

Topic in Social Statistics

Part 1 _B

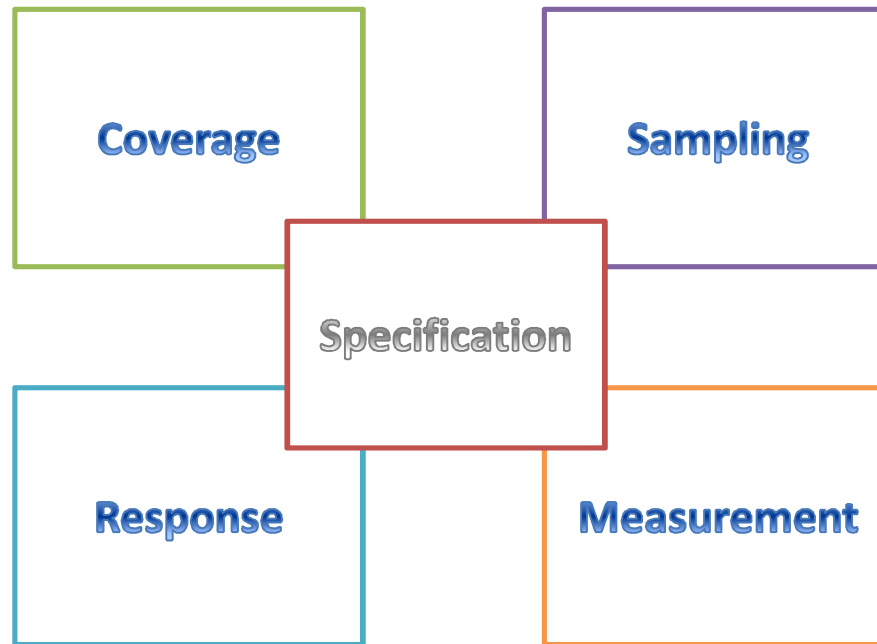
Seppo Laaksonen

Basics on Questionnaire Designing in Surveys

Cornerstones of survey research (modified from Salant&Dillman 1994)

There are big variations in coverage of surveys (under- and over-coverage), also due to classification errors

Avoid/minimize non-response error, and hope that it is as ignorable as possible



Some sampling error occur, hopefully as random and low as possible

Avoid/minimize measurement error, which occurs when a respondent's answer to a question is inaccurate

Questionnaire designing

I thus concentrate on measurement in surveys. This is the big part of Data Collection but I exclude the three other cornerstones almost completely even though there are interactions between them, e.g.

- Non-response may be due to a bad measurement (no motivation to participate, inappropriate survey mode, bad/invalid questions in a questionnaire).
- Coverage cannot be equally achieved with all types of survey data collection modes. E.g. phone numbers are not available for all people, web is not used by all.
- Using interviewers or not i.e. using self-administered answering: there can be an influence both for non-response and measurement, and even for coverage.

Questionnaire designing

What it is in this presentation?

Next I first discuss the whole process of the data collection after the potential respondents have been selected by sampling and a more or less complete list of them is available (This is not the case always like in ad hoc surveys).

The second area is close to the first one but focus more on survey modes.

Thirdly, I will try to give a summary on designing the questions themselves for the survey.

The fourth section is focused on examples of various types, but mainly on the European Social Survey (ESS) that has been very familiar to me over 10 years. In 2001 I started in its sampling expert panel but I have much used its all material in my research and teaching.

Go on to <http://www.europeansocialsurvey.org/>

Questionnaire and questioning

Basically four questions is required to be answered:

(i) How to contact/approach to a potential respondent?

- Mail
- Direct contact at home or another location
- Phone, text message, ...
- Email
- General invitation in media, web, social media, poster
- Automatic invitation in the website to those who just are present there
- Respondent has been contacted at street, shop, sport event, cultural event, training course, outside the voting location (Exit polls)
- Respondent has been contacted from outside (driving, walking etc at certain area/point); she/he does not necessarily know that she/he has been picked up at survey data. This may lead to confidentiality problems if entering sensitive information in the data but counts or other aggregates are not so sensitive (e.g. How many students entered today the building?).

