

CONFERENCE OF EUROPEAN STATISTICIANS

Workshop on Statistical Data Collection

WP.5-3

10-12 October 2017, Ottawa, Canada

22 August 2017

How to motivate and catch hard-to-reach groups in household surveys

Vesa Virtanen and Heidi Kangasniemi (Statistics Finland)

vesa.virtanen@stat.fi, heidi.kangasniemi@helsinki.fi

Abstract

Response rates in direct data collections are going down in both voluntary household (e.g. in LFS 68 per cent) and statutory enterprise surveys (on average 78 per cent) in Finland. In some challenging surveys with short data collection periods the non-response rate can be near 50 per cent (like in the Consumer Confidence Survey). Especially the not-contacted group is growing. The data collection infrastructure is changing: mobile data collection will liberate surveys from time and place and flat screens, new e-identification formats for contacting citizens will take place. Mixed-mode and web surveys are becoming common. New ways to contact citizens have also been put into operation. It can be done by email, text-messages, mobile leaflets, responsive mobile questionnaires, motivation videos, infographics, contacting at different times of the day, and through social media. Statistics Finland goal is to keep the survey response rates and quality at a high level and keep the respondent burden as low as possible, e.g. by testing forms, concepts and questions. In our report, we were evaluating the benefits and disadvantages of incentives as a whole. Respondent profiles and service paths are a fresh approach as well. In this article, we will evaluate how we have succeeded in collection and motivating work in recent years and what new ideas are now in practice.

How to motivate and catch hard-to-reach groups in household surveys

Virtanen Vesa* and Kangasniemi Heidi**

* Statistics Finland, vesa.virtanen@stat.fi

** Statistics Finland, heidi.kangasniemi@helsinki.fi

Abstract: Response rates in direct data collections are going down in both voluntary household (e.g. in LFS 68 per cent) and statutory enterprise surveys (on average 78 per cent) in Finland. In some challenging surveys with short data collection periods the non-response rate can be near 50 per cent (like in the Consumer Confidence Survey). Especially the not-contacted group is growing. The data collection infrastructure is changing: mobile data collection will liberate surveys from time and place and flat screens, new e-identification formats for contacting citizens will take place. Mixed-mode and web surveys are becoming common. New ways to contact citizens have also been put into operation. It can be done by email, text-messages, mobile leaflets, responsive mobile questionnaires, motivation videos, infographics, contacting at different times of the day, and through social media. Statistics Finland goal is to keep the survey response rates and quality at a high level and keep the respondent burden as low as possible, e.g. by testing forms, concepts and questions. In our report, we were evaluating the benefits and disadvantages of incentives as a whole. Respondent profiles and service paths are a fresh approach as well. In this article, we will evaluate how we have succeeded in collection and motivating work in recent years and what new ideas are now in practice.

1 Hundred-year-old Finland in the light of statistics

Independent Finland will become 100 years old on 6th December 2017 – very soon by the way. When Finland gained independence in 1917, the Central Statistical Office of Finland, presently Statistics Finland, had already for over 50 years collected statistical data in Finland and from Finns. At the threshold of Finland's centenary celebrations, Statistics Finland collected a list of things in which Finland is one of the best in the world. The comparisons speak for themselves: Finland is a good country.

Finland is the most stable country in the world. The Fund for Peace, [Fragile States Index 2017](#)

Finland is the second best country in social progress in the world. The Social Progress Imperative, [2017 Social Progress Index: Finland](#)

Finns' trust in other people is the second highest in Europe. Eurostat, [Average rating of trust](#)

Finland is the second most gender equal country in the world. World Economic Forum, [Global Gender Gap Index 2016](#)

Finland is the second best in the world in using information and communication technologies to boost competitiveness and well-being. World Economic Forum, [Networked Readiness Index](#)

Finland has the world's third lowest infant mortality rate (deaths at the age of under one year). UN, SDG Indicators, [Infant mortality rate](#)

See more: http://tilastokeskus.fi/ajk/satavuotiassuomi/suomimailmankarjessa_en.html



