



TNS Infratest
Sozialforschung

INVEDUC

Investing in Education in Europe: Attitudes, Politics and Policies

Final report

to the

University of Konstanz

78464 Konstanz

Reference Nr. 2013F04E

Invitation of bids for the conduct of a representative survey of mass public opinion in 8 European countries

submitted by

TNS Infratest Sozialforschung

Landsberger Straße 284

80687 München

Miriam Gensicke
Josef Hartmann
Nikolai Tschersich

Munich, 7 July 2014
67.04.125580

Contents

1	Survey design	3
2	Questionnaire Development	5
2.1	Contents of the questionnaire.....	5
2.2	Translation process	5
2.2.1	Translatability assessment	6
2.2.2	Translation and quality control	8
2.3	Finalization, programming and testing of the questionnaire.....	10
3	Design of Sampling	11
4	Fieldwork Preparation.....	13
4.1	Organisation of fieldwork in the TripleC system	13
4.2	Preparation seminar and briefing of the interviewer	15
5	Fieldwork execution and results.....	17
6	Weighting	20
7	Data check and data file.....	23

Annex:

- Annex I: Master questionnaire
- Annex II: Language versions of the questionnaire
- Annex III: Preparation presentation

1 Survey design

The aim of the study "Investing in Education in Europe: Attitudes, Politics and Policies" (INVEDUC) is to understand the extent to which voting preferences influence the political process, especially as regards the relationship between educational and social policy. To this end, in a first step in a multi-country survey conducted in eight countries information about voters' preferences concerning level, distribution and regulation of public spending for education, for different educational areas and for other social areas is collected.

The universe of the survey is the adult population aged 18 or older in the selected countries, resident in private households and available by phone (landline or mobile) – regardless their ethnicity or nationality.

The survey was conducted by computer assisted telephone interviews (CATI) in the eight countries and coordinated centrally by the TripleC centre of TNS in Brussels. The average interview duration was planned to be 25 minutes.

The main topics of the survey were:

- Subjective measures on general political preferences and attitudes in the area of politics (e.g. party affiliation) and policy-making (role and scope of government)
- Subjective measures, particularly personal attributes as e.g. generalised risk aversion, perception of efficacy to steer own life etc.
- Preferences on overall public expenditure levels in education policies
- Differentiation of spending priorities on the main pillars of the educational system, such as pre-school and early childhood education, general school education, vocational education and training and universities and other higher education
- Priorities on government investment, as regards education policy vs. classical social policy transfer programmes
- Financing of investments
- General socio-demographic characteristics
- Measures of educational resources/formal qualification levels

The eight countries were selected by the University of Konstanz. The following table gives an overview of the selected countries and the target of net interviews per country.

Table 1: Overview of countries and target of net interviews

Selected countries	Target of net interviews
Denmark	n = 1.000
France	n = 1.000
Germany	n = 1.500
Ireland	n = 1.000
Italy	n = 1.000
Spain	n = 1.000
Sweden	n = 1.100
United Kingdom	n = 1.300
Total sum	n = 8.900

2 Questionnaire Development

2.1 Contents of the questionnaire

The draft questionnaire was developed by the University of Konstanz, further elaborated in intensive discussions between the team of the University of Konstanz and of TNS Infratest Sozialforschung and tested in a pilot survey. The aim was a systematic and tailored research instrument with specifications of the different dimensions on attitudes towards educational policies. Every question or questionnaire module was designed in a way to maximise the likelihood of delivering valuable measures to answer the research questions. For some concepts several versions were included in the pilot questionnaire in order to find an optimal operationalization.

2.2 Translation process

The elaboration of high quality national questionnaire versions is an absolute must for any survey aiming at cross-national comparability. Otherwise, it will not be clear whether differences in findings are due to differences in reality or just caused by a lack of equivalence in the translations. A not fully equivalent translation of one single key word may render a whole question or even sequence of questions worthless for cross-national comparison.

In the past two decades, increased attention has been given to the translation and adaptation design of data collection instruments and best practices have been described in the literature and implemented by field practitioners (cf. e.g. Harkness, J. et al. 2010¹; Willis, G. et al, 2010² or Dept, S. et al. 2010³). For the current survey, we applied a version that leads to a high quality output while at the same time being cost and time efficient. The applied model, at first, put much emphasis in upstream work, i.e. the production of a clear, well translatable source questionnaire which helps to avoid any ambiguities in the translations. The process was structured in four distinct parts, the first three parts being carried out by our linguistic cooperation partner:

- **Translatability check**, which consisted in defining processes that will ensure the production of national versions of questionnaires that meet stringent equivalence standards;
- **Translation**;
- **Linguistic quality control**, which consisted in checking whether the equivalence standards were met and in undertaking corrective action when they were not;

¹ Harkness, J., Villar, A.; Edwards, B. (2010): Translation, Adaptation, and Design, in: Harkness, J. et al. (eds.): Survey Methods in Multinational, Multiregional, and Multicultural Contexts 117-140.