Questionnaire and questioning

(ii) How the information is saved/uploaded into the file?

- Interviewer asks, and saves the answers into the paper or another manual file
- Interviewer asks, and saves the answers into the electronic file.
- Respondent reads, looks and/or listens the questions and saves the answers into an electronic file (self-administered survey answering)
- Respondent reads, looks and/or listens the questions and the interviewer saves the answers into a file
- IT system submits the questions to a respondent and she/he answers by email or by web questionnaire
- IT system collects the data automatically from the data base of the respondent (should be accepted by the respondent or his/her representative); this is typical for business surveys and employer-employee surveys.

Questionnaire and questioning

(iii) **What kind of formats do the questionnaires use?** Note that the format can be converted into a new format after initial data collection.

- Paper that can be filled in manually or printed from an electronic file.
- Electronic local format such as a memory stick
- Text message, email, annex of the email,
- Specific driver on the web, open or closed (closed is the best solution, of course with a unique password for each potential respondent)

(iv) **How to submit the data?**

- If the data are already uploaded into a electronic file, it is ready.
- The paper responses can be submitted by mail or after scanning by email etc.
- Electronic files can also be submitted forward by mail or email or uploading on an appropriate location.

Questionnaire and modes

Key terms with acronyms

Postal Survey, independently or together with another tool; main tool in the case that a potential respondent can read, understand the questions and fill in them

Face to Face Interview (f2f)

PAPI = Paper and Pencil Interview, typically with *f2f*

CAPI = Computer Assisted Personal Interview (phone or *f2f* or Skype)

CASI = Computer Assisted Self Interview (often together with *f2f* for sensitive questions, like ESS supplementary questionnaire)

CATI = Computer Assisted Telephone Interview (this can be conducted in a specific so-called CATI centre or managed by an individual interviewer)

TSI = Telephone Self Interview

Web/Internet Survey (CAWI = computer-assisted web survey that is the only rational choice for web)

CAI = Computer Assisted Interview (general term for almost all except postal survey + PAPI)

Questionnaire and modes

Summary of survey modes

Single-mode, Uni-mode, One-mode: just one mode only used (mail, f2f, phone, web, downloading, ...). Traditional strategy in human surveys.

Multi-mode: two or more modes for different sub-target populations (strata) or domains. If the frames are different for different target populations, it is practical even that these target populations are possibly overlapping. In business surveys common: e.g. electronic questionnaire for large and paper questionnaire for small businesses.

Mixed mode survey: two or more modes for one target population (or for one gross sample, e.g. first web, next phone, or **first f2f, next mail for supplementary questions like ESS**, or first web, next f2f, or **first web, next mail** or mail vs web by the choice of the respondent. I never heard about three modes used as mixed-mode survey in practice.

Designing questions for the questionnaire

This is of course a very big area in the whole survey process. The questions are operationalizations of the measurement desired to study. That is, they should be valid. At the same time, the measurement instruments should be reliable.

How to succeed well in measurement? My practical strategies:

- (i) **Try to find appropriate questions from the earlier studies.** This is often possible unless your field is something very new, never tested or validated. Still, it is good to make a search for good questions and question models. Go on to look at the surveys of your interest areas. There are of course a lot of websites too, like europeansocialsurvey.org and the question bank of the UK <http://survey.net.ac.uk/sqb/>.
- (ii) If you have a new type of target, it is still good to exploit those general models, but next **try to develop your own strategy**, not alone but in the team, and then to test it in your neighbourhood first and after that to pilot = pre-test the whole questionnaire with similar people (taking a small sample) as the real survey will be concerned.

