

Data falsification in the European Social Survey?¹

Ineke Stoop (SCP), Roberto Briceño-Rosas (GESIS), Achim Koch (GESIS), Caroline Vandeplass (KU Leuven)
October 2018 (version 4)

1. Survey quality in the European Social Survey and data falsification as a threat	2
2. ESS methods, procedures and structure	3
3. How to detect falsification	4
4. How to prevent falsification	7
5. Falsification, undesirable interviewer behaviour and interviewer effects	9
6. Where are we now in the ESS	12
Appendix A: Recent literature on interviewer falsifications	13
Appendix B: Studies on ESS data: interviewer effects and undesirable interviewer behaviour ..	14

¹ This paper is based on discussions within the Core Scientific Team (CST) of the European Social Survey, analyses conducted by the ESS-ERIC KU Leuven team, comments from members of the ESS-ERIC Methods Advisory Board, discussions with external scientists, and discussions at an AAPOR session on falsification.

1. Survey quality in the European Social Survey and data falsification as a threat

The European Social Survey (ESS) has been founded to provide accurate data about the values, opinions, attitudes and beliefs of Europe's citizens, to provide a bedrock of hard information which is both reliable and extensive. Academics, statisticians, politicians, legislators, policy makers, business leaders and ordinary citizens should be able to trust these data

As Europe faces grand societal challenges including immigration, changing family and relationship arrangements, climate change, welfare reform, declining political trust, increased populism and persistent health inequalities (amongst others), the ESS provides robust data that illuminate changes and stability in the social fabric of Europe. To do this, it has set high quality standards, focuses on comparability across countries and over time, and has implemented a methods and methodological research programme to ensure that the infrastructure remains 'state of the art' and helps to position Europe as a global leader in terms of comparative social science.

As every survey, the ESS has to cope with errors such as sampling errors, coverage errors, nonresponse bias and measurement errors. The ESS specifications, procedures and guidelines aim at minimising these errors, often from a total survey error perspective, and to make rational choices in survey design acknowledging trade-offs between different types of errors. In addition, one of the core values is to be transparent about errors, and give users access to questionnaires in every language, paradata (including contact form data recording the entire interview process), and analyses of sampling and measurement error and nonresponse bias.

Less available is information on data falsification in the ESS. AAPOR (2003) defines interviewer falsification as the "intentional departure from the designed interviewer guidelines or instructions, which could result in the contamination of the data." Increasingly, attention is paid to other agents of falsification, i.e. supervisors or organisations.

The aim of this paper is to fill this gap and to provide data users with all available information about the prevention and detection of data falsification in the ESS at present and in the future.

Falsification by the interviewer, according to Robbins (2018) comprises:

- Fabricating all or part of an interview: the recording of data that are not provided by a designated survey respondent and reporting them as answers of that respondent,
- Deliberately misreporting disposition codes and falsifying process data,
- Recording of a refusal case as ineligible for the sample,
- Reporting a fictitious contact attempt,
- Deliberately miscoding the answer to a question in order to avoid follow-up questions,
- Deliberately interviewing a non-sampled person in order to reduce effort required to complete an interview,
- Otherwise, intentionally misrepresenting the data collection process to the survey management.

Falsification by the organisation comprises:

- Fieldwork supervisor who chooses not to report deviations from the sampling plan by interviewers,
- Data entry personnel that intentionally misrecord responses,
- Members of the firm itself who add fake observations to the data set,

- Fabricating questionnaires,
- Duplicating questionnaires.

In its report on [interviewer falsification](#) AAPOR (2003) stated: “The literature suggests that where appropriate methods are used, interview falsification is rare, involving only a small percentage of interviewers and a substantially smaller percentage of interviews.” According to more recent literature (see Appendix A), however, interviewer falsification can pose a serious problem in terms of data quality (see also the special issue of the [IAOS journal](#)).

Interviewer falsification can have serious consequences on survey quality. It can affect all areas of the survey life-cycle in which interviewers are involved and increase the size different types of errors. Sampling errors will be affected if selection probabilities cannot be calculated correctly if respondents are not selected according to the prescribed random mechanism. Nonresponse bias will be affected if incorrect disposition codes are assigned to sample units, and measurement errors will be affected if interviewers skip questions or misrecord answers to avoid follow-up questions. Fabricating interviews (fully or partially) do not only affect one the fabricated cases, but can have an impact on the overall data quality by decreasing or increasing variation of resulting statistics, and affect survey estimates used for research. Falsification in general will lead to erroneous survey results, potentially misinformed policy decisions and decreasing trusts in surveys in particular and social science in general.

