



# Second European Quality of Life Survey

## Recommendations for the methodology

**Overview**

**Introduction**

**Sampling**

**Fieldwork**

**Questionnaire**

**Data**

**References**

This report is available in electronic format only and has not been submitted to the standard Foundation editorial procedures.

# Overview

The European Quality of Life Survey 2003 was an important contribution to social research on the European Union. It provided a unique snapshot of living conditions and quality of life in the EU member states, accession states and candidate countries on the eve of the eastern enlargement and for the first time enabled the analysis of key social scientific and policy issues in the EU25 and accession states to be conducted on a consistent empirical basis.

The EQLS 2003 was well conducted within its resource limitations. But these limitations were real, especially in that its budget was modest and its timetable short. The European Social Survey (ESS) Round 3, for example, which covered only 24 countries (compared to 28 countries in the EQLS), had a fieldwork budget that was more than four times greater than that of the EQLS (the ESS is now widely viewed as the methodological gold standard of research-oriented sample surveys in Europe). The budget, time and other resources devoted by the ESS to preparation, central coordination and oversight also far exceeded that which was available for the EQLS-1.

The detailed methodological assessment of EQLS 2003 carried out in the present project and in earlier analyses (Kohler, 2006, 2007) revealed that its key methodological weakness lay in its higher risk of sampling bias compared to other similar surveys. In a comparison with five other social surveys, including the ESS and the Eurobarometer 62.1, the EQLS 2003 showed the highest rate of failure against so called 'internal criteria of representativeness' and the highest over-representation of women. The major cause of sampling bias in the EQLS 2003 was a combination of sampling methods and fieldwork arrangements used by the national fieldwork agencies. Any extra resources available for EQLS 2007 should be used to bring real methodological improvements, even though these are not likely to raise EQLS 2007 to the methodological standards of the ESS.

The purpose of this report is to recommend the key improvements in methodology that could be incorporated into EQLS 2007, including an assessment of the methodological advisability of increasing the sample size. The particular concern is to identify feasible and cost-effective means to reduce the risk of sampling bias while remaining within the fieldwork budget likely to be available for EQLS 2007. The budgetary side of this objective is difficult to fulfil with any precision, since the costs of the available methodological options depend on aggregating cost details for 29 countries – costs that are not available to us and that in any event might vary between fieldwork agencies even in a single country. Therefore we have sought to identify a range of methodological improvements that could be implemented at least to some degree within budgetary constraints, though the exact extent of their implementation would depend on cost.

The key conclusions are as follows:

1. Do not generally increase the sample size. The addition of further sample cases is not a cost-effective means of improving EQLS sample quality since it does not address the central challenge of reducing the risk of sample bias.
2. Reduce reliance on 'random walk' sampling, where possible in favour of strict probability sampling. The sample arrangements that will be feasible within cost constraints will vary from country to country. In some cases, no alternative to random walk sampling will be possible. However, fieldwork agencies should be requested to propose what they consider the best possible sampling approach to adopt in their country, taking account of the fieldwork budget.
3. Maximise the efforts to reduce selective unit non response. This includes the use of advance letters and specific means to convert refusals.
4. Undertake substantive efforts to decrease the measurement bias. Personal briefing and thorough back-checking of interviewers are among the means to reach this goal.
5. Assume a central and direct role of the Foundation in the commissioning of fieldwork and its monitoring.

These key conclusions imply a flexible approach to sampling and fieldwork, which is difficult to implement in the tender structure of the Foundation. We suggest some minimum requirements to be pointed out in the tender specification, which however should be accompanied by the development of a standardised questionnaire to collect detailed information on the planned sampling design and fieldwork regulations of the local fieldwork agencies. Based on the answers to these standardised questionnaires the applicant with the best overall package might be chosen. This implies that the Foundation should not necessarily select the cheapest applicant, but the applicant with the best overall package within the given budget. We recommend that the Foundation seeks for help by survey experts for the selection of the best applicant.

Overall, the proposed strategy cannot be implemented fast enough to start fieldwork in March/April. A start of the fieldwork before mid April is however not sensible because of spring holidays. The next alternative fieldwork period will be September 2007.

Besides the key conclusions summarised above, recommendations on various aspects of the EQLS methodology are made. These include proposals for

- the revision of some socio-demographic questions,
- the calculation of the response rate,
- the translation of new questions,
- the pretests of old and new questions,
- the number of call backs,
- the fieldwork time frame,
- the documentation of the fieldwork process, and
- the design of the data set.

# Introduction

This report collects recommendations designed to improve the data quality of the second round of the European Quality of Life Survey (EQLS-2) in comparison to its predecessor, the European Quality of Life Survey of year 2003 (EQLS-1). Most of these recommendations are given without further justification. However, the recommendations are based on a comparison of the EQLS-1 with the European Social Survey (ESS). The report on this comparison (Kohler et al., 2006) contain a substantial amount of back ground information about why many of the recommendations have been chosen.