² Willis, G. et al. (2010): Evaluation of a Multistep Survey Translation Process, in: Harkness, J. et al. (eds.): Survey Methods in Multinational, Multiregional, and Multicultural Contexts 141-156.

³ Dept, S., Ferrari, A., Wäyrynen, L. (2010): Developments in Translation Verification Procedures in Three Multilingual Assessments: A Plea for an Integrated Translation and Adaptation Monitoring Tool, in: Harkness, J. et al. (eds.): Survey Methods in Multinational, Multiregional, and Multicultural Contexts 157-173.

- **Final quality control:** In this final step, the documentation of the decisions taken in the translation process were undergone a final check by the project team at the University of Konstanz and TNS Infratest.

In the following, the above mentioned steps of the translation process are described more in detail. It should be mentioned, that the University of Konstanz translated the English master version to the German version, and that our linguistic partner cApStAn was responsible for the translation of the Danish, French, Italian, Spanish and Swedish versions. For Ireland the English master version was used.

2.2.1 Translatability assessment

In a multi-national survey, a fluent and clear wording in English does not yet guarantee that the different language versions are also clear and without ambiguities. Some English wordings that are very clear in their meaning for native speakers may be ambiguous to translate into other languages because of the large number of different meanings English vocabulary often has.

To this end, the master questionnaire was undergone a thorough Translatability Assessment at our linguistic partner cApStAn before the translation started. The Translatability Assessment consists in submitting the draft version of the questionnaire items to experienced linguists covering the language groups. One expert went through the exercise of producing a draft translation of all items in all languages. These translations were not intended for further use, but they helped contributors identify and describe the issues translators will be confronted with. Comments were then collated by a senior linguist at cApStAn.

A set of 13 translatability categories was used to report on the translation, adaptation and cultural issues identified. The definitions of these categories are presented in Table 2 below. Whenever possible, alternative wording was proposed. This new formulation gave item developers ideas on how to circumvent the problem. Sometimes the linguists also suggested inserting a translation note to clarify what is meant by a given term or expression, or to indicate the type of adaptation that may be necessary.

Table 2: Translatability categories

STRAIGHTFORWARD	No potential translation or adaptation problems identified during the advance translation of this segment into languages from at least two language groups.
KNOWN DIFFICULTY, KNOWN WORKAROUNDS	A translation/adaptation difficulty has been recognised in this segment and has been encountered in the past. Satisfactory solutions to this issue have been successfully implemented.
POTENTIAL ISSUES	The current wording or content of this segment is likely to give rise to translation or adaptation problems in some languages, to the extent that functional equivalence may be difficult to achieve.
POTENTIALLY AMBIGUOUS	The current wording or content of this segment could be interpreted in more than one way and it is desirable to disambiguate the source version of this segment before submitting it for translation/adaptation.
UNNECESSARILY COMPLEX	The current wording or syntax of this segment is somewhat contorted, for example due to use of several clauses, questions embedded in questions or unnecessary use of passive voice. The source version can be simplified without loss of meaning.
REQUIRES REVIEW	The current source version of this segment is not suitable for translation/adaptation and needs to be edited before submitting for translation/adaptation.
POTENTIAL CULTURAL ISSUE	The semantic content of this segment may be difficult to adapt in a particular cultural or language group.
DOUBLE-BARRELLED	A question touches upon more than one issue, yet allows only for one answer. Many double-barrelled questions can be detected by the existence of the grammatical conjunction "and" in them.
AGREEMENT ISSUE	There is either an agreement issue within the segment (e.g. subject-verb agreement, or sequence of tenses, or a pronoun-antecedent agreement) or an agreement issue between two segments (e.g. no grammatical match between a question and response options).
CONSISTENCY	In this segment, a different term, expression or form of address has been used versus other occurrences of similar content; and this inconsistency seems to be unintentional.
REDUNDANCY	This segment contains a tautology or unnecessary repetition. Removing it would not alter the meaning of the segment.
POSSIBLE ADDITION	The current wording or syntax of this segment is elliptic or unclear, and its implicit meaning is likely to get lost in translation. This could be solved by adding a word or a piece of information.
LOGICAL PROBLEM	This segment contains a logical problem or there is a logical problem between this segment and another segment, and this issue seems to be unintentional.

The consolidated translatability report in Excel format was sent to TNS Infratest and the University of Konstanz, who eliminated ambiguities, Anglo-Saxon idiosyncrasies that may be very difficult to render in certain languages, double-barrelled questions, cultural issues or unnecessary complexity. In a nutshell, the translatability check helped to fine-tune the initial English version of the items so that it became a well translatable source version.