Fig 1. The logo of the 100 years old Finland

2 New communication channel for citizens and respondents

There is a lot of expectations for the new digital suomi.fi [“finland.fi”] service to be launched in Finland at the end of this year. It presents an opportunity for household surveys as well. Suomi.fi e-Identification enables the citizens – and respondents – to be recognised in a secure way by using various identification media such as banking IDs and mobile certificates (<http://tutustu.suomi.fi/en/>). We hope that it will be the same kind of success as the My Kanta pages where every citizen can find their health records and electronic prescriptions in one place (<http://www.kanta.fi/en/kanta>).

Suomi.fi Messages enable digital communication for authorities – and statistical authorities as well – regardless of the end-user’s preferred communication channel and message format. For Statistics Finland it will be a new communication channel for communicating with respondents.



Fig 2 The new Suomi.fi online has services for all citizens in all life situations

3 Respondents have more options to answer – response rates still dropping

In this article we focus on household surveys and, especially, respondent groups with low activity and, e.g. service paths for them. The overall survey situation has partly improved and partly worsened for statistical data collections in Finland since last year when the UNECE data collection workshop was held. Better in the means of data collection methods: more data is gathered through registers, other databases and also by utilising big data. There is a continuing development in household surveys to shift from CATI (computer assisted telephone interviewing) to mixed-mode (web + CATI) data collection. This is done to provide respondents more options in response modes and to cut costs. So far the road has been rocky.

In direct household data collections, about 200,000 households supplied information to Statistics Finland in 2016. The main data collection method in Statistics Finland’s household surveys was telephone interviews. In 2016, the response rate was about 68 per cent in the Labour Force Survey, 60 per cent in the Survey on Income and Living Conditions, about 52 per cent in the Consumer Confidence Survey, and 52 per cent in the Use of Information and Communications Technology (mixed-mode) Survey. All these figures are lower than one year ago. Over the past 15 years, response rates have declined in household surveys by about one per cent per year (e.g. LFS) [1, 2, 3, 4].

E.g. in the Use of Information and Communications Technology Survey, people aged 20 to 29 were least active to respond. In the Labour Force Survey the response rate is significantly lower if the

respondents have a lower education level and the youngest age groups have the highest non-response rates. In the Consumer Confidence Survey, over 70 per cent of the non-response was due to non-contact. Non-contact has grown most in the younger age groups (15 to 44 years) [1,2,3]. To guarantee good data for statistical purposes especially among young and less educated persons, a lot of effort has been put in the field of data collection methods, data collection management, services and communication channels.

4 Respondent groups and service paths

In the near future, most of the answers to household surveys assumedly come by web. This requires functional survey technology and easy-to-use user interfaces. The user and respondent experience has to be pleasant. Otherwise respondents disappear from the telephone, mobile device, laptop or PC portal and the survey pages. Contacting respondents varies as well. Currently, almost everyone gets a letter, some of them get a reminder, a text message, a telephone call, or an email, some of them are contacted by an official e-identification interface and some of them are still contacted face-to-face. How to handle so many different data collection, contacting and interface paths? One way is to think more deeply of the user, user experience and user services.

Statistics Finland is at the moment remodelling its web services and web pages. This remodelling was based on previous user studies (9) and analytics. According to these studies, 6 groups – different kinds of user personas – were identified: Influencers (politicians, leaders and activists), Analytics (researchers and economists), Communicators (media and consultants), Adapters for everyday life (regular citizens, students and home owners), Organisation information users (job seekers etc.) and Respondents.

According to Karhu [5], users have different backgrounds and attitudes. They use different channels and devices to get in touch with Statistics Finland, also when sending data and responses. They need to use different kinds of services and they have different kinds of goals and problems when using the web pages. They vary in patience and require different things from the user experience. Different user personas need different service paths. We have to know what needs to happen before the user starts to use the services. What does the user do? What material does the user need? What are the high points and low points of the user experience? What are the greatest impact points? Improving these points have the biggest effect on the total user experience – and response rate as well.