For these reasons, the European Social Survey is increasingly paying attention to methods of detecting and preventing falsification. Many of these efforts take place behind the screen. Discussion at the 2018 Denver meeting of the [new AAPOR data fabrication task force](#), and the publication of a recent paper by Blasius and Thiessen (2018), who claim they identified fraudulent interviewers through statistical methods in a number of ESS countries in the 2010 survey, encouraged the Core Scientific Team to give an overview of the ESS present practices in this area and also to set up a work package on falsification in the new ESS work programme.

As an introduction, Section 2 of this paper gives a short overview of ESS methods, procedures and structure. Almost all of the underlying information and documentation can be found at the [ESS website](#). Section 3 lists how falsification can be detected. Section 4 gives an overview of measure to prevent falsification in a cross-national study. As “odd” answer patterns may have many possible causes, only one of them being falsification, Section 5 gives an overview of undesirable interviewer behaviour and interviewer effects that lead to answer patterns that could incorrectly suggest falsification. Section 6 summarises where we are now in the ESS and what could and should be done to minimise the risk and consequences of falsification and to enhance data quality and comparability across countries.

2. ESS methods, procedures and structure

Face-to-face surveys are often seen as the gold standard in survey research. They may have some disadvantages – they may be less suitable to elicit sensitive information from respondents, they are expensive and fieldwork may take a long time – but the number of advantages is large. Response rates are higher than in other modes. Interviewers are especially effective when the survey is long, more than one person in the household has to be interviewed or when additional information has to be collected. Trustworthy interviewers can invite people to share their thoughts and ideas, and in longitudinal surveys a rapport is established between respondent and interviewer which can enhance the participation in subsequent rounds and the quality of the data. Satisficing, i.e. giving the answer that requires the least

effort, is usually less likely in face-to-face surveys than in self-administered surveys. Face-to-face surveys also make it possible for the functionally illiterate and people without internet access or skills to participate. In addition, a key factor in a cross-national survey, face-to-face is the only mode that can be used for all groups and is applicable in all countries.

As the ESS intends to be representative of the entire populations across the diversity of the European societies, it relies heavily on the quality of interviewer behaviour in the different countries.

In the ESS it is the role of the Director and the Core Scientific Team (CST) to set standards, provide guidelines, control quality, report on deviations, and suggest improvements. A large number of measures are in place to enhance quality, e.g., guidelines for recruiting respondents, extensive testing of the questions, guidelines for interviewer briefing, monitoring of fieldwork, and analyses of the process of obtaining interviews (using the contact forms data) and the answering process during the interviews (see for an overview of these analysis Appendix A).

The countries participating in the ESS each appoint their own National Coordinator (NC) and select their own survey agency. The ESS requires the NC to be familiar at first hand with survey methodology and procedures, and to be willing to oversee the work of the survey agency including the progress of fieldwork. In addition, ESS requires the survey agency appointed in each country to be capable of, and have a track record in, conducting national probability-based surveys to the highest standards of rigour by means of face-to-face computer-assisted personal interviewing (CAPI).

Over time fieldwork in the different countries has been conducted by a variety of organisations: statistical offices, major market research organisations, not-for-profit organisations, and universities. Several of these organisations adhere to their own quality frameworks. Statistical offices adhere to the European Statistics Code of Practice, many survey agencies are ISO-20252 certified. All survey agencies have to adhere to the [Declaration on Ethics of the International Statistical Institute](#).

It has been considered whether the ESS should appoint one single survey agency to conduct the fieldwork in all participating countries. There is no survey organisation, however, that has good CAPI capacities in every participating country, and moving to one agency would also mean that the ESS would lose national agencies that do a very good job.

Thus, the ESS is a face-to-face survey, relying heavily on interviewers, in a decentralised structure combined with intensive centralised quality assessment, control and reporting. The decentralised structure, the distance between interviewers and the CST, and the long time period between fieldwork and quality assessment, can make it difficult to detect and prevent data fabrication.

