

Experiences gained as a result of TurkStat Survey on ICT Usage in Households and by Individuals carried out by CATI for the first time in 2020

Bilal Kurban^a (corresponding author), Aslihan Kabadayı^b, Ayhan Doğan^c, Dr. Cihat Erce İşbaşar^d

^a TurkStat Expert, Science & Technology Statistics Unit, bilal.kurban@tuik.gov.tr; ^b TurkStat Expert, Science & Technology Statistics Unit, aslihan.kabadayi@tuik.gov.tr; ^c TurkStat Expert, Head of Science & Technology Statistics Unit, ayhan.dogan@tuik.gov.tr; ^d TurkStat Expert, Head of Sectoral Statistics Department, erce.isbasar@tuik.gov.tr

The study comparing two data collection modes Computer Assisted Personal Interviewing (CAPI) and Computer Assisted Telephone Interviewing (CATI) and evaluating the results of the CATI has been reported from the perspective of Survey on Information and Communication Technology (ICT) Usage in Households and by Individuals which has been conducted by Turkish Statistical Institute (TurkStat) Science and Technology Unit. It is an annual survey basically involving access to and use of ICTs, use of the Internet and other electronic networks for different purposes and ICT competence and skills in respondents' households. Data collection mode of the survey in 2020 has been transition to CATI mode from CAPI mode due to the Covid-19 pandemic. This is the first study in the TurkStat to permit direct assessment of CAPI and CATI mode effects. Little evidence is found of mode effects in terms of survey quality. Almost no negative quality difference observed as a result of mode change, and the cost was reduced. Therefore, in line with the positive opinions of the regional offices, it has been decided to continue the survey implementation via CATI in the coming years, too.

Key words: Covid-19, data collection, CATI, CAPI, mode effect, TurkStat.

The views, comments and opinions expressed in this paper are those of the author and do not necessarily reflect the opinions or policies of TurkStat. The same rule applies for all the work done on the basis of this paper.

1. Introduction

Information and communication technology (ICT) affects daily life of people in many ways, both at work and in the home, for instance, during communicating, following the news, interacting with public authorities, buying products online or being entertained.

Access to and the use of ICTs is a prerequisite for success in the labor market, education, the use of e-services in health care, transportation, social security system, public administration and trade, social and cultural activities, search and processing of information etc. Although digital exclusion was due to lack of access to technology a few decades ago, this is no longer the case in the modern society (Eurostat, 2020). However, there are still many people do not have technology or digital access in the world. Access to digital technologies provide as primary tool for digital transformation. Digital transformation of society refers to a process where humans are reshaping the way society 'works' by ways of interpreting and understanding society, including the usage of digital technologies in daily life (Norqvist, 2018).

Survey on ICT Usage in Households and by Individuals has been conducted by TurkStat since 2004 annually by Eurostat methodology and aims to provide statistics for monitoring and evaluation of information society strategies. Data are collected from the individuals aged 16-74 in households which were selected by sampling method. Before 2020 Survey on ICT Usage in Households and by Individuals carried out via CAPI mode. Since the Covid-19 outbreak is impacting on the ability of TurkStat to collect and process data in the traditional manner, data collection mode of the 2020 survey has been transition to CATI mode.

This paper aims to offer some analytical and strategic implications regarding the application of two different modes in TurkStat Survey on ICT Usage in Households and by Individuals. The paper is to compare the results of the survey conducted with the CATI mode for the first time in 2020 and the results of the survey applied with the CAPI method which was before 2020 in terms of data quality. The study will also include the experiences gained during the CATI implementation.

TurkStat is planning to redesign all its national household surveys in a way that data collection for them is achieved also via web as well as effective use of CATI mode. TurkStat has a platform in order to implement these two modes also. It is expected that this study will provide support for the application of both modes in household surveys conducted by TurkStat. This study will try to explain the challenges encountered and experiences gained from the perspective of TurkStat ICT Usage in Households and by Individuals Survey, whose data collection mode has been switched from CAPI to CATI by an agile decision due to the covid-19 pandemic.

1.1. Methodology

For the reasons mentioned in the introduction, it has been decided to conduct a feedback survey in order to get the opinions of interviewers of regional offices about master survey (Survey on ICT Usage in Households and by Individuals) implementation which was carried out by phone for the first time between April and May 2020. Following the decision taken, a questionnaire was designed to be filled in by the supervisors who took part in the master survey in the regional offices and this questionnaire was applied in September 2020. The paper not only evaluates the results of the feedback survey but also intends to compare the data

quality of two different data collection modes used in Survey on ICT Usage in Households and by Individuals.

1.1.1. Feedback survey questionnaire

In order to obtain interviewer opinions by a feedback survey, the following three questions were asked them for the overall and each question of the master survey:

- › Level of difficulty in receiving responses (and reason(s) for difficulty),
- › Which mode is thought to be more suitable for the survey question, telephone or face to face (and suggestion (s) to continue the master survey to be conducted over the phone in the following years as well) and
- › Whether the survey question was understood when the question was first read on the phone (if not, how was the question probed or explained).

In this study, the results of this feedback survey will be evaluated together with the results of the Survey on ICT Usage in Households and by Individuals.