Developing questions for the survey

Answer the following seven questions at least:

1. How well the question is concerned your research target and well formulated in this direction (validity)?
2. Is the question definitely beneficial; the question should give some value added for your analysis?
3. Are you sure that the respondents understand the question as your target for the question is?
4. Has the respondent reasonably information to answer correctly (e.g. terminology is known well, help tools might help)?
5. Are the respondents willing to answer the question?
6. Should the question be presented to all or for a certain group of the sample?
7. Can you find other information in order to analyse the answers given and how reliable they are?

Other general features for questionnaire designing

Size of the questionnaire is a big issue. Of course, it should not be too long but not either too short in this case you will lose useful information. How to get an optimal size?

1. First, it does not matter much if you will have all possible questions in the list but after that you have to go to choose the best combination avoiding overlapping, among others.
2. Finally, your team will make the first selection but this should be piloted with 'real respondents.'
3. Always check that the question is valid and the estimates intended to compute using it are applicable.
4. It is possible that some subjects cannot be measured with one question but needs a pattern of questions using the same formulation. Do not avoid this since a good pattern may help in getting a good reliability for your measurement? Later, you maybe create some factors from this pattern and interpret these correctly.
5. Estimate the time needed filling in the questionnaire and take also into account the response burden of the average interviewer.

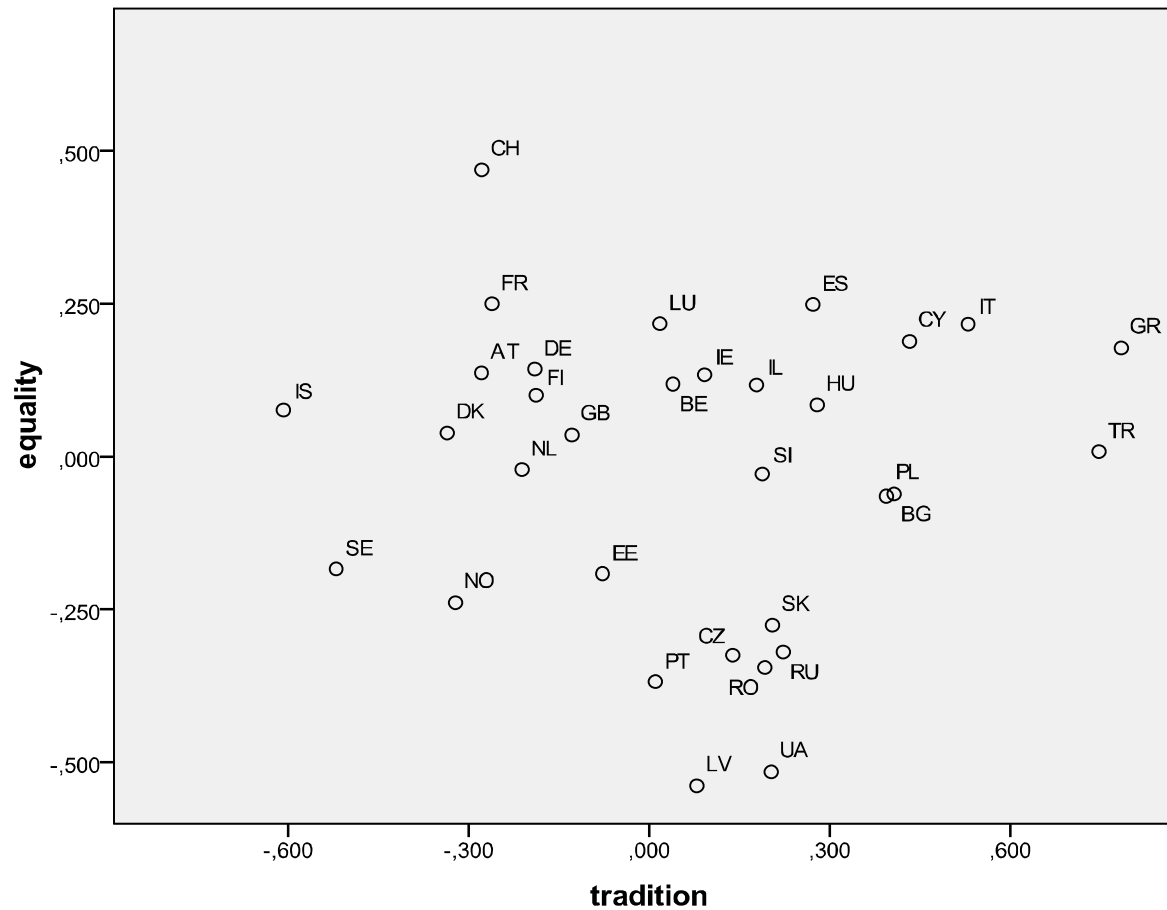
Good to remember that there are good validated patterns for several fields like Shalom Schwarz human values in the ESS. His pattern consists of 21 questions such as

