

## **INNOVATIVE DATA COLLECTION STRATEGIES: ISTAT SURVEY ON « WELLBEING AND SAFETY OF PEOPLE »<sup>1</sup>**

Federica Pellizzaro, Lorella Sicuro

**Abstract.** The ISTAT survey "Wellbeing and Safety of People" analysed subjective well-being and perceived safety in Italy. It involves a sample of about 25,000 individuals aged 16–75, with 21,000 Italians interviewed via CATI (telephone) and 4,500 foreigners via CAPI (face-to-face). A key innovation was the contact strategy for CATI interviews: participants received an informative letter and a reminder, encouraging them voluntarily provide a phone number through a toll-free number or an ISTAT webpage (via tax code or QR code). Another methodological novelty was the formation of interview quartets, prioritizing individuals who had proactively shared their contact details, thereby reinforcing the link between voluntary action and participation. This paper has the goals to assess the impact of the communication strategy and reminders on response rates and participation patterns by age and geographical area. The approach offers useful insights to improve citizen engagement in future surveys and could serve as a replicable model to reduce territorial inequalities in participation.

### **1. Introduction**

The Italian National Institute of Statistics (Istat) survey “Wellbeing and safety of people”, included in the National Statistical Programme (code IST-02260), has as its main objective an in-depth examination of women's safety, with particular attention to psychological, physical and sexual violence, as well as to the perception of safety in everyday life contexts. The survey, which was also carried out in other European countries for Eurostat (Eurostat, 2021), involves a random sample of about 25,500 women between the ages of 16 and 75, of whom 21,000 are Italian (interviewed by telephone - CATI technique) and 4,500 are foreign (interviewed face-to-face - CAPI technique). To protect the participants, the survey was presented with a generic name and not explicitly referring to gender-based violence, to avoid risk situations in case the communication was read by family members, including potential perpetrators of violence. The general wording also avoided defensive reactions, contributing to a climate of trust and greater willingness to participate. An element of methodological

---

<sup>1</sup> The paper is the result of the common work of the authors. In particular, paragraph 1, 3, 4.4 and 5 is attributed to Lorella Sicuro; paragraph 2, 4.1, 4.2 and 4.3 to Federica Pellizzaro.

innovation for the CATI part of the survey concerned the voluntary transmission of telephone numbers by the selected women, either via a toll-free number or by filling in a short online questionnaire (also via QR code). In a context marked by distrust of unknown calls, this strategy proved effective in obtaining correct numbers and ensuring direct contact. The communication was constructed according to principles of effective persuasion (Cialdini, 2002, Petty and Cacioppo, 1986). In particular, three persuasion mechanisms were considered especially relevant. First, the principle of authority was activated through the official letter signed by the President of Istat, thereby reinforcing institutional credibility and trust. In addition, the communication clearly explained the purpose of the survey, the safeguards for data protection and the practical relevance of participation. Second, the appeal to social utility and collective benefit emphasised the contribution of individual participation to the improvement of public policies, fostering a sense of civic responsibility. Third, elements of commitment and consistency were embedded in the request to voluntarily provide a telephone number, encouraging an initial active step that could increase the likelihood of subsequent participation in the interview. This strategy made it possible to collect around 36,000 valid numbers, strengthening the quality of the sample and improving operational efficiency. In addition, in the formation of the 21,000 planned quartets of interviewees, priority was given to those who had spontaneously provided the address, to maximise the phone coverage while preserving sample stratification. This paper aims to evaluate the effectiveness of the communication strategy adopted for the CATI component of the survey, with particular attention to: (i) the impact of reminder letters on the timing and volume of voluntary contact provision; (ii) the socio-demographic determinants of telephone number transmission; and (iii) the presence of territorial disparities in the voluntary provision of contact information. By addressing these research questions, the study contributes to the assessment of how structured communication strategies may influence engagement in sensitive official surveys.