Before the translation commenced, the updated source version was entered in a Translation, Adaptation and Verification Monitoring file in Excel format (TAVM file), and item-per-item guidelines were finalised jointly by cApStAn and the item writers. In addition to guidelines retained from the Translatability Assessment, there were guidelines to list adaptations that are mandatory, desirable or ruled out. They were designed to draw the translators' attention to possible terminology problems, translation traps, literal matches and patterns (in question stems and items, or between questions and response categories)

2.2.2 Translation and quality control

The professional translators selected for this assignment were aware that it is of paramount importance to produce national versions of instruments that meet stringent cross-linguistic and cross-national equivalence standards: They had to have experience in translating survey questionnaires and, more importantly, they had to be familiar with cApStAn's approach, which consists in using the TAVM file to document the way they have dealt with translation and adaptation notes. State-of-the-Art technology was used to ensure consistency: Computer-aided tools allow importing the glossary and using the translation memory to ensure consistency in recurring terms and expressions.

The next step was a fully documented verification, i.e. not proofreading, but linguistic quality control (LQC) and equivalence check. The verifier had to document each intervention in an Excel file using cApStAn's Verifier intervention categories. The definitions of these categories are presented in Table 3 below.

Table 3: Translation evaluation categories

OK	No intervention is needed. The verifier has checked and confirms that the text element or segment is equivalent to source, linguistically correct, and – if applicable – that it conforms to an explicit translation/adaptation guideline. This category may also be used to report an appropriate but undocumented adaptation.
ADDED INFORMATION	An information is present in the target version but not in the source version, e.g. an explanation between brackets of a preceding word.
MISSING INFORMATION	An information is present in the source version but omitted in the target version.
MATCHES AND PATTERNS	1) A literal match (repetition of the same word or phrase) or a synonymous match (use of a synonym or paraphrase) in the source version is not reflected in the target version. Most important: literal or synonymous matches between stimulus and item. 2) A pattern in multiple choice items is not reflected in the target version (e.g. all but one option start with the same word, proportional length of responses options.)
INCONSISTENCY	A recurring element across units (e.g. an instruction or prompt) is inconsistently translated.
ADAPTATION ISSUE	An adaptation is an intentional deviation from the source version made for cultural reasons or to conform to local usage. An adaptation issue occurs when an adaptation would be needed but was not made, or when an inappropriate or unnecessary adaptation was made.
REGISTER / WORDING ISSUE	1) <i>Register</i> : difference in level of terminology (scientific term >< familiar term) or level of language (formal >< casual, standard >< idiomatic) in target versus source. 2) <i>Wording</i> : inappropriate or less than optimal choice of vocabulary or wording in target to fluently convey the same information as in the source. This category is used typically for vague or inaccurate or not quite fluent translations.
GRAMMAR / SYNTAX ISSUE	1) <i>Grammar</i> : grammar mistake that could affect comprehension or equivalence, e.g. wrong subject-verb agreement, wrong case (inflected languages), wrong verb form. 2) <i>Syntax</i> : syntax-related deviation from the source, e.g. a long (source) sentence is split into two (target) sentences or two (source) sentences are merged into a single (target) one; or another syntactic problem due e.g. to overly literal translation of the source.
MISTRANSLATION	A wrong translation, which seriously alters the meaning. A <u>mistranslation should always be reported with a back-translation</u> . Note: a vague or inaccurate translation should rather be classified as a Register/Wording issue (or sometimes a Grammar/Syntax issue). This category covers cases where the source has been misunderstood, but also copy/paste errors that unintentionally result in a wrong text element or segment.
GUIDELINE NOT FOLLOWED	An explicit translation/adaptation guideline for a given text element or segment was overlooked or was not addressed in a satisfactory way.
LEFT IN SOURCE LANGUAGE	A text element or segment that should have been translated was left in source language.
MINOR LINGUISTIC DEFECT	Typo or other linguistic defect (spelling, grammar, capitalization, punctuation, etc.) that does not significantly affect comprehension or equivalence. Correcting such errors is usually not controversial and can be made in track changes without documenting them.
ERRATUM/UPDATE MISSED	An erratum or update notice has been overlooked.
LAYOUT / FORMAT ISSUE	A deviation or defect in layout or formatting: disposition of text and graphics, item labels, question numbering, styles (boldface, <u>underlining</u> , <i>italics</i> , UPPERCASE), legibility of captions, tables, number formatting (decimal separators, "five" versus "5"), etc. In computer-based materials, this includes truncated words in the preview, undesired scrolling, etc.

In addition, verifiers make sure that no residual errors, typographical errors, punctuation, grammar and syntax problems are present. This is, however, only a small subset of their task description. More importantly, they have to carefully compare each target segment to the corresponding source segment and check content equivalence and consistency. Possible deviations are described in the TAVM using the Verifier intervention categories, and comments are inserted in English. When a correction is suggested, the updated version has to be entered into the relevant column of the TAVM. The advantages are that (i) documentation is centralised and easily accessible to all parties; and (ii) the final version is in one column of the Excel file and can be imported into the online authoring tool. cApStAn's project manager, a senior linguist, then scrutinised the verifier's output and clarified all potentially controversial issues with the translator and the verifier before final delivery.