1.1.2. Data collection

The questionnaire of the feedback survey was designed to be filled in by the supervisors who took part in the master survey in the 26 regional offices of TurkStat and this questionnaire was applied in September 2020. A data entry form was created for this feedback survey via the MS Access and sent to the regional offices. It was applied on total of 46 supervisors.

1.1.3. Data analysis

The questionnaire forms filled out by the regional offices were combined on the return via SAS, and the results were created as frequency distributions. Answers given to open-ended questions were grouped. These will be discussed in later sections.

1.2. Data Collection Mode

CAPI is a data collection mode which involves the interviewer entering responses directly into a portable laptop computer. CATI is a telephone surveying technique in which the interviewer follows a script provided by a software application. Data collection method of Survey on ICT Usage in Households and by Individuals which has been applied annually since 2004 by TurkStat was CAPI until 2020. By interviewing face to face with household individuals, queries in both household and individual questionnaires were done and entry of the data was performed by through laptops.

The Covid-19 outbreak is impacting on the ability of statistical authorities to collect and process data in the traditional manner. Face-to-face interviews have been replaced with telephone and web-based surveys, nowadays (Eurostat, 2020). Due to the pandemic, it was more preferable to work remotely instead of door-to-door visits to protect the health of both field interviewers and respondents. TurkStat has been reflecting on its past experiences and has commenced transitioning to other modes of primary data collection.

This led to the conclusion that the TurkStat Survey on ICT Usage in Households and by Individuals switched to CATI before it had been carried out entirely in CAPI for years. For the first time, it was applied with the CATI method in 2020.

2. TurkStat Survey on ICT Usage in Households and By Individuals

TurkStat Survey on ICT Usage in Households and by Individuals is an annual survey conducted since 2004. With the survey data are collected on the use of information and communication technologies, the internet, e-government and electronic skills in households and by individuals.

2.1. Purpose

The main purpose of the survey is to help better understanding how the digital technologies are transforming the world of households and individuals. In addition to this, the survey aims to obtain data on following topics:

- › ICT technologies in households,
- › Access to and use of ICTs by individuals and/or in households,
- › Use of the Internet and other electronic networks for different purposes by individuals and/or in households,
- › Barriers to the use of ICT and the Internet,
- › E-commerce,
- › Use of e-government and
- › ICT security and trust.

2.2. Coverage

All settlements in Turkey are covered in sample selection. Scope of the survey covers individuals aged 16-74 and households with at least one member in the age group 16-74 years old. All private households of Republic of Turkey, which has 16-74 years of age and over persons who are Turkish citizens and foreigners are covered. Institutional population (dormitories, rest homes for elderly persons, special hospitals, military barracks and recreation quarters for officers etc.) are not covered.

2.3. Sampling Unit

Below is a summary of the sampling of the survey:

- › First Stage Sampling Unit is clusters composing of the average 100 addresses.
- › Second Stage Sampling Unit is households residing in the address selected from clusters in sampling.
- › Observation Unit is individuals aged 16-74.
- › Estimation size is Turkey in general with Statistical Region Level 1 is taken.
- › Sampling method is two stage stratified cluster sampling.
- › Sample size is total of 17,530 sample households throughout Turkey in April-May 2020.

2.4. Sources of the Data

Data are collected from the households which are selected by defined sampling method. Household questionnaires are filled by interviewing individuals aged 18 and older and who can give information about the household. Individual questionnaires are collected from individuals aged 16-74.

Telephone numbers of selected sample households and individuals were compiled from internal and external administrative sources. Although great success has been achieved in obtaining phone numbers from administrative data, during implementation (as expected) some of these proved to be inaccessible or inaccurate, and they were assigned to the relevant non-response codes predetermined by survey methodologists.

2.5. Data Collection

Before the Covid-19 outbreak, CAPI was used as data collection mode for all households surveys including Survey on ICT Usage in Households and by Individuals conducted in TurkStat. After 2020 data collection has been transition to CATI mode.

2.6. Compilation Practices

Questionnaires are prepared based on model questionnaire which are designed and recommended by Eurostat. Model questionnaires are adapted to the conditions in Turkey and also regulated in accordance with requirements of national institutions and also hold the National Information Society Strategy institutions/ organizations are organized in line with the needs.

The survey has been carried out in April and May of each year. Reference period is the first quarter of the year.

3. 2020 vs. Before 2020 Comparisons According to Obtained Results

3.1. Survey Design

As in previous years, in 2020, the questionnaire prepared on the basis of the model questionnaire designed and recommended by Eurostat was implemented. A CATI specific questionnaire was not designed in 2020. For this reason, there was no difference between the years before 2020 and 2020 in terms of questionnaire design.

Regional offices collect Survey on ICT Usage in Households and by Individuals data through the desktop application of the survey portal (DDI-based statistical production platform developed by TurkStat Information Technologies Experts). The desktop application designed to work offline in netbooks and notebooks and allows collecting data on electronic platform without requiring constant internet connection. While creating DDI and Rule files for 2020 Survey on ICT Usage in Households and by Individuals data collection, no CATI specific development, customization and modification could be made due to time constraints. In 2020, as in previous years, data obtained by regional offices were combined and analyzed by central office of TurkStat.