- Thinking up new ideas and being creative is important to him. He likes to do things in his own original way.
- It is important to him to be rich. He wants to have a lot of money and expensive things.
- He thinks it is important that every person in the world should be treated equally. He believes everyone should have equal opportunities in life.
- It's important to him to show his abilities. He wants people to admire what he does.
- It is important to him to live in secure surroundings. He avoids anything that might endanger his safety.

When analysing this pattern with exploratory factor analysis, I have obtained four factors with following short variable labels: Tradition, Equality, Enjoy and Success. The next page gives a cross-classification of the two first factors, averages by country.

It is often very rational to use validated patterns, not trying to create a new pattern that needs much work and good validation. Of course, in the best case, you will be famous after several years, known as

Two first factors of human values of Schwarz based on the ESS data of the four first rounds (2002-2009).
Averages of factor scores



Other general features for questionnaire designing

Order of questions is also essential but there is no only one optimal order. Take into account at least these points.

1. The first real question is important. It is good that it already is concerned one big issue of the survey but should not be too difficult to answer. Naturally, this question should be for all respondents, not for some.

2. Where to put so called background questions (gender, age, education, occupation, workplace). Some prefer to put all these in the end but some others to divide in a quite early stage as far as non-confidential questions are concerned but confidential questions (salary, income, ...) are good to put in a very end.

3. Each survey should have certain specific key areas/subjects and main concentration should be paid to these. Hence it is good to carefully design the questionnaire so that such key questions are in an optimal place.

Other general features for questionnaire designing

Fortunately, possible questions are only of two types. They are concerned either

1. *Facts (age, gender, living area, industry class, occupation, salary, income, partnership, ...)*

Or

2. *Subjective features (attitudes, opinions, assessments, purposes, ...).*

Naturally, questionnaire layout is often essentially different between these two questions. Basically, the first ones are easier to formulate and their scaling is often initially clear. Of course, there can be used different categories for categorical variables like for partnership or occupation and education. Continuous variables can be asked as continuous or they can be categorised, even age or salary. Naturally, using them as continuous it gives more options to categorise these later but how well it is possible to get correct answers, this is another point.

Other general features for questionnaire designing

As far as attitudes, opinions, assessments, purposes, ... are concerned there are in literature different standards for example concerning scales. I think that in history the scales have been shorter than currently for example in the ESS. I give some examples from show cards used in it by *f2f*. Show cards help a respondent to understand better the scaling of the question. These cannot be used in CATI at all, and the interviewer should be able to explain scales in the best way. I am not convinced how well this works. Naturally in self-administered surveys the respondent have in fact such show cards at hand, she/he see the scaling well. Layout is still a big issue, e.g. How far from real answers are the 'don't know' answers?

I go now to

http://www.europeansocialsurvey.org/index.php?option=com_content&view=article&id=63&Itemid=405

Other general features for questionnaire designing

The response burden is thus a big problem even in short questionnaires to some extent. The expected time is always told in advance for a respondent, and in the web survey there should be a measurement indicator that tells how far you are in replying.