3. How to detect falsification

Robbins (2018) gives a complete overview of detecting falsification. He distinguishes the following general methods:

- Observational Methods
- Recontact
- Data Analysis Methods (see also Birnbaum, 2012)
- Comparing Results to Benford's Law

- Unusual Patterns in the Data
- Rare Response Combinations
- Undifferentiated Response Patterns
- Short Paths through the Survey
- Missing Data or Incomplete Interviews
- Use of Paradata
 - Interview duration
 - Duration between interviews
 - Close to deadline
 - Time of day
 - Surge of interviews
 - Missing phone numbers

According to Robbins (2018, p. xxx) “..... statistical approaches offer important leverage and can identify observations that are more likely to have been falsified. Yet, many do not offer definitive proof that fabrication has occurred. As an example, even in cases where the valid and fabricated observations are known, it is not possible to perfectly predict whether the data are real or fraudulent. Menold and Kemper (2014) performed an experiment where they collected real survey data and then had the interviewers fabricate the same data to look for differences. As they sought to identify techniques to separate the real data from the fabricated data, they noted that no method guarantees complete accuracy. At best, their use of multiple measures predicted whether an interview was real or fabricated with about a 75 percent rate of accuracy.”

In the ESS a number of these methods have been implemented, either centrally by the CST, or by the survey agency. At present little information is available on the detection of falsification by the survey agency before the delivery of data to NSD, the central data processing agency and ESS archive. In 2018 two meetings have been organised to assess the size and impact of this problem, namely a special session of National Coordinators and a Field Directors Meeting. Not all countries attended these meetings. At these meetings it turned out that survey agencies very occasionally found out cases of data fabrication and took measures.

So what has been done, and what can be improved in the ESS in the area of detection of falsification:

a) Recontacts: Back-checks

In the ESS it was and is obligatory to call back on 10% of the respondents and 5% of the ineligibles (see section 10.7.3 of the [Specifications for Round 9](#)). Up till now only general recommendations were given on which information to collect during back-checks, and what to report on. There are also no rules on how to select the cases to be re-contacted, and the back-checks are conducted by the survey agency.

An analysis of information on the back-check process showed that procedures differ greatly among countries, that definitions are not entirely clear, and that there is no central information on the results of the back-checks, except at a very basic level, and on the measures that are taken nationally, as a result of the back-checks.

In 2019 the back-check procedure will be reviewed and tightened.

b) Analysing paradata: Contact forms

The ESS contact forms keep track of the recruiting behaviour of the interviewers. The information is often collected on paper and available for analysis well after the fieldwork period. In R9 a Fieldwork Monitoring System App and Portal will be used, making it possible to monitor fieldwork progress on a weekly base and to detect anomalies in a timely manner, during fieldwork rather than after data delivery, to allow for interventions. It is expected that in Round 10 only digital contact forms will be allowed.

c) Respondent sample composition

The composition of the final respondent group can be checked with internal and external criteria (see Koch, 2016). Deviations from expected patterns can be an indication of undesirable interviewer behaviour. Koch found that these deviations were larger when sampling frames were used in which the interviewer had more leeway to select the designated respondents.

In countries with individual samples it is also possible to check whether the birth date from the sample frame is identical to the birthdate provided by the respondent. Countries that perform this check do not provide the complete birth date to the interviewer.

d) Time stamps

As of Round 9 only CAPI interviewing is allowed, and paper questionnaires will not be used anymore. This makes it possible to collect time stamps during the interview and monitor the progress through the survey. The speed with which questions, blocks and the entire questionnaire are answered can give an indication of undesirable interviewer behaviour (speeding, skipping introductions) or just plain data fabrication. In addition, time stamps can be linked to the contact data. It is to be expected that in the future more detailed time stamps will be collected, and that these will be analysed already during data collection.

e) Data quality

After fieldwork analyses are conducted to identify data quality issues such as straightlining, odd patterns, inconsistent outcomes etc. These issues have been found in the ESS, as can be seen from the list of reports and papers mainly produced by KU Leuven in Appendix B. The ESS recommend conducting these analyses on intermediate files during fieldwork, but at present this is not a requirement.

f) (partial) Duplicates

An analysis of partial duplicates is conducted by NSD after fieldwork. Based on their detection of duplicates additional analyses have been conducted in other areas (speeding, inconsistent patterns) by KU Leuven. Based on the combined evidence all cases from interviewers who produced suspicious cases have been removed. This has happened in three countries in the Round 8 of the ESS.

g) Observation: Audio-recording

Audio-recording of interviews is an effective method of detecting fraud. This is technically possible because the ESS is entirely CAPI. Privacy issues (the consequences of the new GDPR are not yet entirely clear in all participating countries), and concerns about a decrease in response rates, are the main reasons why this has not been made mandatory in the ESS yet.