As part of the deliverables, cApStAn also provided a translation and adaptation report describing the procedures in detail and listing problems encountered and solutions found.

2.3 Finalization, programming and testing of the questionnaire

The pilot survey and its results are described in detail in the pilot report. Therefore this information is not repeated in this report. Based on the results of the pilot the questionnaire was revised and cut down a few minutes by deleting a few questions. All changes of formulations were translated once again, these translations were verified by a second translator and finally by TNS Infratest Sozialforschung and the University of Konstanz.

The programming of the master questionnaire (for the pilot survey as well as for the main phase) was done centrally in Munich by the scripting team. The testing of the master questionnaire was carried out by the research team of TNS Infratest Sozialforschung. After the script of the final master questionnaire was tested and finished, the national versions of the questionnaire were generated and also tested by the researcher team in Munich. Beyond that, the local field institutes also had a look on the wording and tested their national version of the questionnaire. In case of recommendations of the local institutes these proposals were discussed with cApStAn and the team of the University of Konstanz. If there was an agreement the proposals were adapted in the final version of the questionnaire.

3 Design of Sampling

The universe of the survey is population aged 18 or more. The survey is conducted per CATI. One of the key challenges of CATI surveys has become the increasing share of residents who belong to the „mobile only“ segment of the population, i.e. persons who live in households who do not have landline telephone connections. The percentage of the mobile only population varies between the different countries. E. g. in Germany it is about 10 % of the population, in other countries, as for example Denmark, Ireland, Italy or Spain the respective share is estimated to be 30% or higher. Variation across social strata and regions implies a selective coverage problem for CATI surveys that only rely on landline sampling.

Therefore a landline/mobile mix of the gross sample was foreseen. As sampling frame we used a set of RDD numbers in all countries – except Sweden (see below). The procedure to generate a RDD sampling frame is as follows: Listed telephone numbers from a recent point in time are drawn from a database. The database that is not limited to single number providers can be used to identify area codes and active blocks of telephone numbers as a part of the process of creating a RDD database. In this process the two last digits of the numbers are deleted and replaced by 00 to 99. By this, also numbers not listed have a positive probability of being selected.

In Sweden the sample frame was different compared to the other countries, because there exists an address register containing 90 % of all Swedish aged 16 years or more. In this register not only the landline numbers are available, but also all registered mobile numbers (except prepaid cards). The high percentage of 90 % of registered persons in this address register was the reason to draw a random sample of persons on this base and not to use a RDD-sample. This means that the Swedish sample was a sample of individuals and not of households as it was for landline numbers in the other countries.

To cover all regions of each country proportional to the number of inhabitants the numbers in the gross sample were stratified regionally by using the nuts 2 regions.⁴ This means, that the gross sample was allocated to cells representing the nuts 2 regions according to their share in the population. Within the cells the numbers to be dialled were selected at random. Thus, in each cell, i.e. region, interviewing can be monitored and steered.

If the number selected was a mobile number, the target person for the interview was the owner of the mobile phone. On the contrary, if the number was a landline number, the target person could be every person living in the household and being at least 18 years old. Therefore, in this case in a second step the target person had to be selected by chance. This was done using the kish selection grid, a computer based random selection procedure to identify the respondent among all persons 18 years or older living in the household.

This description shows that the mobile sample is a random sample of persons already in the first step whereas the landline sample is a random household sample in the first step and only in the second step a sample of persons. This design effect of the landline sample has to be borne in mind

⁴ Regional stratification was possible only for the landline part of the sample. Only landline telephone numbers can be linked to regions. Mobile numbers were drawn at random without ex ante regional stratification. For weighting and analysis the information about the region was asked for in the interview.



when weighting the samples: The design effect has to be corrected for by using a design weight (c.f. chapter 0).

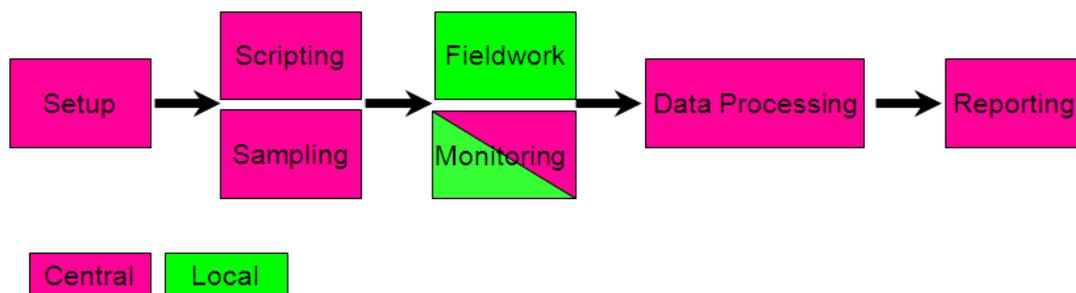
4 Fieldwork Preparation

4.1 Organisation of fieldwork in the TripleC system

For fieldwork organisation the centralised CATI system TripleC, i.e. our Connected Call Centre, was used. TripleC was developed by TNS for multi-country surveys. The TripleC system uses a central infrastructure – technological and personnel – to organise and manage fieldwork and a local infrastructure to conduct fieldwork:

- **The central infrastructure allows for the centrally coordinated processing of the survey** and guarantees best practice in view of the setup of the sample management and CATI system, the sample management, the monitoring and steering of fieldwork, the data and para-data processing as well as the reporting about fieldwork progress.
- **The usage of the local infrastructure to conduct and supervise fieldwork guarantees best practice in view of the concrete realization of the interviews:** The whole team, i.e. interviewers, supervisors, fieldwork managers and project managers, works in the respective country. The interviewer can use this point in recruiting respondents, interviewer management is done locally, supervisors, fieldwork managers and project managers train the interviewers on site, guide, mentor and supervise them as they know country specific characteristics. The national expertise is used to organise and steer fieldwork as well as to react on problems or idiosyncrasies, and, finally, quality of the interviewing is controlled locally in the first place.

Figure 1: Work flow of the TripleC solution



With regard to the organizational and personnel structure the system is organised by a central TripleC project management team which is specialised in the coordination of multi-country CATI surveys. In the project INVEDUC the TripleC team was responsible for the coordination of sampling and fieldwork, strictly following the rules and procedures that were specified by the project team at TNS Infratest. The programming of the CATI-script and of the country versions was done centrally in Munich by the scripting department. Then the team of TripleC implemented the scripts in their central CATI system.

The TripleC team worked directly with the fieldwork operations in all countries. It was the first contact for the national fieldwork partners in all matters related to fieldwork preparations and fieldwork implementation. In matters of questions they could not answer the researcher team was involved.

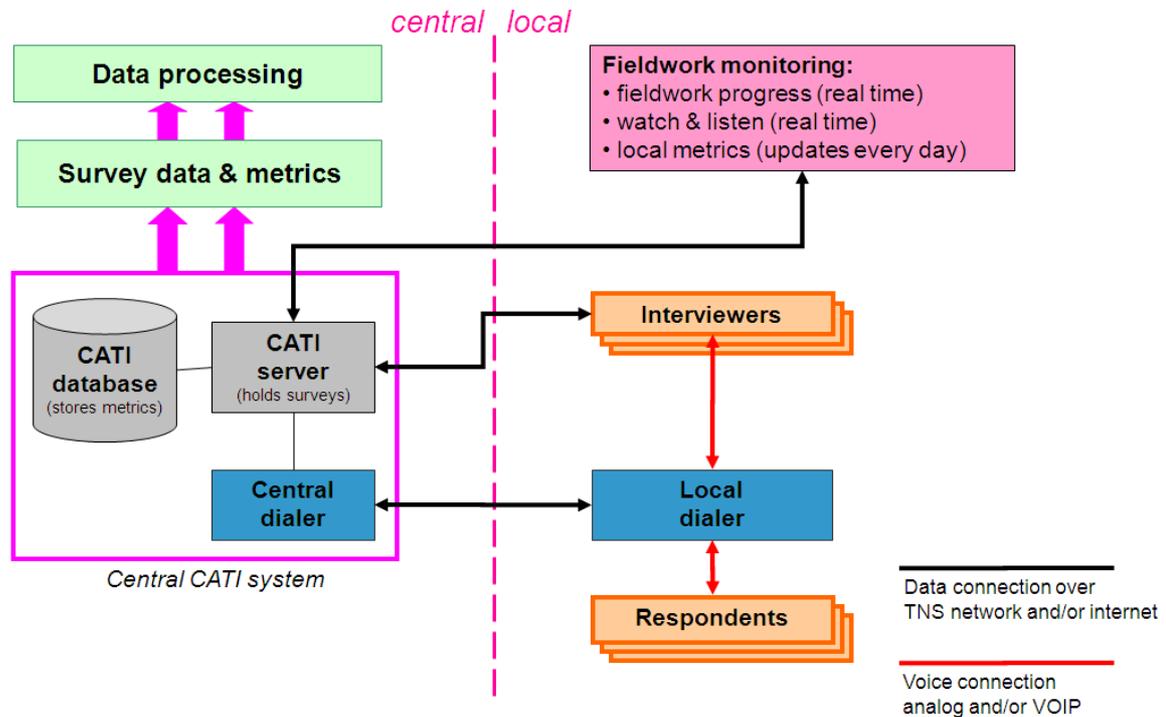
The TripleC team had at its side **two further teams** with definite responsibilities.