We have discussed two general scenarios. The first scenario was to increase the sample size. Increasing the sample size reduces the sampling variance of the survey, which is one source of overall survey error (Groves, 2004, 10). Unfortunately, increasing the sample size does not (necessarily) reduce the sampling bias, which was the main problem of EQLS-1. Moreover, even assuming that the budget allows to invest in additional observations, the decrease in sampling variance will be only small unless the budget increase is rather large.

The second scenario was to invest in better methodology. The literature on survey methodology provides guidelines for increasing the accuracy of a survey, which include various methods to decrease the sampling bias. The overall idea of these methods is to organise the survey in a way that increases the means to monitor the work of the interviewer and the work of the fieldwork agencies, and to use these means. This second scenario is more promising. It tackles the observed problems of the EQLS-1 directly, and it helps to decrease not only the sampling variance, but also some of the other sources of survey errors.

The downside of the second scenario is that it is more difficult to budget. Which of the above regularities can be implemented within the budget of the EQLS-2 is difficult to say. The costs and prospects of success very much depend on the experience of the different fieldwork agencies.

We strongly recommend the second scenario. The Foundation should be aware that it will not be possible to improve the data quality without choosing one of these scenarios. In this sense there is a link between them. If the Foundation does not choose scenario 1, it must choose scenario 2. If it chooses scenario 2, there will be no financial resources left over for scenario 1.

## Sample size

Setting a minimum required sample size is necessary. In the EQLS-1 the minimum sample size was 1,000 observations, and around 600 observations for smaller countries. For the EQLS-2 the Foundation is considering the advantages of increasing the minimum sample size to 1,500 observations, and 800 observations for the smaller countries. As higher sample sizes increase the precision of estimators, and allow more in-depth analysis of subgroups, we very much appreciate this plan. Nevertheless, after evaluating all pros and cons, we recommend to abstain from this general increase of this minimum sample size. Instead, we recommend to only increase the sample size of the small countries, and to invest the resources into:

- strict probability sampling and
- better fieldwork regulations.

As this is the most far reaching decision of this report, it needs to be justified in some detail. We started from the assumption that the EQLS-2 is designed for three basic needs:

1. Estimation of country averages of reasonable precision,
2. investigation into trends over time, and
3. estimation of group specific averages within country.

Let us now discuss the possibility to achieve these goals in more detail.

## Sampling variance, sample size and sampling design

The achievement of the mentioned goals depends on (a) the survey bias, and (b) the sampling variance (Groves, 2004, 1–36). The increase of the sample size will help to achieve the three goals because it decreases the sampling variance. However, the increase of the sample size will not change the survey bias. Our recommended strategy might possibly decrease the sampling variance and the survey bias.

To understand our recommendations, the Foundation must be aware that the sampling variance does not solely depend on the sample size. The sampling variance also depends on the sampling design. A multistage sample has a higher sampling variance than a simple random sample. It is often the case that a multistage sample with 1500 observations has a lower precision than a simple random sample with 1000 observations. Survey statisticians therefore have developed the concept of ‘effective sample sizes’. The effective sample size is the number of observations that would be drawn with a simple random sample to get the same precision as a given sample. For example, if a specific multistage sample has 1000 observations, it might be enough to draw a simple random sample of 500 observations to get estimators of the same precision. In this case the effective sample size of the multistage sample would be 500. Another way to express this is the so-called design effect. Conceptually the design effect is the division of the net sample size ( $n_{net}$ ) by the effective sample size ( $n_{eff}$ ),

$$deff = \frac{n_{net}}{n_{eff}} \quad (1)$$

The design effect expresses how much larger the sample size of a multistage sample must be to get the same precision as a corresponding simple random sample. In the example above, we would get a design effect of 2. Design effects of this size are not uncommon in practice.

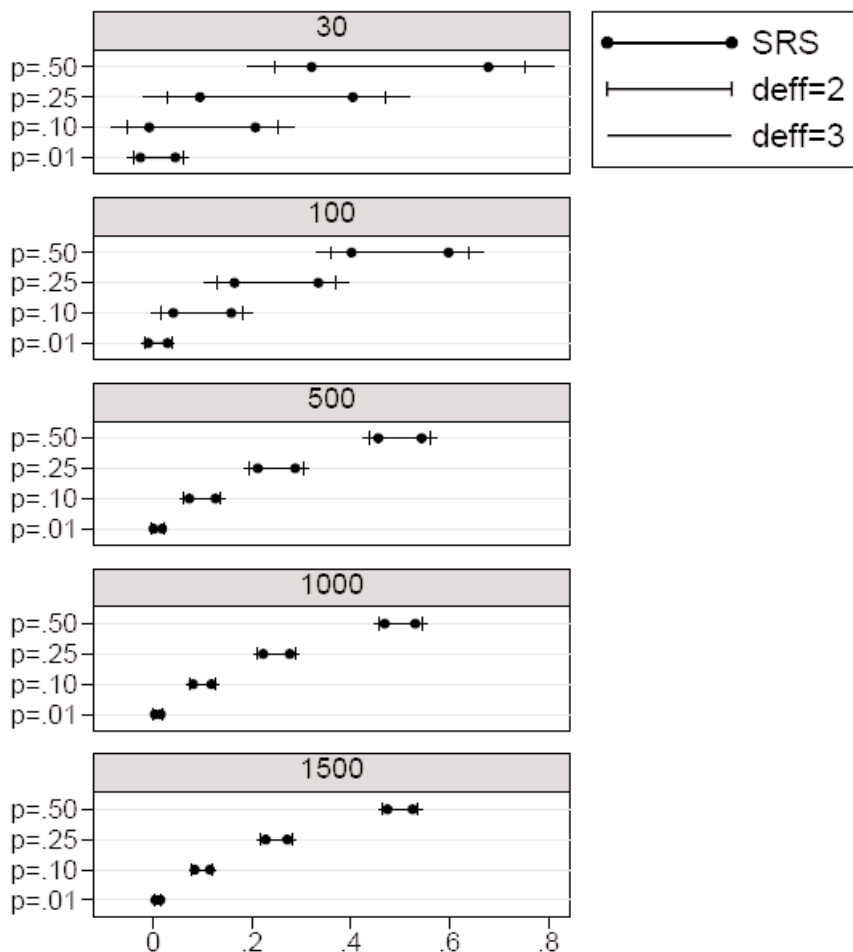
The Foundation can achieve a decrease of the sampling variance either by increasing the sample size, or by minimizing the design effect. That is to say, the Foundation can increase the effective sample size by applying a sampling design with smaller design effects.