After this web page (stat.fi) user group work, the writers of this article continued to create more respondent groups. We were especially interested in the hard to reach respondent groups we found through the non-response PxWeb database and non-response visualization application [4].

Service design is the activity of planning and organising people, infrastructure, communication and material components of a service in order to improve its quality and the interaction between the service provider and its customers. The purpose of service design methodologies is to establish best practices for designing services according to both the needs of customers (respondents) and the competencies and capabilities of service providers (survey or statistical organisation).

The aim is to improve the service path for both the respondent and the survey organisation employee. During the data collection period there is a certain number of respondent touch points and the respondent journey from start (e.g. invitation to the survey) to the finish (happiness with the service as a whole or feelings after filling the questionnaire) is a multistep adventure.

5 Hard-to-reach respondent groups

According to this respondent service path idea we created five hard-to-reach respondent groups. We selected only young respondents to this examination because they have the lowest response rates and something has to be done to turn the tables.

This model contains the following elements:

- Picture and personal information: Characteristic (illustrational, not actual people) photo and name for the respondent type and his/her age, title and working sector.
- Service Channels: service channels and devices used by the respondent type.
- Services: services of Statistics Finland used by the respondent type.
- Digitality: utilisation degree of devices (such as computers or smart phones) and services (such as online services and applications).
- Clicks: The amount of mouse clicks, the respondent type expects to have to make to get to log in (3 = as easy as possible, 5 = chooses time for phone or personal interview, >9 = prefers strong identification, as many clicks as is needed), and descriptive adjectives for expectations related to the user experience of the services.
- Story: descriptive story of the life situation of the respondent type.
- Citation: slogan of the respondent type/ descriptive citation for one's attitude.
- Goals: Reasons why a respondent contributes.
- Problems: respondent type's problems related to a statistical investigation.
- Attitude: descriptive attributes related to attitudes towards responding.
- User experience: the amount of mouse clicks, which the respondent type expects to do to get to the login page, and descriptive adjectives for expectations related to the user experience of the services.
- Message of Statistics Finland for the respondent type: short message from Statistics Finland for the respondent type.

We named five respondent types to illustrate young hard-to-reach respondent groups. These were *Adam Apathetic*, *Susan Suspicious*, *Daniel Digital native*, *Irene Immigrant* and *Michael Marginalized*.

