this survey. Due to its characteristics, this experience can be considered a very particular methodological experimentation that touches on many topics of the theory and practice of social research (Marradi, 2007; Corbetta, 1999; Campelli, 1996; Statera, 1995; Agnoli, 1994). Regardless of the great interest of the survey's object itself, a further reason of methodological interest lies in the specificity of the context of the interviews (both the interviewers and the interviewees selected in the same professional context) and also in the unexpected reaction of many interviewees, which can be configured as a sort of serendipity (Merton & Barber, 1992). Indeed, it is a rare event that in an experimental survey both the interviewers and the interviewees are selected in the same professional context.

To better understand the tasks performed by the interviewers, we can analyse the whole process as organised in four steps. The interviewers had to:

a) follow an in-depth training on the questionnaire, including simulated interviews on possible criticalities related to the interview and specifically to the HHG;

b) contact the people to be interviewed;

c) carry out the interviews;

d) give their personal feedback on the experience.

During the in-depth training, the interviewers attended some meetings, in which the many activities to be carried out were described. The training was focussed on the possible criticalities related both to the questionnaire and to the recoding of the family relationships among the household members. In this regard, it is worth emphasising that the three different techniques that had to be tested (PAPI, CATI, and CAWI) involved a different kind of relationship with the respondents, and consequently different kind of work for the interviewers. In PAPI and CATI interviews, the interviewers had to submit the questions to the interviewees. They had to fill in the questionnaire themselves, codifying the kinship relations, as declared by the respondents. In the CAWI interview, they had to observe (in a non-participant mode) the respondent's behaviour during the compilation of the questionnaire. For all the three techniques, the

interviewers had to provide support – if requested – and ask the questions of the cognitive test. After the interview, they implemented the data file, adding their own observations.

In order to be skilled to cope with possible difficulties, very demanding interviews were simulated during the training phase. They were conducted on a set of hypothetical families, characterised by particularly complex bonds. These simulations were carried out with all the data collection techniques to be used in the real test. Following the training, the interviewers contributed to the improvement of some aspects related to the wording of some questions (Pitrone, 2005), which regarded, for example, the need to insert some further information in the PAPI questionnaire, useful to make it easier to carry out the interviews correctly.

As already said, interviewers and interviewees had been chosen in the same work context. This could create, from a methodological point of view, some doubts whether colleagues would agree to be interviewed more easily than in real surveys. Surprisingly, it was not as easy as expected. In fact, some colleagues expressed the same doubts as expressed by individuals during official surveys, *e.g.* concerns their privacy (Facioni, 2017). Furthermore, this research experience confirms the power of the *situational context*, from the point of view of the phenomenological theories (Garfinkel, 1967). For example, the whole group of interviewers had the feeling that the colleagues, during the interview, had changed their attitude towards them. They were not colleagues, but only individuals in an interview context. They returned to being colleagues only at the end of the cognitive test. This real serendipity represents a further element of interest in this research, which was carried out for a totally different purpose.

The interviewers debriefing highlighted that:

- the flow of the interviews was very simple to manage. Only for paper questionnaires it was suggested to make it clearer that HHG questions had to be asked row by row (and not column by column);

- respondents often requested clarifications on some general concepts not related to the management of the HHG (definitions of 'household', of usually 'resident household member', *etc.*)

Finally, interviewers observed the habit to consider 'relatives' also members that are not proper 'relatives' (for instance 'child of cousin'), which is probably due to a cultural sentiment. Anyway, from the statistical point of view, such an error of classification would not have an impact on the identification of the 'nuclei'.

4. Conclusions and perspectives

The main doubts about the household grid approach regarded the time to collect the information and the perceived respondent burden caused by the repetition of questions.

The test results showed that the interview length was contained for all the techniques (about 5 minutes, with the majority of time spent for GHI Section); responding was also considered easy and not burdening.

Concerning quality, the system of checking rules implemented for CATI and CAWI electronic questionnaire did not cause problems and guaranteed the correctness of data. Paper forms also were error-free. The only suggestion with regard to CATI and CAWI modes was to show a synthetic family scheme on the upper section of each screenshot, so that respondents can easily see whether they gave some wrong answers in the previous questions.

In addition, the test results confirmed the reduction of ambiguity in identifying family nuclei, thanks to the higher precision in the identification of intra-household relationships given by the household grid.