Invitation to survey letters sent to sample households before the beginning of the survey in 2020 as before. Invitation letter contained of the purpose of the survey, importance, usage of the results of ICT data, confidentiality and mandatory items and application method. In 2020, for the first time of the CATI application, telephone numbers of individuals in the sample households were obtained from administrative sources and forwarded to the regional offices along with the address information of the sample households and individuals living in there (such as name-surname, age, gender).

A special telephone line was purchased by TurkStat for phone calls. First of all, an informative short message was sent to sample individuals on behalf of TurkStat for CATI application. In the message, in addition to the information contained in the invitation letter previously sent to the households, the name and surname of the interviewer who would call for the survey implementation and the period when would apply the survey were also included.

The interviewers called the individuals living at the sample addresses and, if respondents were ready for the interview, they completed the questionnaire via telephone. An appointment was set for unavailable individuals and the questionnaire was completed by calling again on the appointment date and time for them. The phone numbers of the individuals whose phone information was not in the sample address list were tried to be reached through the local authorities.

3.2. Effects on Response Rates (both for households and individuals)

It is of interest to know whether CATI induces different co-operation behavior from CAPI. In this study, it was assumed that mode of interviewing would be unlikely to affect the propensity for the interviewer to achieve contact at a sampled address. Response rates, calculated on the base of addresses at which contact was made, have therefore been compared.

There were slightly more refusals amongst the CATI mode for households non-response rate in 2020 (Table 1), though the difference, taking into account the clustered nature of the sample,

does not achieve significance. It is observed that there is a decrease in the non-response rate of individuals compared to the previous year.

Table 1. Non-response rates

	2017	2018	2019	2020
	via CAPI			via CATI
Households	5.1%	6.4%	6.7%	8.8%
Individuals	8.1%	8.7%	9.4%	8.5%

This study provides no evidence that overall non-response rate was affected by mode of data collection, and it is relevant to note that the response rates in subsequent years, when the survey was carried out entirely in CATI, were not notably different from those in earlier - CAPI years.

3.3. Comparison of Achieved Profiles

In order to carry out the analyses of mode effects on responses reported, it was desirable to control for any differences in sample composition for responders between the CAPI and CATI modes. Such differences could have arisen due to sampling variance or differential non-response. The intention was therefore to identify any significant differences in composition between the two samples.

The profiles of the two achieved modes were therefore compared in terms of demographic and grouping variables. Assuming that the variables investigated are not subject to any mode-induced response effect -in other words that respondents would give the same answer to each question whether interviewed in CATI or CAPI.

Table 2. Comparison of the achieved mode profiles

	2019 via CAPI	2020 via CATI
Gender		
Male	49.8%	49.8%
Female	50.2%	50.2%
Age group		
16-24	18.3%	17.9%
25-34	21.5%	21.2%
35-44	21.5%	21.5%
45-54	17.5%	17.4%
55-64	13.4%	13.7%
65-74	7.9%	8.3%
Education level		
No school completed	12.0%	10.3%
Primary school	30.1%	30.3%
Primary education or junior high school	19.1%	19.2%
High school and equivalent	20.4%	21.0%
Higher education	18.4%	19.2%

When the results of the survey applied with CAPI in 2019 and the survey applied with CATI in 2020 were compared in terms of gender, age group and education level, it was determined that there was no significant difference. The biggest difference is in respect of education level. The CATI mode contains fewer individuals who were no school completed (10.3%, compared with 12.0% in the CAPI mode), and more individuals who were high school and equivalent and higher education graduates (high school and equivalent with 21.0% and higher education with 19.2% in the CATI mode compared with 20.4% and 18.4% in the CAPI, respectively).

3.4. Effects on Responses to Questions

When it is taken into account digital transformation and developing technology, it is seen that the results of the selected indicators such as in the below table increase, as expected. There is no difference due to the mode difference in terms of indicators.

Table 3. Comparison of the achieved mode profiles

Selected indicators	2019 CAPI	2020 CATI
Internet users	75.3%	79.0%
Access to Internet	88.3%	90.7%
Broadband	87.9%	89.9%
Fixed broadband (ADSL, cable, optic fibre, etc.)	49.1%	50.8%
Mobile broadband (3G, 4,5G via mobile phone/ smart phone or modem)	86.9%	86.9%
Usage of e-government services	51.2%	51.5%
Individuals purchased goods or services over the Internet	34.1%	36.5%

3.5. Other effects

The paper also looks at effects on interview length, willingness of the respondents to be interviewed, interviewer and respondent reactions and survey administration and costs.

3.5.1. Interview length

The mean length of interview of Survey on ICT Usage in Households and by Individuals, as recorded by interviewers, was 19 minutes in CAPI in 2019, and 36 minutes in CATI in 2020. It is thought that both respondents and interviewers are responsible for the duration of an interview.

Interview length can be assumed to be influenced by the way an interviewer performs their task during the interaction and communication process with a respondent.

The most important reason for the longer duration for CATI is considered to be the first contact and persuade the respondents to the survey. In the CAPI application, when the interviewers visit to a sample household, they apply the questionnaire to all individuals in the household after informing the household for the survey. There is no need to re-inform and persuade individuals for the survey. However, in the CATI application, this informing and persuasion process is repeated for each individual.