A number of survey agencies in individual countries do tape at least a number of interviews. Listening to audio files is a very time-intensive activity. It can be expected, however, that in the near future an

analysis of audio recording can be automated. In addition, it will not be necessary to listen to all interviews, but only when there are indicators of fraud. Ideally, listening to audio files should take place when the survey is still in the field.

h) GPS-tracking of interviewers

A final means of detecting fraud is when GPS-trackers are used. A number of countries do use GPS-trackers to monitor their interviewers, but the resulting information is used on an operational level and not centrally available. GPS-tracking will be recommended in the ESS, but it is not foreseen that it will be mandatory in the near future.

All measures described above are not fool proof. Interviewers can say the respondent refused to provide a telephone number (if this happens a lot to a particular interviewer this could by itself be an indicator of fraud), direct back-check calls to the phone of a friend, can complete questionnaires in their car in front of the respondent's house, can reset the clock of their laptop, etc. Still, more measures can, will and should be taken than in the past, because undetected data fabrication is a serious threat to data quality and – if finally detected – to trust in surveys.

4. How to prevent falsification

At least as important as the question how to detect falsification is the question how to prevent falsification. Robbins (2018) distinguishes between measures that can be taken before and during fieldwork. Before fieldwork begins the survey questionnaire is a key issue (length, sensitive, complex questions). The impact of the questions can differ across countries. In addition, the training is important and the remuneration schemes. During fieldwork oversight is critical, which can be especially difficult in multi-country, decentralised studies. Intermediate data files can help to detect falsification halfway, and prevent future falsification during the remainder of the fieldwork. In theory, falsification could be detected in real time, making it possible to retrain or replace falsifying interviewers. One problem is that in many cases a number of interviews by the same interviewer are required to detect fraud.

One assumption is that the more measures in place to timely detect falsification, and the more interviewers, supervisors and organisations are aware of these measures, the less likely it is that data fabrication will occur. Every measure can be circumvented, and no method is guaranteed to prevent fraud, but an increased awareness of the risk of undesirable behaviour and possible consequences when it is found out will certainly have an effect.

In the ESS the following measures to prevent falsification have been implemented.

a) Transparency

The ESS makes all survey information available, including standards, procedures, training material, and guidelines and quality reports, and also including contact form data and interviewer IDs (that cannot be related to individual identities). This makes it possible to identify odd patterns in recruitment, interviewer effects, etc., and makes the ESS vulnerable to criticism from external partners. The CST don't think this is a threat but an asset, and welcome third partners to analyse our data, point to weaknesses and make it possible to improve data quality.

b) Openness and trust

Openness about possibilities of undesirable interviewer behaviour and falsification is a first step towards improvements. In sessions with NCs and survey agencies in 2018, where representatives of a larger number of ESS countries were present, this topic has been discussed in an open, nonaccusatory way. In these discussions it turned out that at a national level many steps have been taken to detect and prevent fraud, and that occasionally interviews had been fired. It also turned out that these extreme cases were rare. Openness and trust in national survey agencies should not mean a blind trust. It would mean, however, that reports of falsification by national parties are neither punished nor seen as a lack of quality, but as an indication of quality control at a national level. The CST therefore welcome receiving information on the detection of undesirable interviewer behaviour and falsification at a national level, before data delivery.

In the preparation of fieldwork the risk of fraud should be discussed with the National Coordinators and the survey agencies.

c) Good questionnaire

Undesirable interviewer behaviour, like rephrasing or simplifying questions, skipping introductions or questions, speeding, etc. may occur often when questionnaires are seen as long and boring, and questions are hard to understand, difficult or overly sensitive. In the ESS a lot of efforts have been made to make the questionnaire interesting and not too long, and to make sure interviewers are well-trained and well-briefed to guide all respondents through all questions. Still, this is a serious challenge given a questionnaire that intends to measure many different aspects, change over time and compare across cultures, languages and countries.

d) Interviewers

Interviewers are a key element in the ESS. Good interviewers make the ESS a success, less proficient interviewers threaten the quality of the survey and bad or even fraudulent interviewers are a big risk. It is important that interviewers are aware of key survey quality issues, and that the CST acknowledge the difficult tasks they have and the problems they encounter.