- At first, it was supported by the **statistics unit** that implemented the sample design and assisted with the practical sampling process. TripleC sampling took care for the final gross sample selection, for the sample monitoring during the field work period and the realization of interviews in accordance with the regional stratification matrix. Methods and procedures for these tasks were agreed during the setup-phase with the local project team at TNS Infratest Sozialforschung.
- Secondly, the **technical support team** had the task to implement the CATI script, which, as already said, was programmed by the scripting team in Munich, in the CATI system and to keep the system running. Any technical issues were addressed and solved here. It was in close contact with the TripleC team and the coordination team in Munich, with which it discussed any problems as they arise.

The „heart“ of the TripleC solution is the Central CATI system, which consists of the central CATI server, the central CATI database, and the central dialler. The first element of the Central CATI system, the central CATI server, holds the surveys, i.e. the sample management system and the CATI questionnaire. The CATI database is the location where the data stemming from the nationally conducted interviews and the respective contact process data and paradata are stored. On the central dialler the gross sample, containing the telephone numbers from all countries, is saved.

The central dialler allocates a country specific sample to the local dialler, whenever a country works on the survey. The Local dialler, then, contacts the potential respondents and, in case of a successful contact, connects respondent and interviewer. The interviewer has access to the national version(s) of the questionnaire stored on the CATI server. And via the CATI server, all the data gathered in this process, be it contact process data, be it paradata, be it interview data, are transferred to and stored on the CATI database. In addition, the CATI server is the tool to monitor fieldwork in real time: Progress is supervised, sampling is steered, interviewing is watched and listened in. The organization of the process as described also implies that the data are available virtually immediately and can be used for steering the fieldwork as well as for analysis.

Figure 2: TripleC technical solution by TNS for connected call centres



4.2 Preparation seminar and briefing of the interviewer

For preparation of the fieldwork two measures were taken:

- (1) Briefing of field managers and supervisors by the research team of TNS Infratest and the University of Konstanz via webinar
- (2) Personal training and instructions to interviewers by the field managers

Briefing of the field managers:

The personal briefing of the field managers in the eight countries was done by the central project management team of TNS Infratest on 10th April 2014 in a WebEx seminar (internet based presentation combined with a telephone conference). The research team of the University of Konstanz used the opportunity to participate and presented the background and main aims of the survey. The main issues of this WebEx seminar were the following:

- Introduction of the client's team and the project team in Munich
- Overview and design of the project
- Aims and background of the study
- Sampling and selection of the target person
- Definition of the respondent and strategies to get into contact with the target respondent
- Importance of convincing the target persons
- Questionnaire: Changes after the pilot, hints on specific questions
- Organisational hints by TripleC
- Time schedule

The presentation of the preparation seminar (Annex III) was sent to the field work managers in order to be used it for the interviewer training.

Interviewer training

The training of the interviewers and their introduction to the specific requirements of the survey was done in the local telephone studios by the local field work managers and by the supervisors. The project-specific training was based on the information of the presentation they received in the training webinar.

Before starting to telephone the survey all interviewers carried out a test interview to familiarize themselves with the questionnaire, the wording and the routing. The project specific training was provided to all interviewers working on the Inveduc survey.

5 Fieldwork execution and results

Fieldwork of the pilot started in most countries on schedule on 15th April. In Denmark the interviewers were trained on the 16th April and the first interview was conducted on 17th April. Due to the Easter holidays Spain and Ireland started on 22nd April. France was the first country that finished fieldwork on 22nd May. Except Ireland the other countries finished fieldwork between the 25th and the 29th May. Due to a public holiday on 2nd June Ireland conducted the last interview on 4th June.

The following table outlines the duration of the fieldwork in the eight countries.

Table 4: Fieldwork period in the countries

	Fieldwork period
Denmark	17.4. - 25.5.
France	15.4. - 22.5.
Germany	15.4. - 29.5.
Ireland	22.4. - 4.6.
Italy	15.4. - 28.5.
Spain	22.4. - 26.5.
Sweden	15.4. - 27.5.
United Kingdom	15.4. - 28.5.

All in all 8.905 interviews were conducted, 29 % on mobile phones. The target of net interviews differed between the countries. Most of the countries had a target of 1.000 net interviews, in Germany the target was 1.500, in UK it was 1.300 and in Sweden it was 1.100 interviews.

The average duration of the interviews was 25 minutes, the shortest interview took 11 minutes, the longest lasted 77 minutes. Due to the fact that there is no big difference between the interviews related to filtering, this is quite a large range.

The average duration also differed between the countries. The following table shows, that the shortest average duration could be observed in UK, Ireland and Spain (23 Minutes). In four countries the duration was 25 or 26 minutes respectively. The longest interviews in average were conducted in Germany with 28 minutes in average.