Besides, the characteristics of respondents and questions (e.g. age, education, complexity of the question and response categories, memory retrieval, etc.) might affect the length of interview. The educational level and age of respondents may both be relevant in this study. It could be assumed that older people need more time to answer questions and as a consequence will need more time. It is seen that the proportion of 55-64 and 65-74 age group in CATI in 2020 is higher than the proportion of 55-64 and 65-74 age group in CAPI in 2019. Furthermore, when the education level is compared, it is observed that individuals more educated in 2020 than in 2019. With regard to interview length, it is expected that a negative effect from age and a positive effect from the respondent's educational level.

3.5.2. Willingness of the respondents to be interviewed

As the field application period of this survey fall into the pandemic period, it was observed that the respondents preferred to make the CATI application rather than face to face interview and were willing to take part in the survey. In addition, even though interviewers are officials, being visited by interviewers at respondent's resident is caused anxiety and insecurity for respondents. For this reason, CATI application for this survey was chosen as a best option for both respondents and interviewers.

TurkStat collects data from 81 provinces through 26 regional offices for each survey. Considering the size of the country and the geographical distribution of the sample addresses, visiting the sample addresses is very difficult for interviewers. It was observed that the interviewers were more willing to apply the questionnaire with CATI.

3.5.3. Interviewer and respondent reactions and survey administration

Interviewers generally reacted favorably to CATI, and reported that respondents did likewise. Implementation was different with CAPI, but there were no major problems once staff had got used to the new systems.

3.5.4. Costs

Since the studies on the cost table of the CATI application continues, the cost comparison of the CAPI and CATI modes cannot be made with exact figures yet. However, it is seen that the travel allowance, travel and vehicle expenses were paid to officials in the CAPI application in 2019 were not included in the CATI application in 2020. This unpaid allowance, travel and vehicle expenses reduced the total cost of the survey in substantial amount.

4. Regional Offices Feedback Survey

There are 26 regional offices of TurkStat (corresponding to NUTS II level regions), which are responsible mainly for data collection and fieldwork. The regional offices are mainly involved in the collection and initial examination of the basic data.

In TurkStat, it has been decided to conduct a feedback survey in order to get the opinions of interviewers of regional offices about master survey (Survey on ICT Usage in Households and by Individuals) implementation which was carried out by phone for the first time between April and May 2020. Following the decision taken, a questionnaire was designed to be filled in by the supervisors (together with the interviewers) who took part in the master survey in the regional offices and this questionnaire was applied in September 2020.

4.1. Design of the Feedback Survey

Feedback survey questionnaire was designed as an MS access form. The design of the questionnaire is as in the figure below:

Figure 1. Design of the regional offices feedback questionnaire

The screenshot shows a Microsoft Access form titled "Regional Offices Feedback Questionnaire on 2020 Survey on ICT usage in households and by individuals Collected by Phone". The form is divided into two main panes:

- ICT Usage Survey Questions (Left Pane, Gray):**
 - Question ID: C9
 - Question text: How often on average did you use the internet in the last 3 months?
 - a) Every day or almost every day
 - b) At least once a week (but not every day)
 - c) Less than once a week
- Supervisor and Interviewer Feedbacks (Right Pane, Green):**
 - 1. How much did you have difficulty in getting a respond to the question?**
 - Very much
 - Somewhat
 - Neutral** (Selected)
 - Not much
 - Not at all
 - 1a. Reason(s) for having difficulty:** [Text input area]
 - 1b. Suggestion(s) for reducing difficulty for the next implementations:** [Text input area]
 - 2. Which mode of data collection do you think is more suitable for the question?**
 - Telephone (CATI)** (Selected)
 - Face to face (CAPI)
 - 2a. Suggestion(s) that can make the question suitable for the phone:** [Text input area]
 - 3. Was the question understood when it was first read to the respondent?**
 - Yes (Selected)
 - 3a. How did you probe the question/what examples did you give?** [Text input area]

At the bottom of the form, there are four buttons: "First", "Previous", "Next", and "Save and Exit".

The questionnaire form given the screenshot above consists of two parts. The first part (left gray side), namely ICT usage survey questions, is the part where the master survey questions appear and is read-only. Those completing the questionnaire (supervisors) are not expected to make any changes in this section. The second part (right green side), namely supervisors and interviewers feedbacks, is the part where opinions would be written and filled in for each question. A total of 47 questions were asked to provide feedback, the first of which was for the evaluation of overall survey and the remaining 46 questions were the master survey questions.

The order of the questions was the same as in the Survey on ICT Usage in Households and by Individuals field application.

Supervisors and interviewers feedbacks part consists of three main questions and supervisors of regional offices were separately expected to complete the questionnaire form and respond those same questions for each of the 47 questions. Those questions are:

1) How much did you have difficulty in getting a respond to the question?
For those who say “Very much” or “Somewhat” to this question asked for:

- a) Reason(s) for having difficulty and
- b) Suggestion(s) for reducing difficulty for the next implementations.

2) Which mode of data collection do you think is more suitable for the question?
For those who say “Face to face (CAPI)” to this question asked for:

- a) Suggestion(s) that can make the question suitable for the phone.

3) Was the question understood when it was first read to the respondent?
For those who say “No” to this question asked also:

- a) How did you probe the question or what examples did you give?