Interpenetrated designs, in which interviewers work in different areas, and more than one interviewer works in each area, can help to identify interviewer effects, and also to detect fraud. However, this measure may be expensive in large countries.

Reducing the interviewer workload can mitigate the effect of undesirable interviewer behaviour, but could also reduce the risk, because the pressure on interviewers to deliver will be smaller. On the other hand, good interviewer training and briefing will be more expensive when interviewers conduct only a small number of interviews. Involvement and commitment is likely to be higher when interviewers have a substantial workload. In any case, the analysis of the effect of interviewer workload and interviewer effects in the different countries should be a permanent part of the ESS quality assessment.

Because interviewers are so important, they should be remunerated fairly for their efforts. This means that they need to be sufficiently paid for conducting the interview, but also for completing the contact forms, traveling to sample persons who are not at home and trying to persuade people who will steadfastly refuse to participate. Just paying interviewers for completed interviews might be conducive to fraud. Studies on the structure of financial payments are not totally clear on the effects of different

payment schedules, partly because of confounding factors such as survey mode, economic climate, national practices, etc., but fair payment is a *sine qua non*.

Also important will be training and briefing, emphasizing the need to pose the questions as formulated, select the right person according to the guidelines, and making interviewers aware of the importance of data quality and the methods that are in place to detect undesirable behaviour and fraud.

5. Falsification, undesirable interviewer behaviour and interviewer effects

As mentioned above, falsification may be hard to prove, and of falsification. Robbins (2018) emphasises that falsification should not be confused with variations in interviewer characteristics and experience. He also stresses that falsification implies intent: *“At times, what appears to be fabrication may be the result of unintentional errors in data collection that could be solved through additional training or other changes in data collection methods. However, seeking to determine if the problem is fabrication or an unintentional error is a critical part of the diagnostic process, as it provides important information about how to correct the issue in a specific survey and what steps must be taken to prevent it in future surveys.”* (Robbins, 2018, p. xx).

Even though the ESS has high quality standards, unintentional errors, undesirable interviewer behaviour and national variation on interviewer characteristics can be a source of error. It can be assumed that these factors have a greater impact on data quality than plain falsification. In the ESS we have come to acknowledge a heuristic distinction between a) deliberate falsification, b) unintentional errors, c) undesirable interviewer behaviour which is avoidable, d) undesirable interviewer behaviour driven by context, and e) natural variations in interviewer characteristics and behaviour that may produce outcomes that could suggest falsification but are an (almost) unavoidable effect of face-to face interviewing. The following examples can be given, mostly relevant for the design and specifications of the European Social Survey:

- a) *Deliberate falsification* includes:
 - curbstoning (falsifying entire interviews),
 - partially duplicating interviews,
 - identifying a sample person incorrectly as ineligible or not being able to participate because of language problems,
 - selecting available household members as respondents instead of a random member of the household because they are more cooperative or more often at home when the interviewer calls (possibly resulting in an overrepresentation of women, as shown by Koch (2016),
 - incorrectly recording answers to filter questions to reduce the duration of the survey, etc.
- b) *Unintentional errors* includes erroneously interviewing the wrong person, recording the wrong day for the interview, keying the wrong answer category, etc.
- c) *Undesirable interviewer behaviour that could be avoided* includes speeding through the interview, being opinionated and disapproving of respondents' answers, skipping introductions to questions, not handing over show cards and not properly reading answer categories, etc.

- d) *Undesirable interviewer behaviour driven by context* includes speeding through the interview when the respondent seems close to breaking off the interview, chatting with respondents who are unsure about their answers or are happy to receive some personal attention, scoring DK without properly reading the question when a respondent has repeatedly told that they really don't know anything about the topic, rephrasing a question when the respondent misunderstands, etc.
- e) *Natural variation in interviewer characteristics and behaviours* includes the effect of differences between interviewers that could result in outcomes that look as if falsification or undesirable behaviour has taken place. Some interviewers, for instance, could be particularly adept at obtaining the cooperation of women (or the elderly or the higher educated). Other interviewers might be very skilled in persuading people to participate who are not at all interested in the topic of the survey (resulting in many DKs). As a result, because of the fact that interviewers often are assigned to particular areas or PSUs, this could result in differential nonresponse errors or measurement errors in different areas.