## **2. Communication to participation and delivery**

The women selected to take part in the survey had received the letter to their address of residence, presenting the survey and requesting their cooperation in providing their telephone number for the purposes of contact for the interview. The transmission of the telephone number could be done by accessing a special web page of Istat, authenticating with one's id tax code, or by contacting a toll-free number, active six days a week, communicating the unique code indicated in the letter. Furthermore, it was specified that the survey would be conducted between January and June 2025 by interviewers from the company CSA (Centro Statistica Aziendale

S.r.l.), on behalf of Istat. The contact by the interviewers to the selected persons would be made by telephone on weekdays, with a wide time slot, using a telephone number that was explicitly indicated in the letter, thus making the contact identifiable. To persuade the woman to provide the telephone number, the social relevance of the survey, the value of participation and the concrete benefits for the respondents and the community were emphasised. It was also emphasised that providing the telephone number would ensure flexible participation, adaptable to family and work commitments, and that the data provided would be used exclusively for statistical purposes and in a confidential manner. Furthermore, the contribution of individual participation to the improvement of public policies was emphasised, promoting a sense of social utility, active involvement and belonging. Ample space was devoted to the protection of personal data and compliance with statistical confidentiality and privacy regulations, also clarifying the mandatory nature of participation.

The informative letter was sent by ordinary mail at the beginning of November to 126,000 women, and a subsequent reminder at the end of November to 110,000 who had not yet provided their telephone number. The letters were sent in three languages: Italian and German for persons in the province of Bolzano, Italian and Slovene for persons in the Slovene-speaking areas (Friuli Venezia Giulia).

### **3. Materials and methods**

For the purposes of the survey, two samples of women were extracted: the sample consisting of Italian women for whom the telephone interview (CATI technique) and the sample of foreign women with a face-to-face interview (CAPI technique). For the CATI interviews, a sample of 126,000 Italian women aged between 16 and 75 was drawn from the Municipal Population Registers of the National Institute of Statistics (IST-02492), stratified by region and municipal type. An oversampling was therefore carried out to have an adequate number of replacement units and to reach the expected sample size of 21,000 women. The analysis envisaged the use of descriptive and inferential statistical techniques, with the aim of exploring the mode and time of response of the women involved in the CATI survey to the request to provide a telephone number for the interview. In particular, the territorial distribution of the respondents who transmitted their telephone number was analysed, with a focus on the differences by age group and geographical area. Specific attention was given to the way in which the respondents provided their telephone number: through self-compilation or by contacting the toll-free number. This distinction also made it possible to explore some aspects related to the level of digitalisation of the respondents. The time trend of the responses was also examined, considering the

interval between the receipt of the information letter and the transmission of the telephone number, with the aim of assessing the possible effect of a reminder sent during the survey.

Besides, it was described the strategy adopted for forming the interview quartets.

Referring to the entire sample of 126,000 women, the Chi-square test was applied between the categorical variables telephone number transmission (yes, no) and geographical division, telephone number transmission (yes, no) and age group, with the aim of examining whether providing one's telephone number was statistically significantly associated ( $p < 0.05$ ) with age group or geographical distribution. To investigate the propensity to provide one's telephone number as a function of geographical area and age group, a logistic regression model was estimated. The model had as dependent variable the information on providing one's telephone number (yes, no), as independent variables geographic affiliation (five geographic distributions: North-West, North-East, Centre, South, Islands) and three defined age groups (16-34, 35-64, 65-74). Interaction effects between age group and geographical area were not included in the final specification to preserve model parsimony and ensure interpretability, given the primarily descriptive objective of the analysis. The results of the model were interpreted through the analysis of statistically significant estimated Odds Ratios ( $p\text{-value} < 0.05$ ). All analyses were conducted using SAS software.

## 4. Results

### 4.1. Statistical analysis of the voluntary transmission of telephone numbers

The overall rate of women who voluntarily provided their telephone number was 28.6%, i.e. one in four women. Response rates varied considerably across regions, ranging from over 40% in Valle d'Aosta and around 39% in the Autonomous Province of Trento to below 20% in several Southern regions such as Sicilia, Campania and Calabria (Table 1).