An interesting side effect relating to response burden is so-called satisficing:

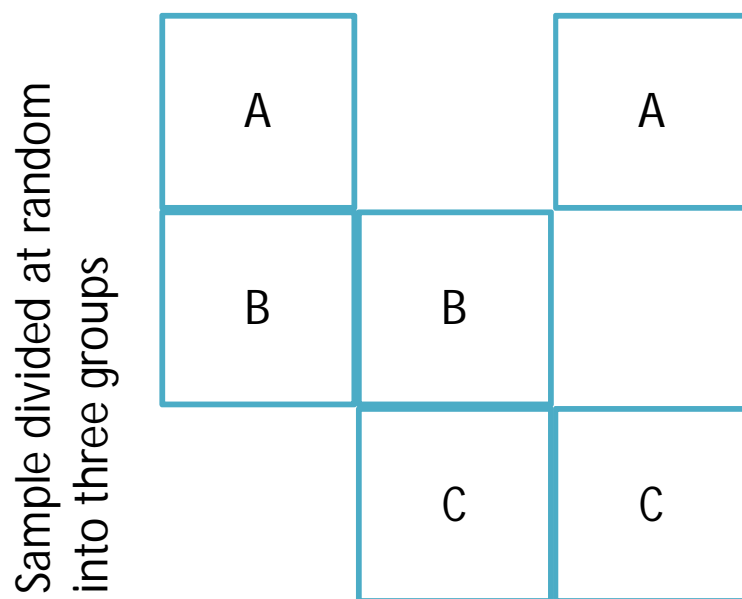
-Since I have started to reply but it seems to take a longer time as expected or the questions are too boring, I will continue to answer even I am too tired to think really my answers.

-OK, I will answer although too many questions are not smart; my solution is that I will put there 'I don't know' or 'I don't like to say my opinion' or I put an average score.

This satisficing is not much examined and it is not easy either. Try to work for it.

Rotating questionnaire

A solution to response burden and satisficing could be a rotating questionnaire in the case the questionnaire seems to be too long and there is danger for increasing measurement errors. The below scheme illustrates this strategy. Here the whole questionnaire has been divided into three sub-questionnaires, A, B and C. Basic background questions are included in all subs.



This strategy gives opportunity to obtain good results also concerning association between all sub-questionnaires.

Very little exploited. **WHY NOT?**

The ESS questionnaire piloting uses a partially similar strategy. In the end I give examples of this piloting.

NO ANSWER

There are different options for no answer like:

- Don't know
- Isn't willing to answer (refusal)
- The question does not concern him/her/it
- Unable to answer correctly
- No time to answer correctly
- Lost answer
- Other reason

Naturally, it is a purpose to avoid 'no answer' answers and formulate the questionnaire so that the number of such answers is as minimal as possible except in the third case that can be often a key indicator in some cases (like if this is a second question relating to details of crimes occurred for example when the first question gives already information that whether crime was occurred).

Thus if the relative number of 'no answer' answers is high, the questionnaire or interviewing is not working well, and any estimates should not be published.

NO ANSWER

Filtering "Don't Know"

There are three ways of dealing with "don't know".

Standard format. The "don't know" option is presented to the respondent, but is recorded if the respondent volunteers it.

Quasi filter. The "don't know" option is included among the possible responses.

Full filter. First the respondent is asked if they have an opinion. Then, if yes, they ask the question.

Basically similar filtering can be applied for the 'Isn't willing to answer' alternative. Some questionnaires do not give opportunity to give any of these alternatives at all. I have stopped to fill in such questionnaire but tried to write somewhere my opinion: "I cannot tell my opinion, since your questionnaire is catastrophic. Do not publish anything based on such violated data!"