Although this paper focuses on the detection and prevention of falsification in the ESS, it should be clear there are many types of undesirable interviewer behaviour, that errors and falsification are both undesirable but not identical, and that there are many interviewer effects that could suggest falsification but are a natural consequence of face-to-face interviewing, and that may be difficult to avoid. This means that there may be a thin line between the categories mentioned above (e.g., between c and d), and that similar anomalies in the data (partial duplicates, many DKs, an overrepresentation of women among the respondents) may have different causes.

As mentioned above, good interviewers can be a great asset and contribute to high data quality. But, not all interviewers are equal – the human factor will always play a role – and not all interviewers are equally good. Hence, there are a large number of studies showing interviewer effects in the ESS (listed in Appendix B). This section comprises a short overview of interviewer effects and possibly undesirable interviewer behaviour that cannot be designated as falsification. Some of the issues here may be unavoidable, and some of them may have a detrimental effect on survey quality. This overview is not complete, but may give an idea of the problems that are encountered in cross-national surveys.

a) Respondent recruitment

In the recruitment phase selection effects may play a role. Interviewers may vary in the degree of success in obtaining cooperation from different types of respondents. Some interviewers may be more effective with elderly people, women, the higher educated, minority ethnic groups than others and some interviewers may be highly successful in recruiting reluctant, uninterested respondents resulting in a high rate of satisficing (and more Don't Knows as a result). This effect is hard to avoid, although more training could help interviewers to recruit a wide range of persons.

However, if interviewers purposely concentrate their efforts on easy-to-get respondents than on hard-to-get sample units they may reinforce nonresponse bias. They may also classify refusals as ineligible to avoid having to work the case further. And they may also select other persons than the designated sample person, e.g., more willing persons, or those that are more often at home. The focus on easy-to-get respondents is undesirable, whereas the interviewing of someone else than the designated respondent can be seen as data fabrication.

b) Interviewer and area effect

People living in the same area (or, more technically, in the same PSU) may be more similar than other people. Part of this similarity can be caused by the fact that people living in the same area are often interviewed by the same interviewer. Part of the area effects may thus be interviewer effects. These may be exacerbated by different interview conditions in different areas. Unsafe neighbourhoods may make evening calls less attractive resulting in an underrepresentation of employed respondents for some interviewers.

c) Social desirability

Interviewer characteristics may affect the answers respondents give. Respondents may adapt their answers to characteristics of the interviewer:

- More positive about emancipation to female interviewers,
- Less racist to black interviewers,
- Less alcohol use and more religious behaviour from Muslim respondents reported to Muslim interviewers,
- Lower weight reported to thin interviewers compared to overweight interviewers,
- Lower item nonresponse on income when interviewer has no problem in reporting own income.

These effects may be hard to prevent, except by interviewer training and fostering an open, non-judgmental atmosphere during the interview, or by introducing Audio-CASI.

d) Deviations from standardised interviewing

The ESS places great emphasis on the importance of standardised interviewing. One of the reasons is that ESS is a survey conducted in more than 30 European countries, with different levels of survey sophistication, conducted by different types of survey agencies, and by interviewers with different levels of experience and competence. To minimise the effect of this variation, interviewers are urged to strictly adhere to the questions as formulated by the questionnaire design teams and as translated via strict procedures in all languages.

Still, interviewers may digress from standardised interviewing:

- Rephrasing or skipping questions the interviewer perceives as too difficult or too sensitive,
- Leaving out 'boring introductions',
- Not giving the respondents time to think, possibly resulting in straightlining or other satisficing behaviour,
- Answering filter questions negatively for the respondents in order to skip a long list of follow-up questions,
- Speeding.

Sometimes interviewer behaviour like this may be hard to prevent (what is an interviewer expected to do when the respondent is under serious time pressure and wants to end the interview and there are still a number of questions to go?) but in general these interviewer behaviours will increase the measurement errors of the survey, and can thus be seen as undesirable.

6. Where are we now in the ESS

Interviewer falsification is a permanent concern of the ESS, and the CST are increasingly aware of the problem. The ESS strives for permanent improvement and has, using the measures described above, found indications of falsification in some cases in a small number of countries. Changing the survey mode, and conducting the survey without interviewers, will definitely not result in a higher quality.