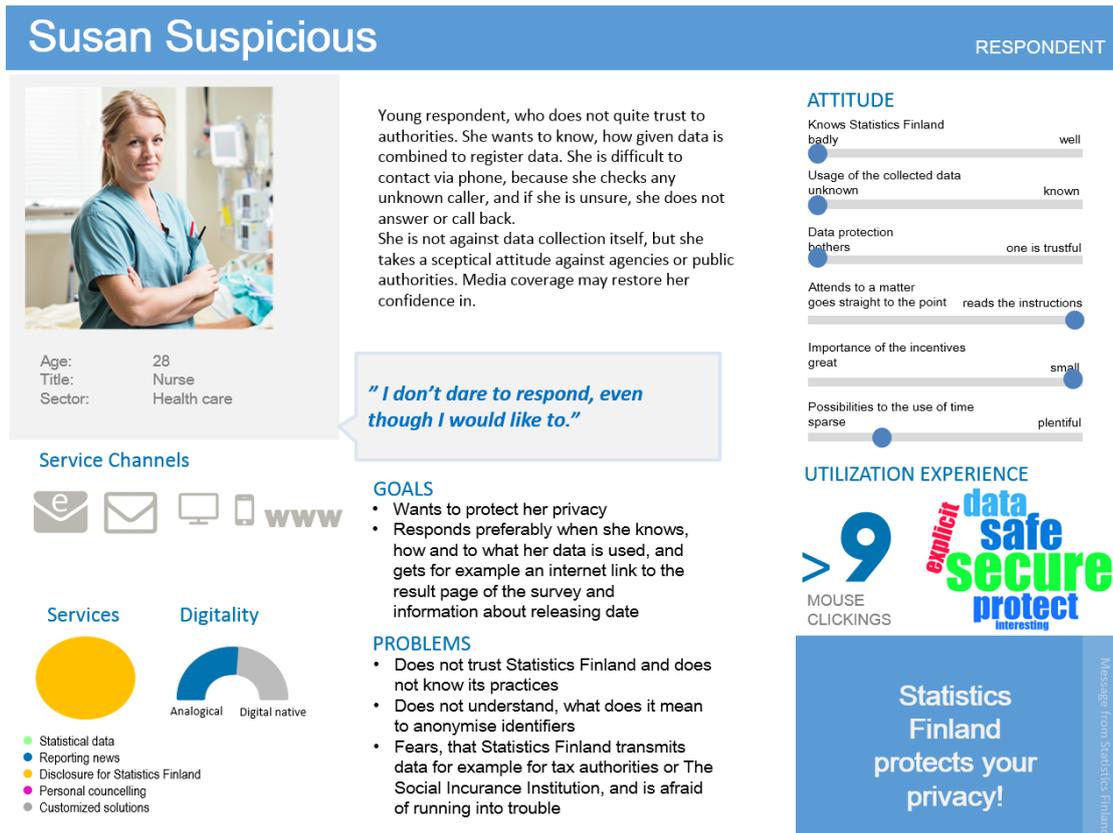


Fig 3 One example of low activity respondent characteristic (illustrational, not actual people)

For example we describe Susan Suspicious like this: *Young respondent, who does not quite trust authorities. She wants to know, how provided data is combined with register data. She is difficult to contact via phone, because she checks any unknown caller, and if she is unsure, she does not answer or call back. She is not against the data collection itself, but she has a sceptical attitude towards agencies or public authorities. Media coverage may restore her confidence.*

GOALS: *She wants to protect her privacy. Susan prefers to respond when she knows how and for what her data is used, and gets, for example, an Internet link to the result page of the survey and information about the releasing date.*

PROBLEMS: *Susan does not trust Statistics Finland and does not know its practices. She does not understand, what it means to anonymize identifiers. Fears, that Statistics Finland transmits data, for example, to tax authorities or the Social Insurance Institution, and is afraid of running into trouble.*

Message of Statistics Finland to the respondent type is: Statistics Finland protects your privacy! We should communicate this information in our data collection materials through the service channels she is using.

6 The use of incentives in household surveys – benefits and drawbacks

Incentives are used in inquiries and interviews to increase response rates. Incentive use has not been a widely adopted practice at Statistics Finland. As response rates have continued to decline in surveys, especially among young and less educated groups, a new need has risen to look into the ways in which incentives could be used more systematically [6]. Incentives can be classified according to their different features: monetary or non-monetary, prepaid or promised (sent initially or contingent on the returned response) and tailored or non-tailored (tailoring meaning a strategy where some groups are rewarded and others not) [7,8].