The table 2 shows how the telephone number was communicated for age groups, analysing the difference between autonomous compilation by the user and transmission via a contact centre. Of the total 35,984 cases, the vast majority - 26,756 people, or 74.36% - chose to provide their telephone number autonomously, without the intervention of an operator. In contrast, 25.64% (or 9,228 cases) provided their number through the help of the operator. These data indicate a clear preference for the autonomous mode of telephone number communication. Relating the channel chosen and age, it is observed that as age increases, the propensity to use autonomous modes decreases, while reliance on the contact centre increases. In the youth and

adult groups (16-24, 25-34, 35-44), the percentage of autonomous compilation is very high, exceeding 88%, with a peak of 93% among 25–34-year-olds. A progressive decline in self-compilation is observed in the intermediate age groups (45-54, 55-64): from 79.9% among 45–54-year-olds, down to 64.2% in the 55-64 age group. Among the over-65s, only 49.8% filled in autonomously, while the remaining 50.2% relied on the contact centre.

In addition, analysing how the telephone number was communicated for geographical area, although the differences are not substantial, the data suggest that users in the North-East are more inclined to use the self-completion channel (79.5%), while those in the North-West tend to rely slightly more on assistance from the Contact Center (72.9%). The other geographical areas fall within a narrow range around the national average, showing moderate uniformity in channel preference. (Table 3)

**Table 1** – *Voluntary transmission of telephone numbers by women who received the informative letter for participation in the survey, absolute values and percentages.*

Region	Yes	No	Total	%
Abruzzo	1,325	3,853	5,178	25.6
Basilicata	1,254	3,564	4,818	26
Calabria	855	4,713	5,568	15.4
Campania	1,150	6,476	7,626	15.1
Emilia-Romagna	2,365	4,325	6,690	35.4
Friuli-Venezia Giulia	1,977	3,171	5,148	38.4
Lazio	1,912	5,564	7,476	25.6
Liguria	1,708	3,667	5,375	31.8
Lombardia	3,384	6,228	9,612	35.2
Marche	1,763	3,511	5,274	33.4
Molise	1,230	3,438	4,668	26.3
Piemonte	2,324	4,516	6,840	34
Provincia autonoma di Bolzano	997	3,718	4,715	21.1
Provincia autonoma di Trento	1,862	2,908	4,770	39
Puglia	1,470	5,244	6,714	21.9
Sardegna	1,337	4,069	5,406	24.7
Sicilia	1,051	6,125	7,176	14.6
Toscana	2,196	4,254	6,450	34

**Table 1 (cont.)** – *Voluntary transmission of telephone numbers by women who received the informative letter for participation in the survey, absolute values and percentages.*

Region	Yes	No	Total	%
Umbria	1,622	3,322	4,944	32.8
Valle d'Aosta	1,878	2,676	4,554	41.2
Veneto	2,324	4,672	6,996	33.2
Total	35,984	90,014	125,998	28.6

**Table 2** – *Voluntary transmission of telephone numbers by age and communication method, self-completion or Contact Centre, absolute values and percentages.*

Age groups	Self-compilation	Contact Center	Total	% Self-compilation
16-24	3,241	346	3,587	90.3
25-34	4,198	314	4,512	93.0
35-44	4,849	626	5,475	88.6
45-54	5,672	1,425	7,097	79.9
55-64	5,220	2,914	8,134	64.2
65+	3,576	3,603	7,179	49.8
Total	26,756	9,228	35,984	74.4

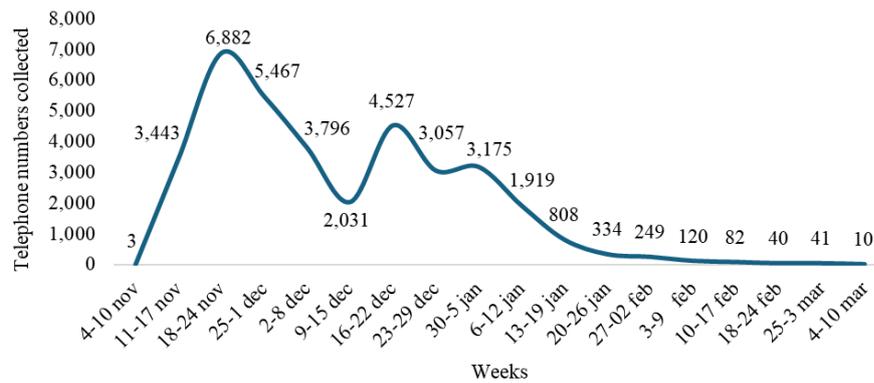
**Table 3** – *Voluntary transmission of telephone numbers by geographical area and communication method, self-completion or Contact Centre, absolute values and percentages.*