Example of human values of Schwartz pattern

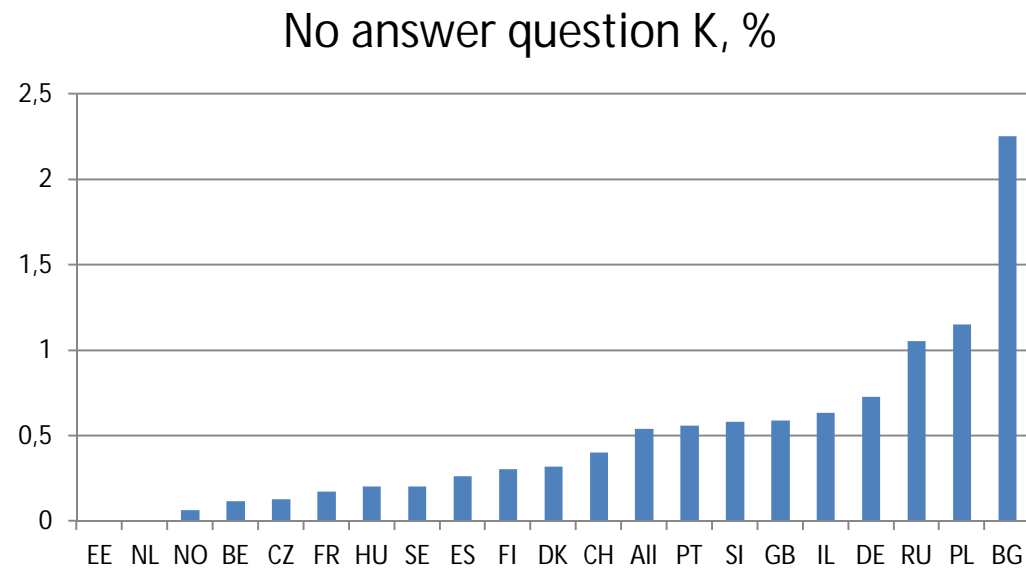
It is good to note that the scales without a mid-point means that the respondent cannot give the 'no answer' answer except keeping all boxes empty. Below is the question K of all 21 human value questions.

How much like you is this person?

	Very much like me	Like me	Some- what like me	A little like me	Not like me	Not like me at all
K						
It is important to her to make her own decisions about what she does. She likes to be free and not depend on others.	01 <input type="checkbox"/>	02 <input type="checkbox"/>	03 <input type="checkbox"/>	04 <input type="checkbox"/>	05 <input type="checkbox"/>	06 <input type="checkbox"/>

Next page you find the percentages of 'no answer' answers by country.

No answers by country in the ESS Round 5 for the question of the previous page, given that she/he answered to the first question of the pattern. This is from the supplementary questionnaire filled in by hand for the paper after the *f2f*. *This could be an indicator of satisficing.*



Example of the different scales ESS test questionnaire, Round 4.

Question in all the three questionnaires: "*Generally speaking, would you say that most people can be trusted, or that you can't be too careful in dealing with people?*"

But there are three scales as follows:

-Core questionnaire with 11 alternatives (without no answers):

0 = You can't be too careful, 10= Most people can be trusted.

-First pilot questionnaire with 6 alternatives:

0 = You can't be too careful, 5= Most people can be trusted.

-Second pilot questionnaire with two alternatives:

1 = You can't be too careful, 2 = Most people can be trusted.