4.2. Implementation

The designed questionnaire was shared with the regional offices and a period of three weeks was given them to fill in the form. Feedback survey was filled out separately by each supervisor who took part in the master survey. Supervisors filled out the questionnaire in consultation with the interviewers.

The questionnaire was designed as an MS access form to get results quickly. In fact, it could have been designed under TurkStat survey portal, but this time it would be both more time consuming and more burdensome.

4.3. Results

The master survey (Survey on ICT Usage in Households and by Individuals) questionnaire consisted of three main parts: household composition (single question to create a household composition), household questions (4 questions) and individual questions (41 questions). In addition to each question included in these three parts, Regional Office Feedback Survey was designed to ask opinions for the overall master survey also. Therefore, the results are been evaluated under four group breakdowns, namely; overall survey, household composition part, questions for household part and questions for individual part.

4.3.1. Level of difficulty in receiving responses

The first question asked to measure the difficulty level in obtaining responds for each question of the master survey was: "How much did you have difficulty in getting a respond to the question?" Distribution of the responds given by regional offices supervisors to this question is presented in the table below.

Table 4. The frequency and percentage distribution of the responds to the “How much did you have difficulty in getting a respond to the question?” question

		Very much	Somewhat	Neutral	Not much	Not at all	Total
Overall Survey	Freq.	4	18	12	11	1	46
	%	8.7%	39.1%	26.1%	23.9%	2.2%	100%
Household composition	Freq.	0	4	4	30	7	45
	%	0.0%	8.9%	8.9%	66.7%	15.6%	100%
Questions for household	Freq.	3	34	41	87	16	181
	%	1.7%	18.8%	22.7%	48.1%	8.8%	100%
Questions for individual	Freq.	18	212	527	928	162	1847
	%	1.0%	11.5%	28.5%	50.2%	8.8%	100%

As can be seen from the table, more difficulties were experienced in the overall survey than survey parts (and questions in it).

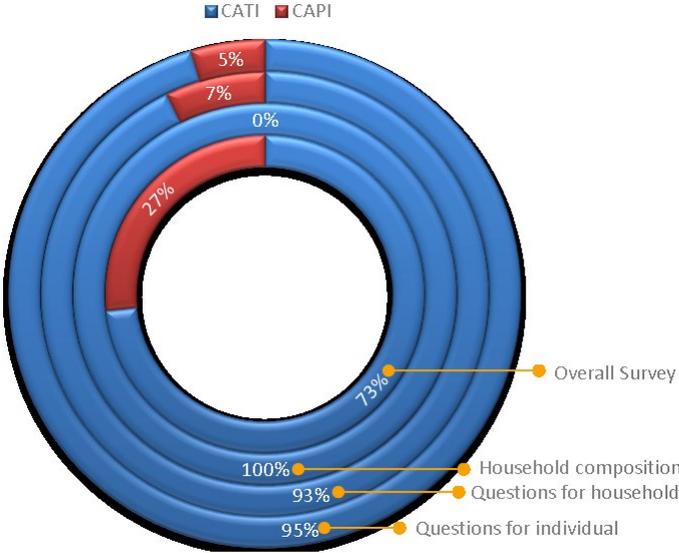
Regional offices stated the followings as the most important reasons for having difficulty in overall survey (in order of frequency):

- › The long duration of the survey implementation due to high number of questions in the survey (even more, depending on the number of individuals aged 16-74 in the household),
- › Incomprehension of the questions by respondents due to the technical nature of the survey,
- › Trust and persuasion problems due to the calls from unknown and non-corporate phone numbers and
- › Regional differences (large households, social structure, etc.).

4.3.2. Which mode is thought to be more suitable?

Regional offices supervisors were asked whether the phone or face-to-face mode was more appropriate for the overall survey and each master survey question. The percentage of the responds to this question was grouped according to the overall survey and the questionnaire parts and presented in the graphic below.

Figure 2. The percentage of the responds to the “Which mode of data collection do you think is more suitable for the question? CATI or CAPI” question

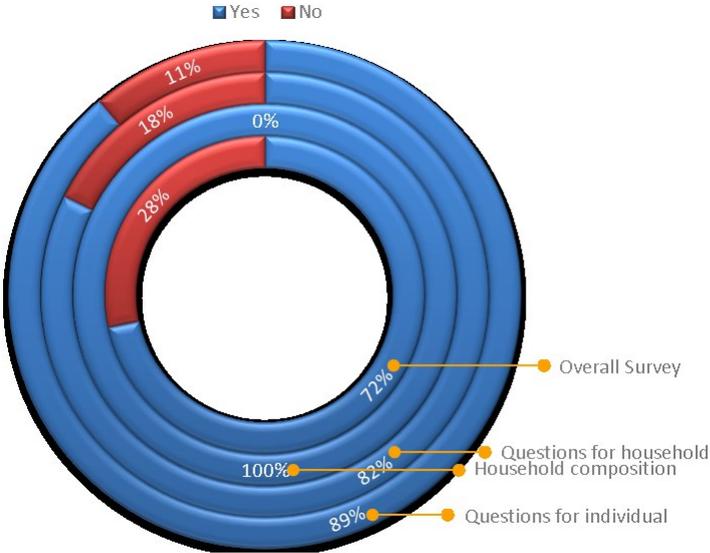


As a result, it has been stated by the regional offices that CATI is more suitable for both overall survey (73%) and for all parts of the questionnaire. Another conclusion to be drawn is that when the master survey questions were evaluated one by one, it was observed that almost all of them were said to be suitable for CATI than CAPI, while this ratio decreased slightly for the overall survey due to the difficulties mentioned in the previous section.