Geographical Area	Self-compilation	Contact Center	Total	% Self-compilation
North-East	4,120	1,063	5,183	79.5
North-West	9,946	3,690	13,636	72.9
Centre	5,477	2,016	7,493	73.1
South	5,418	1,866	7,284	74.4
Islands	1,795	593	2,388	75.2
Total	26,756	9,228	35,984	74.4

#### *4.2. Time course of telephone number transmission*

An analysis of the time trend in responses (Figure 1) clearly shows how communication actions played a decisive role in promoting and supporting participation in the survey. After the initial informative letter was sent out on 30 October, the first two weeks of November were characterised by a very low number of telephone numbers received. This initial trend can be interpreted as a physiological phase, during which the recipients received the communication, began to take notice and become familiar with the instructions provided. Starting from the week of 11-17 November, there was a first important increase in the number of feedback, culminating in a peak in the following week (18-24 November), when 6,882 cards were received. Thereafter, the flow of participation remained steady thanks to the sending of a reminder on 30 November. This second communication intervention had the effect of reinforcing the focus on the survey and prolonging the involvement of participants. In fact, in the following weeks, until 22 December, a regular and continuous flow of mailings was observed, with weekly values of between 3,000 and 5,000 telephone numbers.

During the Christmas holiday there was a physiological decline, probably due to the temporary absence of many people from their homes or the suspension of usual daily activities. However, upon returning from the holidays, there was a slight upturn in the number of responses. As the weeks pass a gradual decline in participation was observed, leading to an almost complete cessation of mailings around mid-March. Overall, the temporal analysis of the trend in responses highlights the effectiveness of the communication tools used, both in terms of initial activation (thanks to the informative letter) and maintaining involvement over time (thanks to the reminder). The overall total of 35,984 telephone numbers collected was achieved mainly thanks to intense activity concentrated in the period between mid-November and the end of December, when the greatest number of responses occurred.

**Figure 1** – Time course of telephone number transmission.

#### 4.3. Construction of the quartets for final sample formation

The methodology developed for the survey focuses on building groups of four respondents (“quartets”) from an initial random sample of six (“sextets”), with the aim of optimizing telephone contact of the CATI campaign. To this end, a reserve pool of contactable individuals was integrated using a priority-based algorithm in three R scripts. These scripts allowed for the creation, updating, and formatting of the quatrain to ensure coverage, stratification and data quality. The process emphasized the reuse of the best available telephone numbers (preferably self-reported, which were placed at the first position of each quatrain and, when possible, also at the second position), ensuring replicability and minimizing the need for manual intervention in future iterations of the survey.

#### 4.4. The propensity to provide the telephone number

The chi-square tests of independence (Tables 4 and 5) were elaborated to assess the existence of statistically significant associations between telephone number sharing (yes/no) and each of the two categorical variables (age group 16-34, 35-64 and 65-74 and geographical areas Northwest, Northeast, Centre, South and Islands). The results showed that the observed differences were statistically significant in relation to both age group ( $p < 0.001$ ) and geographical area ( $p < 0.001$ ), displaying the presence of a relationship between these individual characteristics and the propensity to share one's contact. The logistic regression analysis (Table 6) carried out subsequently made it possible to assess the effect of age group and geographical