Over time more measures to prevent falsification have been implemented, including protocols on briefing, weekly monitoring of fieldwork, and meetings with national teams. However, specifications, documents and guidelines are not enough, and undesirable interviewer behaviour in general, and data falsification in particular, cannot be totally prevented. Prevention, and also detection, is especially difficult in cross-national surveys, where survey culture and sophistication differ across countries, where national partners are needed to collect the data and the chain of command is long. Saying it doesn't happen even in a high quality survey would be naïve.

Besides prevention, greater efforts can be devoted to detecting cases of undesirable interviewer behaviour as early as possible in the process. Detection during fieldwork means other interviewers may be able to take over, detection before the release of the data means that correct weights can be calculated and users don't have to be concerned about data quality. In the ESS over time more measures to detect falsification have been implemented, and in the future even more measures will be implemented, as has been outlined in Section 3. In the 2019-2021 Work Programme of the ESS a separate work package will be introduced to promote desirable interviewer behaviour, to minimise undesirable interviewer behaviour, to prevent falsification as much as possible and to detect it in an early stage.

Appendix A: Recent literature on interviewer falsifications

- Biemer, P., Stringer, C., Thissen, R., Day, O. & Hsieh, P. (2016). Interviewer falsification: Current and best practices for prevention, detection, and mitigation. *Statistical Journal of the IAOS*, 32, 313–326.
- Birnbaum, B. (2012). *Algorithmic approaches to detecting interviewer fabrication in surveys*. [Ph.D. dissertation](#), University of Washington.
- Blasius, J. & Thiessen, V. (2018). Perceived Corruption, Trust, and Interviewer Behavior in 26 European Countries. *Sociological Methods & Research*, 1-38.
- Bredl, S., Winker, P. & Kötschau, K. (2012). A statistical approach to detect interviewer falsification of survey data. *Survey Methodology*, 38 (1), 1-10.
- De Haas, S. & Winker, P. (2014). Identification of partial falsifications in survey data. *Statistical Journal of the IAOS*, 30, 271–281
- Koczela, S., Furlong, C., McCarthy, J. & Mushtaq, A. (2015). Curbstoning and beyond: Confronting data fabrication in survey research. *Statistical Journal of the IAOS*, 31, 413–422.
- Koczela, S. & Scheuren, F. (2016). Editorial. Progress in understanding survey data fabrication. *Statistical Journal of the IAOS*, 32, 277–282.
- Kuriakose, N. & Robbins, M. (2016). Don't get duped: Fraud through duplication in public opinion surveys. *Statistical Journal of the IAOS*, 32, 283–291.
- Landrock, U. & Menold, N. (2016). Validation of theoretical assumptions with real and falsified survey data. *Statistical Journal of the IAOS*, 32, 305–312.
- Menold, K. & Kemper, C.J. (2014). How Do Real and Falsified Data Differ? Psychology of Survey Response as a Source of Falsification Indicators in Face-to-Face Surveys. *International Journal of Public Opinion Research*, 26 (1), 41-65.
- Menold, N. Winker, P., Storfinger, N., & Bredl, S. (2014). Interviewer' falsifications in face-to-face-surveys: impact, reasons, detection and prevention. In: U. Engel, B. Jann & Peter Lynn (Eds.), *Improving survey methods: lessons from recent research* (pp. 86-100). New York: Routledge.
- Robbins, M. (2018). New Frontiers in Detecting Data Fabrication. In: T.P. Johnson, B.-E. Pennell, I.A.L. Stoop, B. Dorer (Eds.), *Advances in Comparative Survey Methods: Multinational, Multiregional, and Multicultural Contexts (3MC)*. Wiley (2018).
- Simmons, K., Mercer, A., Schwarzer, S. & Kennedy, C. (2016). Evaluating a new proposal for detecting data falsification in surveys. The underlying causes of “high matches” between survey respondents. *Statistical Journal of the IAOS*, 32, 327–338.
- Spagat, M. (2016). Comment on “Don't get duped: Fraud through duplication in public opinion surveys”. *Statistical Journal of the IAOS*, 32, 293–294.
- Thiessen, V. & Blasius, J. (2016). Another Look at Survey Data Quality. In: C. Wolf, D. Joye, T.W. Smith & Y-c. Fu (Eds.), *The Sage Handbook of Survey Quality* (pp. 613-629). Los Angeles. Sage.
- Thissen, M.R. & Myers, S.K. (2016). Systems and processes for detecting interviewer falsification and assuring data collection quality. *Statistical Journal of the IAOS*, 32, 339–347.
- Winker, P. (2016). Assuring the quality of survey data: Incentives, detection and documentation of deviant behavior. *Statistical Journal of the IAOS*, 32, 295–303.
- Winker, P., Kruse, K.-W., Menold, N. & Landrock, U. (2015). Interviewer Effects in Real and Falsified Interviews - Results from a Large Scale Experiment. *Statistical Journal of the IAOS*, 31, 423-434.