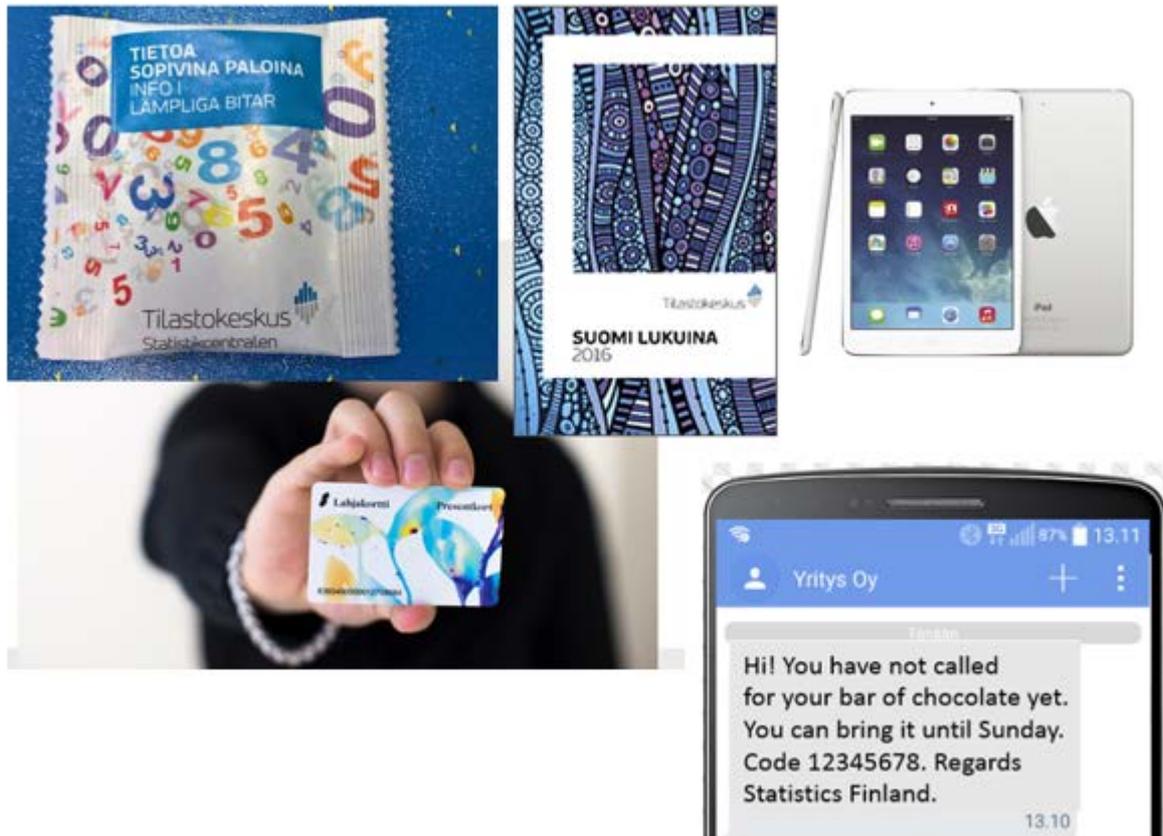


Fig 4 Examples of incentives (chewing gum, pocket statistics, iPad, gift token, and promotional code)

Factors, which people say cause them to answer in surveys, can be classified into three wide categories according to Singer and Ye: altruistic, egoistic and reasons connected to the survey. Altruistic causes include reasons such as believing in the importance of the study (e.g. employment), wanting to be helpful, civic duty, etc. Egoistic reasons include things like enjoying answering to surveys, interest in results or monetary benefit. Reasons connected to the survey are related to the organisation in question, the topic of the survey, etc. Incentives increase response rates most in web or postal surveys.

At Statistics Finland, the use of incentives in data collections has been, in the past, viewed as ethically dubious and possibly harmful to data quality and to the reputation of the office. There has been a strong emphasis on “civic duty” in responding and providing information, especially for official statistics. However, recently there has been a shift in attitudes towards the use of incentives.

One type of incentive under development at Statistics Finland is feedback information. The idea is to provide information to the respondents as a way to reward them for their participation.

The benefits and drawbacks of incentives are:

The up-sides of using incentives:

- o In general, they increase response rates (regardless of the mode of the data collection).
- o They may improve data quality in some cases (better item response, longer answers to open-ended questions, decreased sample bias).
- o The use of incentives can be motivational for interviewers, since it is a tool for them to convince the respondent to participate.
- o Incentives may shorten the time period needed for data collection and thus decrease the costs.
- o Incentives are a means of communicating with the respondent and can enhance recognition of the statistical office. Incentives can communicate to the respondent how important their participation is regarded by the surveyor.