area (independent variables) on the propensity of passing on the telephone number (dependent variable). The model was statistically significant, as confirmed by the Likelihood Ratio test comparing the fitted model with the intercept-only specification ( $p < 0.001$ ), indicating a significant improvement in model fit. Pseudo- $R^2$  indices showed limited explanatory power (Cox & Snell  $R^2 = 0.0208$ ; Nagelkerke  $R^2 = 0.0298$ ), suggesting that age group and geographical area account for approximately 2–3% of the variance in the propensity to provide a telephone number. Although modest, such values are common in large-scale social surveys examining individual behavioural decisions with a restricted set of socio-demographic predictors. Participation in a sensitive survey is likely influenced by additional unobserved factors—such as institutional trust, digital literacy, privacy concerns, and broader cultural attitudes—which were not included in the model. Nevertheless, the estimated effects remain statistically robust and substantively consistent. The results, expressed in terms of Odds Ratio (OR), showed that adult women (35-64 years) have a slightly higher inclination of passing on the telephone number than younger women (16-34 years) (OR = 1.045; IC95%: 1.014-1.077), while older women (65+) show a lower propensity (OR = 0.930; IC95%: 0.895-0.966). These differences were statistically significant, indicating that age plays a role in contact-sharing behaviour. Regarding geographical distribution, compared to women residing in Central Italy (reference category), those resident in the North-West and North-East had a higher inclination of passing on the telephone number, with Odds Ratios of 1.209 (IC95%: 1.165-1.255) and 1.125 (IC95%: 1.085-1.168) respectively. In contrast, the propensity was significantly lower among residents in the South (OR = 0.593; IC95%: 0.571-0.616) and the Islands (OR = 0.520; IC95%: 0.494-0.548), compared to the Centre.

**Table 4** – *Chi-square test between age group and telephone number transmission ( $p < 0.05$ ).*

Age groups	Telephone number transmission		Total	Chi-square test
	no	yes		
16-34	21,638	8,499	30,137	<.0001
35-64	50,363	20,848	71,211	
65+	17,975	6,637	24,612	
Total	89,976	35,984	125,960	

**Table 5** – Chi-square test between geographical distribution and telephone number transmission ( $p < 0.05$ ).

Telephone number transmission			Chi-square test	
Geographical Area	no	yes	Total	
Centre	16,645	7,493	24,138	<.0001
Islands	10,193	2,388	12,581	
North-East	18,788	9,525	28,313	
North-West	17,072	9,294	26,366	
South	27,278	7,284	34,562	
Total	89,976	35,984	125,960	

**Table 6** – Logistic regression model with dependent variable (DP) the transmission of the telephone number (yes, no) and as independent variables the age group (16-34, 35-64 and 65+ with reference mode 16-34) and the geographical distribution (North-east, North-west, Centre, South and Islands with reference mode Centre), related Odds Ratio and 95% Confidence Intervals.

Regression logistic model			
DP: Telephone number transmission (yes/no)	Odd Ratio	95 % Confidence Interval	
35-64	1.045	1.014	1.077
65+	0.930	0.895	0.966
North-East	1.125	1.085	1.168
North-West	1.209	1.165	1.255
Islands	0.520	0.494	0.548
South	0.593	0.571	0.616

## 5. Conclusion

Effective communication is one of the crucial elements for the success of statistical surveys, especially when the active and voluntary participation of those involved is required. The way a survey is presented, the language adopted, the clarity of the message, and the communication channels used can significantly affect response rates and data quality. In this context, the experience of the “Well-being and safety of people” survey offers a case in point. The findings confirm the effectiveness of the communication strategy adopted to encourage the voluntary transmission of telephone numbers. With an overall response rate of 28.6%, the information campaign reached a significant proportion of participants. The communication approach adopted - based on transparency, flexibility and the social

value of participation - proved to be consistent with international recommendations for increasing involvement in sensitive surveys (Eurostat, 2021).