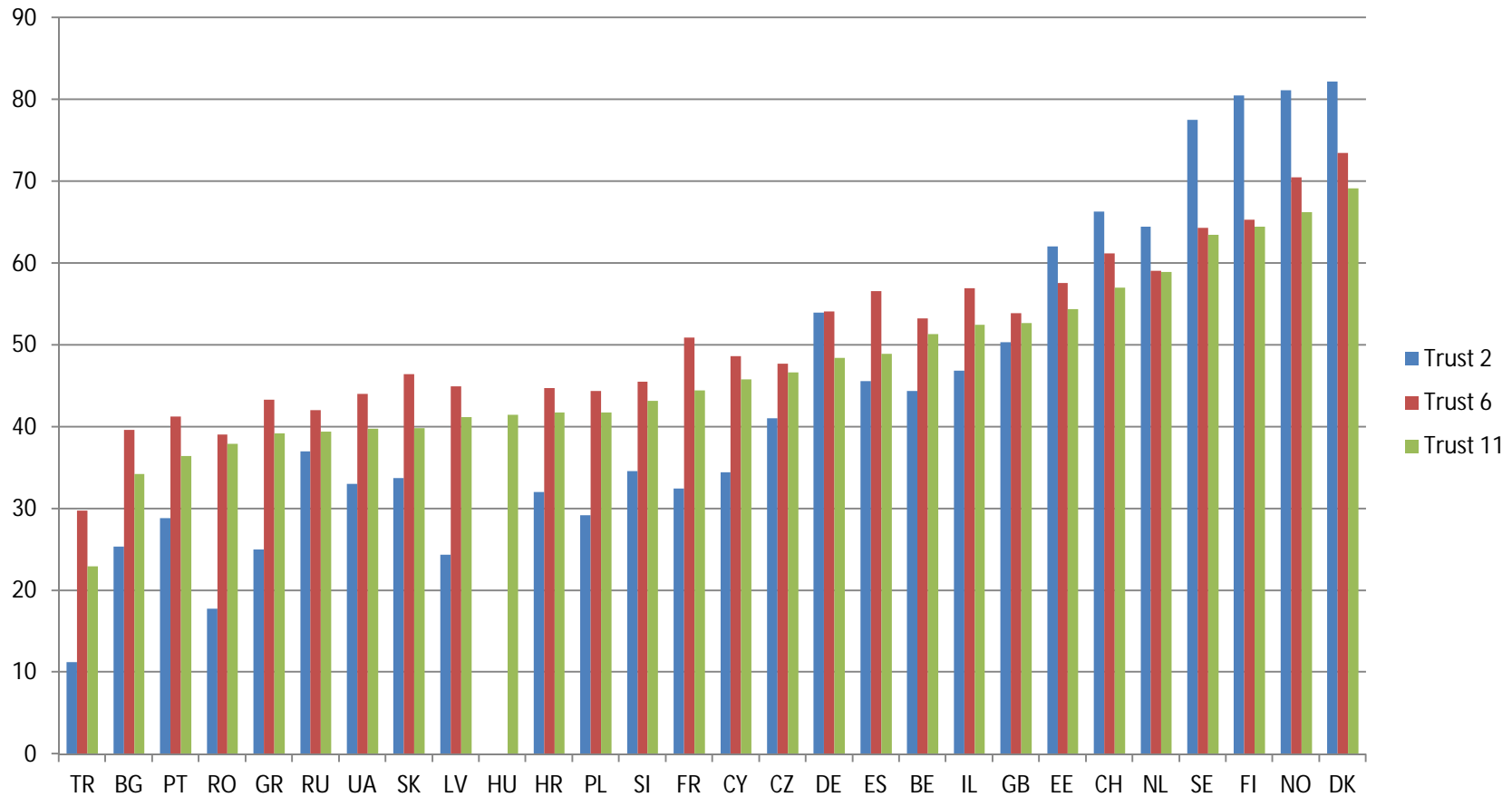
Although the samples are not exactly as accurate, we can quite well compare these three alternatives. On the final page all three questions have been linearly transformed to the same scale so that:

0 = You can't be too careful, 100= Most people can be trusted.

Interpret the results.

Trust other people in the country

Three different scaling for one question. The countries sorted by the scaling with the maximum categories, that is, 11 categories = Trust 11. Respectively 6 categories (Trust 6) and two categories (Trust 2).



Hungary did not participate in the pilot