However, some aspects need to be improved; in particular, all the information addressed to respondents to explain some concepts (like the definitions of 'household' or 'usually resident household member', *etc.*) and the specifications to give correct responses (which members should be included or excluded, *etc.*) should be managed more efficiently.

The good results of the test demonstrate that this approach could be adopted for all the socio-demographic surveys, this way guaranteeing the homogeneity of survey questionnaires and comparability of data as well as a better accuracy in identifying household structures. This solution would be particularly important in the actual context of the Master Sample.

References

Agnoli, M.S. 1994. *Concetti e pratica nella ricerca sociale*. Milano, Italy: Franco Angeli.

Beatty, P.C., and G.B. Willis. 2007. "Research synthesis: the practice of cognitive interviewing". *Public Opinion Quarterly*, Volume 71, Issue 2: 287-311.

Bradburn, N.M. 1978. "Respondent Burden". Paper presented at the 138th *Annual Meetings of the American Statistical Association*, San Diego, CA, U.S.

Campelli, E. 1996. *Il metodo e il suo contrario*. Milano, Italy: Franco Angeli.

Conrad, F., J. Blair, and E. Tracy. 2000. "Verbal reports are data. A Theoretical Approach to Cognitive Interviews". In *Proceedings of the 1999 Federal Committee on Statistical Methodology Research Conference*. Washington, D.C., U.S.: Office of Management and Budget.

Corbetta, P. 1999. *Metodologia e tecniche della ricerca sociale*. Bologna, Italy: Il Mulino.

European Commission. 2009. "Social Variables Standardization". Item 5.3, *Meeting of the European Directors of Social Statistics*. Luxembourg, 23rd - 24th February 2016.

Facioni, C. 2017. "Le opinioni, via web, dei cittadini intervistati in Istat". In Romano, M.C. (a cura di). "L'utilizzo della tecnica CAWI nelle indagini su individui e famiglie": 153-159. *Letture Statistiche – Metodi*. Roma, Italy: Istat. https://www.istat.it/it/archivio/203729.

Frankel, J., and L.M. Sharp. 1981. "Measurement of Response Burden". *Statistical Reporter*, N. 81-4: 105-111.

Garfinkel, H. 1967. *Studies in Ethnomethodology*. Englewood Cliffs, NJ, U.S.: Prentice-Hall.

Marradi, A. 2007. *Metodologia delle scienze sociali*. Bologna, Italy: Il Mulino.

Merton, R.K., and E.G. Barber. 1992. *The Travels and Adventures of Serendipity: A Study in Historical Semantics and the Sociology of Science*. Princeton, NJ, U.S.: Princeton University Press.

Oksenberg, L., C.F. Cannell, and G. Kalton. 1991. "New Strategies for Pretesting Survey Questions". *Journal of Official Statistics – JOS*, Volume 7, N. 3: 349-365.

Pitrone, M.C. 2005. *Il sondaggio*. Milano, Italy: Franco Angeli.

Presser, S., J.M. Rothgeb, M.P. Couper, J.T. Lessler, E. Martin, J. Martin, and E. Singer. 2004. *Methods for testing and evaluating survey questionnaire*. Hoboken, NJ, U.S.: John Wiley & Sons, *Wiley Series in Survey Metodology*.

Sharp, L.M., and J. Frankel. 1983. "Respondent Burden: A Test of Some Common Assumptions". *Public Opinion Quarterly*, Volume 47, Issue 1: 36-53.

Statera, G. 1995. *Logica dell'indagine scientifico-sociale*. Milano, Italy: Franco Angeli.

Sudman, S., N.M. Bradburn, and N. Schwartz. 1996. *Thinking About Answers: The Application of Cognitive Processes to Survey Methodology*. San Francisco, CA, U.S.: Jossey - Bass Publishers; Hoboken, NJ, U.S.: John Wiley & Sons, *Wiley Series in Research Methods for the Social Sciences*.

Willis, G.B. 2005. *Cognitive Interviewing: A Tool for Improving Questionnaire Design*. Thousand Oaks, CA, U.S.: SAGE Publications.

Willis, G.B. 1994. "Cognitive Interviewing and Questionnaire Design: A Training Manual". *Working Papers Series*, N. 7. Atlanta, GA, U.S.: U.S. Department of Health and Human Services, Centers of Disease, Control and Prevention – CDC, National Center for Health Statistics.