4.3.3. Whether the survey question was understood when the question was first read on the phone

The last question of the feedback survey questionnaire was whether the questions asked by telephone were understood when they were first read to respondent. The percentage of the responses received according to the overall survey and the questionnaire parts is shown in the following image.

Figure 3. The percentage of the responds to the “Was the question understood when it was first read to the respondent?” question



When the questions in the master questionnaire parts were considered one by one, it was stated that most of the questions were understood when they were first read to the respondents (for overall survey, this figure (72%) is slightly lower). When the questions that could not be understood when they were first read to the respondent on the phone are analyzed according to their frequency, the following two questions stand out in order of descending frequency, and it was stated that the interviewers felt the need to provide extra explanations to respondent the most for these:

- 1) (*Household question*) What types of internet connections are used at home?
 - a) Fixed broadband connections, e.g. DSL, ADSL, VDSL, cable, optical fiber, satellite, etc
 - b) Mobile broadband connections (via mobile phone network, at least 3G, using (SIM) card or USB key, mobile phone or smart phone as modem)
 - c) Dial-up access over normal telephone line or ISDN
 - d) Mobile narrowband connection via mobile phone (WAP, GPRS)
- 2) (*Individual question*) Did you use storage space on the internet (cloud computing) to save documents, pictures, music, video or other files for private purposes in the last 3 months (e.g. Google Drive, Dropbox, Microsoft OneDrive, iCloud, Amazon Drive)?

5. Conclusions

TurkStat has been planning to redesign all its national household surveys in a way that data collection is achieved via other relatively low cost modes (such as CATI, and web survey) than CAPI, provided that no reduction in the quality of the collected data. Covid-19 pandemic has contributed to this modernization process of TurkStat by adding momentum.

Before 2020, TurkStat Survey on ICT Usage in Households and by Individuals has been carried out through face-to-face interviews, which was the most widely used and traditional and also expensive and time consuming way of data collection. Via face-to-face interviews the data was traditionally collected through door to door visits of field interviewers. Transport costs and travelling time were unavoidable in this case. These costs would further increase if field interviewers had to revisit respondents.

For TurkStat Survey on ICT Usage in Households and by Individuals, it is more appropriate to proceed with the lower cost CATI mode, both in line with the positive opinions of the regional offices and because there is no negative difference observed in the quality of the data collected this year compared to the previous years.

5.1. In Line with the Positive Opinions of the Regional Offices Keeping up with CATI

According to the results of the regional offices feedback survey, a very large proportion of the regional offices preferred to continue with CATI, and therefore TurkStat Survey on ICT Usage in Households and by Individuals will be planned to continue via CATI in the coming years, too. While planning, good practices and the opinions and suggestions of the regional offices for taking action to overcome the difficulties, adapting standardized interviewing method and benefiting from CATI system will be taken into consideration.

5.1.1. Taking action to overcome the difficulties

In the feedback survey, the regional offices expressed not only the difficulties encountered but also the suggestions to overcome these difficulties, which should be taken into account in order to increase the quality of the master survey to be continued with CATI in the future. These suggestions are listed below:

- › Reducing the number of survey questions by
 - Eliminating those that have exceeded a certain percentage and become saturated and redundant,
 - Removing relatively less important sections of survey question and
 - Obtaining some of the survey variables from administrative sources.
- › Increasing the rate of invitation letters reaching sample addresses
 - Very useful in persuasion and
 - Survey questionnaire can be completed relatively easily
- › Calling respondents from corporate numbers and/or specifying the numbers to be called in the invitation letter
- › Sending the survey questionnaire to sample households along with the invitation letter in order them to be familiarized with the survey questionnaire and questions
- › Increasing the examples given to make the survey questions easier to be understood more easily by diversifying them

- › Making the questions simpler and more understandable (using examples in daily life by reducing technical terms)

5.1.2. Adapting standardized interviewing method

As a result of the feedback survey; for the questions which were difficult to get a respond or were not understood at once by respondents it has been observed that interviewers produced different solutions from each other, detailed the questions in different ways, and gave different examples to be able to explain the questions.

Considering that TurkStat Survey on ICT Usage in Households and by Individuals will continue via CATI in the coming years, standardized interviewing method becomes particularly important and should be taken into account during the design and planning phase. Standardized interviewing method consists of four elements.