Appendix B: Studies on ESS data: interviewer effects and undesirable interviewer behaviour

- Beullens, K. & Loosveldt, G. (2013). Assessing Interviewer Variance and its Implications for Data Collection: Report on Interviewer Effects in the Fifth Round of the European Social Survey (ESS-DACE Deliverable 12.10). KU Leuven: Centre for Sociological Research.
- Beullens, K. & Loosveldt, G. (2014). Interviewer effects on latent constructs in survey research. *Journal of Survey Statistics and Methodology*, 2(4), 433–458.
- Beullens, K. & Loosveldt, G. (2016). Interviewer Effects in the European Social Survey. *Survey Research Methods*, 10 (2), pp. 103 – 118.
- Blom, A. G., de Leeuw, E. D. & Hox, J. (2011). Interviewer effects on nonresponse in the European Social Survey. *Journal of Official Statistics*, 27(2), 359–377.
- Blom, A., Lynn, P. & Jäckle, A. (2008). Understanding Cross-National Differences in Unit Non-Response: The Role of Contact Data. *ISER Working Paper 2008-01*. Colchester: University of Essex
- Koch, A. (2016). Assessment of socio-demographic sample composition in ESS Round 6. Mannheim: European Social Survey, GESIS.
- Loosveldt, G. (2008). Face-to-Face Interviews. In: E. de Leeuw, J. Hox & D. Dillman (Eds.), *International Handbook of Survey Methodology* (pp.201- 220). New York: Lawrence Erlbaum Associates.
- Loosveldt, G. & Beullens, K. (2013a). The impact of respondents and interviewers on interview speed in face-to-face interviews. *Social Science Research*, 42(6), 1422–1430.
- Loosveldt, G. & Beullens, K. (2013b). 'How long will it take?' An analysis of interview length in the fifth round of the European Social Survey. *Survey Research Methods*, 7(2), 69–78.
- Loosveldt, G. & Beullens, K. (2014). *Report on Interviewer-Related Variances in the European Social Survey Round 6 (ESS ERIC Deliverable 7.2)*. KU Leuven: Centre for Sociological Research.
- Loosveldt, G. & Beullens, K. (2017). Interviewer Effects on Non-Differentiation and Straightlining in the European Social Survey. *Journal of Official Statistics*; 2017, 33 (2), pp. 409 – 426.
- Matsuo, H. & Loosveldt, G. (2012). *Data Quality in Interviewer Observation Data: Interviewer Burden Perspective from ESS Round 5*. Paper presented at the International Conference on European Social Survey, Nicosia (Cyprus), November 23-25.
- Philippens, M. & Loosveldt, G. (2004). *Modelling Interviewer-Effects in the European Social Survey*. Paper presented at the International Conference on Social Science Methodology, Amsterdam.
- Pickery, J. & Loosveldt, G. (2002). A Multilevel Multinomial Analysis of Interviewer Effects on Various Components of Unit Nonresponse. *Quality & Quantity*, 36, 427-437.
- Stoop, I., Billiet, J., Koch, A. & Fitzgerald, R. (2010). *Improving Survey Response: Lessons Learned from the European Social Survey*. Chichester: Wiley.
- Vandenplas, C. & Loosveldt, G. (2017). Modeling the weekly data collection efficiency of face-to-face surveys: Six rounds of the European Social Survey. *Journal of Survey Statistics and Methodology* 5(2), 212-232.
- Vandenplas, C., Loosveldt, G. & Beullens K. (2017). Fieldwork monitoring for the European Social Survey: an illustration with Belgium and the Czech Republic in round 7. *Journal of Official Statistics* 33(3), 659-686.
- Vandenplas, C., Loosveldt, G., Beullens, K. & Denies, K. (2017). Are interviewer effects on interview speed related to interviewer effects on straightlining tendency in the European Social Survey? An interviewer-related analysis. *Journal of Survey Statistics and methodology*, smx034, <https://doi.org/10.1093/jssam/smx034>.