The following table gives an overview of the number of net interviews and the average interview duration in the eight countries.

Table 5: Number of net interviews and average duration per country

	Number of interviews	Average duration
Denmark	n = 1.000	25 min
France	n = 1.003	25 min
Germany	n = 1.500	28 min
Ireland	n = 1.000	23 min
Italy	n = 1.002	25 min
Spain	n = 1.000	23 min
Sweden	n = 1.100	26 min
United Kingdom	n = 1.300	23 min
Total	n = 8.905	25 min

Response rates

For the output quality of a survey the process of data collection is one important issue. The way it is organised has direct impact on the response rates that can be reached. Data collection has to be organised in a way that helps to ensure that every single unit has a positive and known chance to become part of the sample. Besides employing a sampling frame that maximally covers the population of interest this task implies:

- random selection of households and target persons, respectively,
- close monitoring and steering of fieldwork, e.g. closing regional cells if the target number of interviews is achieved,
- calling at different hours per day and at different days per week,
- time of fieldwork long enough to also reach persons who, for example, are in holiday or ill,
- possibility to interrupt an interview and to continue it on a later point of time,
- employing the name of the client University of Konstanz in introducing the survey,
- an interesting questionnaire.

All these measures help to increase the chance of maximum heterogeneity of the sample, thus minimizing selectivity. Besides this the response rates are also one indicator for output quality and described in the following table:

Table 6: Overview of gross sample and response rates per country

AAPOR Code	Germany		Denmark		Spain		France		United Kingdom		Ireland		Italy		Sweden	
	Total	Rate	Total	Rate	Total	Rate	Total	Rate	Total	Rate	Total	Rate	Total	Rate	Total	Rate
Gross sample (in principle eligible)	5.679	100%	2.762	100%	4.247	100%	3.640	100%	6.336	100%	4.991	100%	3.589	100%	3.298	100%
3 Unknown eligibility (e.g. line busy, no answer)		16%		17%		18%		19%		11%		20%		17%		20%
2.1 Refusals and break-offs		44%		31%		49%		39%		47%		43%		44%		36%
2.2 Non-contact (e.g. respondent not available, answer device)		14%		16%		10%		14%		21%		17%		11%		10%
1.10 Successful / Response rate	1.500	26%	1.000	36%	1.000	24%	1.003	28%	1.300	21%	1.000	20%	1.000	28%	1.100	33%

6 Weighting

In a CATI survey the selection of the target person needs two steps. In the first step, a number – and thereby a household – is selected, in the second step, within the household the target person is chosen. Whereas this is true for landline CATI surveys in mobile CATI surveys the number directly leads to the target person. According to this two-step procedure landline CATI surveys comprise two design components influencing proportionality of the sample. On the one hand, how many numbers there are to reach the household, on the other hand, the number of persons in the household belonging to the target population. Besides these components there is a second complex of causes for problems of representation: On the one hand, with random sampling a sampling error is inevitable, on the other hand, nonresponse may cause selectivity. The causes for these problems have to be checked and corrected for using a weighting model.

Design weighting

Design weighting applies to landline interviews only. It corrects different inclusion probabilities depending on the number of landline telephone number of the selected household and the number of eligible household members. In a mobile CATI sample no design factor is needed because the target person is identified directly; in these cases the design factor is always set to “1”.

In the first step, the disproportionate selection probabilities due to the design components are corrected. Therefore, at first, the number of landline numbers has to be asked for in the interview. The probability that a household is selected is directly proportional to the number of landline numbers the household can be reached by. A weighting factor inversely proportional to this number adjusts for the different selection probability. This factor enters all weighting steps following.

Within the household only one target person is selected at random. The chance of being selected is inversely proportional to the number of target persons in the household, in this survey persons aged 18 years or more. The number of potential target persons is used to adjust the selection probabilities in a second step. The product of this factor and the first design factor transforms the originally household proportional sample into a person proportional sample. This factor is the input factor for the following weighting step to adjust for selectivity.

In Sweden the sample was a sample on individuals so design weighting was not necessary.

Weighting to correct for selectivity

This weighting step applies for both, landline and mobile interviews. As an integrated theory explaining participation behaviour doesn't exist a correction for selectivity cannot be theoretically deduced. Therefore, weighting factors to correct for selectivity have to be based on comparisons of the sample structure to the population structure (age, gender, educational level, occupational status, region (nuts 2) and employment status).⁵ The aim was to adjust the sample structures as good as possible to the population structures known from Eurostat 2012.⁶

⁵ Depending on the different number of net interviews per country the weighting matrix had to be adjusted in some way for each country.

⁶ For Germany additionally the census 2012 was used.