The down-sides of using incentives are:

- o Incentives have complex effects – these effects cannot always be estimated before the data collection because of the multiple factors involved.
- o The cost-benefit ratio can be quite small in promised lottery-type incentives (studies show that their impact on response rate is quite small, some three to four percentage points on the whole, and they vary between age groups; in the youngest age groups lotteries may decrease response rate).
- o Lower data quality (item non response, biased net sample). Biased net sample seems to be the most crucial aspect when it comes to data quality.
- o Negative aura of marketing (loss of reputation to the provider)
- o There are laws and regulations that need to be considered when adding incentives to a data collection.
- o Ethical questions: are incentives fair use of money, ecological sustainability (when sending small tokens such as ballpoint pens)

At the moment, our suggestions for future are: forming best practices, developing feedback information practices, testing the effects of incentives on item non-response, sample composition and other quality measures, testing the cost-benefit ratio of lottery incentives, testing the effects of tokens used in advance letters, centralised bidding on “envelope puff ups”, i.e. small tokens provided with the advance letter and testing promotional codes where respondents – especially when responding by web – gets a code to pick up a gift (chocolate, ice cream or so on). All in all, the use of incentives is very case-specific. Therefore, general guidelines should be broad enough to allow tailoring the use according to the needs of the data collection in question. Instead of ‘guidelines’ the term ‘best practise’ describes what is needed here.

7 New ways to communicate with surveys

At Statistics Finland we have been putting into operation new ways for strengthen the image of collecting data for statistical purposes. Using videos is one method to illustrate to citizens in general and a respondent personally why and for what data are collected. Here are some examples:

We all are statistics: <https://www.youtube.com/watch?v=OzUgc4G1Fdk>

Survey-specific videos have also been used, for example, in the Use of Information and Communications Technology Survey (in Finnish): <https://www.youtube.com/watch?v=yi-FVypoHKE>

Respondents find these videos when they are responding to the survey through the data collection webpage or the link is provided in the information letter of the survey or via Twitter, etc. Infographics are an effective way to concretise the purpose of statistics and data collection. Written text, for example, in many advance letters may be too long and numbing, especially for some young people. An infographic is a graphic that gives you lots of info and stimulus. Here is one example of an infographic, which should encourage respondents to answer to the survey. This infographic is distributed to respondents or at webpages and on Twitter, etc.



Fig 5. One example of an infographic interconnected to the data collection material

The aim is to share common knowledge of the meaning and use of statistics and data collection.

8 Conclusion

Over the past 15 years, response rates have declined in household surveys by about one per cent per year (e.g. LFS). In some surveys, the trend has been steeper. The situation is most difficult among young people, men and less educated respondents. New ways of communication is needed and has been put into operation. Service design, respondent profiles, e-identification, new digital communication channels for communicating with respondents, videos and infographics about data collection are needed. The production of reliable statistics requires this kind of data collection efforts and management.

References

- [1] Virtanen Vesa – [1] Virtanen Vesa – Sarén Nina (2016). Non-response Review 2016: Individuals and Households. Internal report. Statistics Finland.
- [2] Virtanen Vesa – Harjunmaa Hanne (2016): Enterprise as a respondent – Response rate trends in enterprise surveys in 2015. Internal report. Statistics Finland.
- [3] Virtanen Vesa – Lehtinen Milla (2014): Annual Nonresponse Review 2013 – Individuals and Households. Statistics Finland.
- [4] Virtanen Vesa – Oikarinen Jukka (2016). Ensuring the data collection quality of enterprise and household surveys – an overview and new ways. UNECE Workshop on Statistical Data Collection 2016.
- [5] Karhu Laura (2017), Remodeling the web services, Nordic web meeting 2017, Oslo 2017, <http://www.dst.dk/extranet/nordweb/>
- [6] Sarén Nina (2016). Incentives in Statistics Finland. Project work as a part of annual nonresponse report at Statistics Finland.
- [7] Singer Eleanor – Conger Ye (2013). The Use and Effects of Incentives in Surveys. *Annals, AAPS*, vol. 645, January 2013. 112 - 141.
- [8] Lund Kari-Anne – Gulløy Elisabeth (2016). The use of incentives in household surveys: a future strategy to save response rates, or a relic of the past? Statistics Norway. Nordic Statistics Meeting 2016. Nina (2016). Non-response Review 2016: Individuals and Households. Internal report. Statistics Finland.