The statistical analysis revealed strong differentiations in the propensity to send the telephone number in relation to socio-demographic factors. Age was found to be a significant predictor: preference for autonomous transmission (web or QR code) declines from around 90% among younger women to 49.8% among those aged 65 and over, who rely more frequently on the Contact Centre. This confirms the findings of previous studies on the generational digital divide (van Deursen & Helsper, 2015) and underlines the importance of offering alternative channels (such as using a Contact Centre) to ensure equity of access. The predominance of the autonomous mode (74.4%) highlights the growing importance of digitalisation in survey implementation, although human support remains essential for older and vulnerable groups. These findings confirm the need for a multi-channel, inclusive approach combining digital and traditional modes, as recommended in mixed-mode survey research (de Leeuw, 2005).

The results also confirm the effectiveness of the reminder sent, which encouraged a significant number of selected women to voluntarily submit their telephone number, confirming the ability of reminders to maintain interest and further stimulate participation.

Response rates varied geographically, exceeding 35% in Northern and Central Italy and falling below 20% in the South. These disparities may reflect differences in institutional trust, socioeconomic conditions, and digital culture, consistent with recent findings on civic participation in Italy (Istat, 2023). Although the initial sample was probabilistic, the voluntary provision of contact details introduced a self-selection component at the contact stage, as women who shared their telephone numbers may differ systematically from those who did not. However, this did not alter the probabilistic foundation of the design. To limit operational distortions, a fully automated and replicable algorithm was implemented to construct and update contact quartets, maximising telephone coverage while preserving stratification and minimising manual intervention. In conclusion, the strategy adopted by Istat proved to be effective and replicable for future surveys. Voluntary involvement, supported by clear communication focused on social utility, strengthened the quality of the sample and ensured greater compliance. However, the persistence of territorial and generational gaps calls for more targeted policies on the front of statistical culture and digital inclusion, so that more and more people feel motivated to participate and contribute to statistical surveys. Although the communication strategy was designed in alignment with international methodological recommendations (Eurostat, 2021), its generalisability beyond Italy should be considered with caution. The underlying principles—transparency, multi-channel access, institutional authority, and behavioural commitment—are theoretically transferable; however, their

effectiveness may depend on contextual factors such as levels of institutional trust, cultural norms, and prior exposure to official surveys. Survey methodology research suggests that response behaviour is strongly shaped by these macro-level conditions (Groves *et al.*, 2009; Harkness *et al.*, 2010). Cross-national applications would therefore be necessary to assess the robustness and replicability of the strategy across different settings.

## References

- CIALDINI R. B. 2001. *Influence: Science and Practice (4th ed.)*. Boston: Allyn & Bacon.
- DE LEEUW E. D. 2005. To mix or not to mix data collection modes in surveys. *Journal of Official Statistics*, Vol. 21, No. 2, pp. 233–255.
- EUROSTAT. 2021. *Violence against women survey: Methodological guidelines*. Publications Office of the European Union.
- ISTAT. 2023. *Cittadini e fiducia nelle istituzioni – Indagine Multiscopo 2022*. Report.
- GROVES R. M., FOWLER F. J., COUPER M. P., LEPKOWSI J. M., SINGER E., TOURANGEAU R. 2009. *Survey Methodology* (2nd ed.). Wiley.
- HARKNESS J. A., BRAUN M., EDWARDS B., JOHNSON T. P., LYBERG L. E., MOHLER P. P., PENNEL B.-E., SMITH T. W. 2010. *Survey methods in multinational, multiregional, and multicultural contexts* (Wiley Series in Survey Methodology). Wiley. <https://doi.org/10.1002/9780470609927>
- PETTY R., CACIOPPO J. 1986. Elaboration Likelihood Model of Persuasion. In SPRINGER SERIES IN SOCIAL PSYCHOLOGY ((SSSOC)) *Communication and persuasion* Springer, New York, NY. [https://doi.org/10.1007/978-1-4612-4964-1\\_1](https://doi.org/10.1007/978-1-4612-4964-1_1)
- PUBLISHING LIMITED, LEEDS *Communication and Information Technologies Annual (Studies in Media and Communications)* Vol. 10, pp. 29-52. <https://doi.org/10.1108/S2050-206020150000010002>.
- VAN DEURSEN A.J.A.M., HELSPER E.J. 2015. The Third-Level Digital Divide: Who Benefits Most from Being Online? In EMERALD GROUP