- 1) Questionnaire standards
 - › How to provide structured way to compose questions, answer alternatives, introductions, descriptions etc. based for using interviewing techniques
 - › How to help interviewers realizing how to ask different kind of questions
- 2) Interviewing techniques
 - › How to ask questions
 - › How to detail definitions and introductions
 - › How to repeat question if respondent did not understand the question
 - › How to do neutral probing (not leading)
- 3) Interview interaction
 - › How to create positive and active answering atmosphere by following interviewing techniques
 - › How to keep positive attitude through the interview or how to keep intensity of the interview
 - › How to pay attention to the respondent and the interviewing situation and how to end the interview (who was calling, from where, when the results are published etc., positiveness, image of TurkStat)
- 4) Interviewing style
 - › How to articulate and emphasize
 - › How to avoid speaking manners
 - › How to achieve tone neutrality (neutral but natural style of speaking)
 - › How to behave to convince the respondent (persuasion)
 - › How to create right kind of atmosphere and correct interviewing style (pace, volume of the speaking and quality of the voice)

5.1.3. Benefiting from CATI system

CATI is not just about surveying over the phone but it is a system that offers various facilities and possibilities. While the 2020 application had been first planned to be implemented with CAPI as before, with the outbreak of covid-19, it was decided to be made by phone with an agile decision. Therefore, planning and preparation could not be made long enough about CATI. Since there was no CATI system in the regional offices, the facilities and possibilities offered by the system could not be used.

Another result that can be suggested as of this study is the construction of a CATI system in the regional offices for the implementations in the coming years and thus benefiting from the advantages of the CATI system a few of which are mentioned below:

- › Starting and ending the survey interview with respondents,
- › Sharing and controlling samples,
- › Creating an automatic appointment system (creating and managing appointments),
- › Automatic dialing and
- › Recording voices in interviews.

5.2. Almost No Negative Quality Difference Observed as A Result of Mode Change

While changing data collection mode of a survey, identifying, assessing and implementing one of the risk options namely; tolerate, treat, transfer or terminate is critical. If a mode change is to be planned to achieve in a survey, first of all, possible risk factors such as;

- › Coverage,
- › Bias that may result from selective nonresponse,
- › Break in series,
- › Non-response and
- › Proxy should be carefully evaluated beforehand.

5.2.1. Coverage

Telephone data collection inevitably requires a mobile or fixed line phone presence. In telephone surveys, under coverage will occur due to the frame consisting of call lists that do not have phones or units whose numbers are not listed. Under coverage may cause bias in results.

This risk was terminated primarily during the survey design phase by covering all settlements regardless of whether they have a phone or not. Then, the sample settlements and the phone numbers of the inhabitants were matched by linking these records with administrative data.

5.2.2. Bias that may result from selective non-response

Selective non-response may cause bias in results, too. In CAPI mode, “Field interviewers can also distribute their efforts somewhat selectively in order to achieve response. They may be less willing to visit certain neighborhoods, and they will be more reticent at some addresses than others. Tight monitoring can be used to try to curb these tendencies.” (Beukenhorst and Kerssemakers, 2012) There is no such risk in CATI as there is no door-to-door travel. However, there is a risk that households or individuals who do not have a telephone will fall within a certain region (especially relatively underdeveloped), age group (especially the elderly) or household size (especially single-person households).

To eliminate the risk of the bias caused by selective non-response in the Survey on ICT Usage in Households and by Individuals carried out by telephone in 2020, In addition to those for whom no phone number information could be obtained from the administrative sources, inaccurate or unreachable telephones were also investigated by regional office interviewers with the help of local authorities. Eventually, in the analyzes made after the end of the data

collection, no significant difference was observed compared to the previous years as can be seen from the table below.

Table 5. Average household size and age of household members for responders and non-responders of survey on ICT usage in households and by individuals in 2019 and 2020

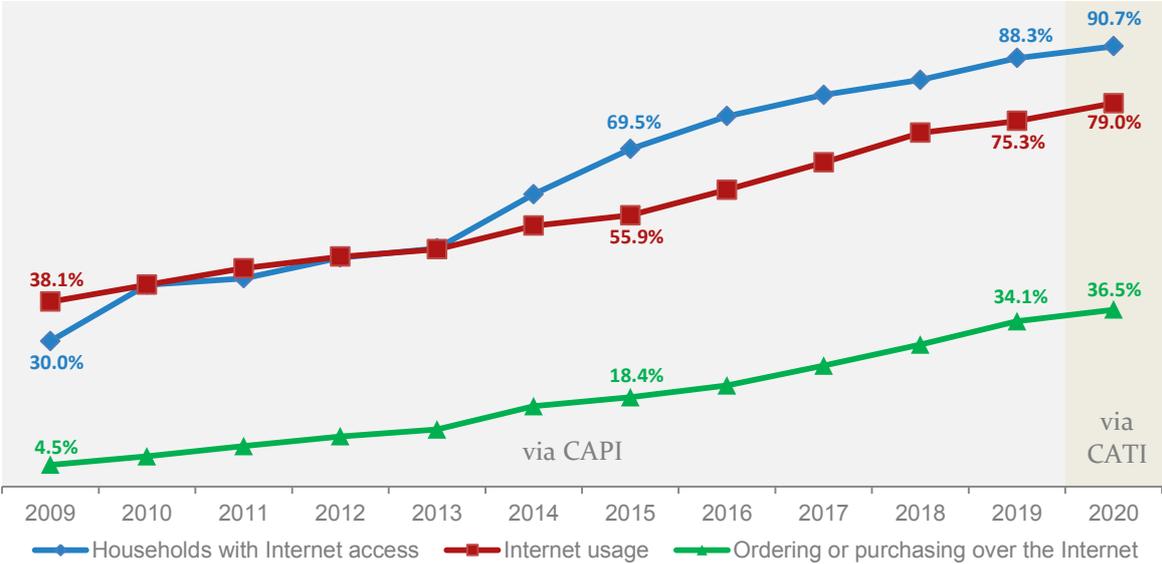
	2019 (via CAPI)		2020 (via CATI)	
	responders	non-responders	responders	non-responders
Average household size	3.91	4.60	3.93	4.55
Average age of household members	41.97	34.10	41.79	34.34

5.2.3. Break in series

The possibility that the current historical series may not be maintained was another risk, perhaps the most important.