The structure weights were determined in an iterative process. In this process the weights were iteratively developed to adapt the sample structure with regard to the structure variables or combinations of these variables, respectively. The input factors were the design factors in this process. This means, that for the landline sample the transformation factors were the input factors, for the mobile sample the input factor was always "1".

The general formula for weighting given in the literature is (Deming and Stephan (1940) or Cochran (1968)):⁷

$$g_j = \frac{1}{\frac{n_j}{n} * \frac{N}{N_j}} = \frac{n * N_j}{n_j * N}$$

g = weight

j = strata

n_j = number (design weighted) of cases in the strata j in the sample

n = number of cases in the sample

N_j = number of cases in strata j in the population

N = number of cases in the population

This general formula is represented by the iterative process. We will illustrate this in the following with an example with two strata. In this example marginal distributions of two structural variables are adjusted iteratively. At first, the weights on basis of strata i are calculated. In the next step, the weights based on the distribution of strata j are estimated. The distribution of the first margin is the distribution entering this step.

$$n_{ij} \Rightarrow m_{ij}^I \Rightarrow m_{ij}^{II} \Rightarrow m_{ij}^{III} \Rightarrow m_{ij}^{IV} \Rightarrow \dots \text{ usw.}$$

1. Margin of strata i (cf. Deming and Stephan (1940), equation 52)

$$m_{ij}^I = n_{ij} * \frac{n * N_{i.}}{n_{i.} * N}$$

2. Possibly restriction of weighting factors if they surmount certain limits
3. Margin of strata j (cf. Deming and Stephan (1940), equation 53)

⁷ Cf. Cochran, W.G. 1968: The effectiveness of adjustment by subclassification in removing bias in observational studies. *Biometrics*. 24: 295-313.
Deming, W.E. and Stephan, F.F. 1940: On a least squares adjustment of a sampled frequency table when the expected marginal totals are known. *Annals of Mathematical Statistics*. 11: 427-444.

$$m''_{ij} = m'_{ij} * \frac{n * N_{.j}}{n_{.j} * N}$$

4. Possibly restriction of weighting factors if they surmount certain limits

where:

n_{ij} = number of cases in cell ij (possibly design weighted)

m'_{ij} resp. m''_{ij} = number of cases in the cell ij after the respective marginal adaptation

ij = strata i and j, i and j indicate the category of the respective strata: for example i may represent five age categories and j the two categories of sex

n = number of cases in the sample

N = number of cases in the population

$N_{.i}$ = number of cases in strata i in the population, irrespective of strata j

$n_{.i}$ = number of cases in strata i in the sample, irrespective of strata j

Now the iteration process is starting: The result of the adaptation to strata j enters the next step aiming at the estimation of weighting factors according to strata i and then again according to strata j and so on. This process is continued until the adaptation reaches a predefined convergence criterion or until no further improvement can be arrived at. If there are more strata in principle the process is the same. The iterative process ensures that deviations caused in a previous step are corrected for in subsequent steps. The results of the iterative process are weighting factors that correct for differing selection probabilities according to the design of the survey as well as for those due to selectivity.

"International weighting"

The weighting steps described just were done on a country by country basis.

In a final weighting step national samples were adjusted in a way that reflects the different ratios between the entire population aged 18+ and the net samples. Statistically speaking this is also a design factor which corrects a deliberately disproportional construction of the sample which yields unequal selection probabilities for the units of enquiry. This "international adjustment" allows using the integrated data set for analyses at national and at international level (for any country cluster or the entire sample) by using one single weighting factor ("integrated weighting factor").

7 Data check and data file

The possibilities of the CATI technique to program plausibility checks wherever this is appropriate cannot substitute a thorough ex post data check and cleaning. This is on the one hand due to the fact that not all theoretically possible checks for consistency and plausibility can be implemented into the CATI instrument in order not to overload the interview. On the other hand, with an increasing complexity of the questionnaire the risk of errors remaining in the CATI instrument is increasing, too. Therefore, the data were checked thoroughly ex post.

The data check was realised at TNS Infratest Sozialforschung in Munich. To this end the research team in Munich developed a data-checking syntax in SPSS. This syntax contains all the specific instructions for checking and editing. As the dataset comprised all countries, the checking syntax was implemented in the same way for all countries. A set/range of acceptable values was established for each variable and each recorded value was checked, in isolation from the rest of the data, for its validity. A check for missing values was also included. Furthermore, checks were carried out with regards to filters, i. e. whether or not a certain question has to be answered or not according to the routing of the questionnaire.

As the data of all countries are gathered and stored in the central database of the central CATI system, in principle, the relevant data can be retrieved and checked at any point in time. A data-check was carried out three times: with interim data sets after 300 realised interviews and after 2.200 realised interviews and with the completed data-set after having finalised fieldwork.

The data file was delivered as a Stata file.

Annex

Annex I: Master version of the final questionnaire

Annex II: Language versions of the questionnaire

Annex III: Preparation presentation