The rate of change in all indicators announced to public came within expected margins with the exception of only one (namely, the rate of having an Internet-connected TV in the household decreased while it was expected to increase), which has also been happening in previous years for some others in which any mode changes have not taken place because this is a field survey based on sampling and there may be sampling-related errors or this may be a reality. The chart of basic indicators shared with the press release is available below.

Figure 4. Basic household ICT usage indicators shared with the press release



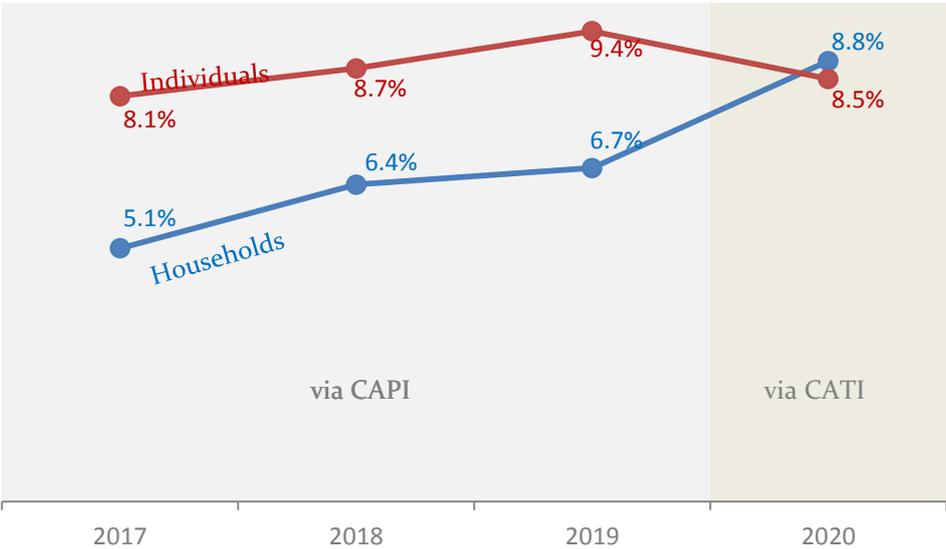
5.2.4. Non-response

It is not always possible to obtain data from all units about which data are collected. People may not voluntarily want to be in an interview, or it may be impossible to contact with them. It is crucial to achieve a satisfactory response rate at an agreed quality level, as missed observation units will result in bias and less accurate statistics.

In the case of TurkStat Survey on ICT Usage in Households and by Individuals carried out by telephone in 2020. No unexpected difference was observed in the non-response rates for both

households and individuals compared to previous years. As can be seen from the figure below, while there was an increase in household non-response in line with the trend of previous years, there was a decrease in individual non-response, which was a pleasing positive result.

Figure 5. ICT usage survey non-response for households and individuals rates change over years



5.2.5. Proxy

There is little reason to allow proxy in the CATI mode, as it is relatively easy to make a phone call, and the same is true for cost. Also, in (combination with) CATI, proxy response results in a lower data quality than CATI without proxy (or CAPI with or without proxy). For this reason, it should be considered very carefully before allowing proxy with the CATI method.

Proxy answering was not acceptable therefore it was not permitted for the survey on ICT usage in households and by individuals. Thus, the emergence of proxy risk factor was terminated from the beginning.

5.3. Cost Reduction Achieved

All stages of the statistical production process are important, but the quality of the data collection stage is of particular importance as it greatly affects the quality of the rest. TurkStat aims to collect and process information more quickly, more accurately, and more effectively than before. Due to these reasons, TurkStat always needs for an accurate and up-to-date data in very short time. Accordingly, to support the right decision-making process, TurkStat maintains trying and utilizing these new data collection modes efficiently.

Sources

- 1- [https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Digital economy and digital society statistics at regional level](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Digital_economy_and_digital_society_statistics_at_regional_level)
- 2- Lars Norqvist, (Department of Political Science, Centre for Principal Development, Umeå University) (2018) Analysis of the Digital Transformation of Society and its Impact on Young People's Lives, Sweden
- 3- <https://ec.europa.eu/eurostat/data/metadata/covid-19-support-for-statisticians>
- 4- <https://epale.ec.europa.eu/en/blog/digital-transformation-society-one-key-topics-forum-future-learning#:~:text=It%20should%20be%20based%20on,for%20accessing%20and%20undergoing%20retraining.>
- 5- https://ec.europa.eu/eurostat/documents/45004/9826397/Peer+Review+Report+of+Turkey_final.pdf/3oficod3-0000-4c4a-83e0-cc4187bb76a3
- 6- <https://ec.europa.eu/eurostat/statistics-explained/pdfscache/2787.pdf>
- 7- Beukenhorst, B. and Kerssemakers, K., (Statistics Netherlands) (2012) Data Collection Strategy General Description and Unimode Designs, The Hague/Heerlen