Minimizing the design effect is however a costly undertaking. The most natural means to increase the effective sample size would be to apply simple random sampling. Unfortunately, simple random sampling tends to be expensive. Moreover, the Foundation should be aware that it will be impossible to reach an effective sample size of 1000 within the likely budget. With a fixed budget, a nominal sample size of 1500 observations will be accompanied by higher design effects, and hence lower effective sample sizes. We therefore recommend that a team of central methodologists participate in the sample designs to be able to reach at least an effective sample size of 750 with a net sample of 1000.

### Estimation of country averages

Let us discuss the precision of country averages with respect to an arbitrary proportion like the proportion of persons that are satisfied with their lives. The precision for such a proportion can be quantified by the width of the 95% confidence interval. Figure 1 illustrates how the confidence intervals around such proportions vary with the variance of the observed characteristic in the target population, the sample size and the so called ‘design effect’. The design effect captures the effect of the sampling design on the sampling variance. Multi-stage samples often have design effects of two and more.

Figure 1: Confidence intervals by observed proportion, sample size and design-effect



Generally, the confidence intervals are large for small sample sizes, for proportions that are closer to 50 per cent, and for samples with larger design-effects (e.g. multi-stage samples). The typical confidence intervals for the EQLS-1 can be found in the fourth graph of figure 1. Assuming simple random sampling, the confidence interval around a proportion of 50 per cent will reach from around 46.9% to 53.1%. If the sample size is increased to 1500 observations, the boundaries of confidence interval will change to 47.5% and 52.5% respectively. Hence, an increase by 500 observations will enhance the precision of such estimators by just one percentage point.

At the same time an increase from 500 to 1000 observations will increase the precision of an estimator by two percentage points. The general message is: the higher the sample size, the more additional observations must be collected to gain the same increase in precision. Increasing the sample size of the small samples by a fixed number of observations will increase the overall precision of EQLS-2 more, than increasing the bigger samples. This illustrates our recommendation to only increase the sample sizes of the smaller countries.<sup>1</sup>

The plot also illustrates that the confidence interval of a multistage-sample with 1500 observations might be larger than the confidence interval of a simple random sample with 1000 observations. This illustrates our above statement that it is possible to achieve the same gain in precision by using simple random sampling than by increasing the sample size.

### Estimation of trends over time

For the investigation of trends over time the question is how large an observed difference between the surveys of 2003 and 2007 must be, to be sure that it is not due to random fluctuation (i.e. 'significant'). Note that finding significant differences in two independent samples is not equivalent to finding non-overlapping confidence intervals. Rather, the answer must be based on the formulae for a mean comparison test. We will do this here under the assumption of simple random sampling.<sup>2</sup> The derivation of these threshold values require some math, which is documented in the following paragraphs. The reader can skip directly to the discussion of figure 2 without loss of information. Let

$|\bar{x}_1 - \bar{x}_2|$  be the absolute value of the difference between the mean of some variable on time points 2003 and 2007.

This difference is significant if 
$$|\bar{x}_1 - \bar{x}_2| > t_\alpha \times \sqrt{\frac{s_{x_1}^2}{n_1} + \frac{s_{x_2}^2}{n_2}} \quad (2)$$

where  $t_\alpha$  is the test value for the significance level of  $\alpha$  (i.e.  $\sim 1.96$  for the 5% level),  $s^2_{x_1, x_2}$  are the variances for the realisations of  $x$  in 2003 and 2007 respectively, and  $n_{1,2}$  are the respective sample sizes. The question then is, how large  $|\bar{x}_1 - \bar{x}_2|$  has to be, to full-fill condition (2). The answer clearly depend on the respective variances, sample sizes, and the significance level. However, for here the main question is, weather an increase of the sample size of EQLS-2 from 1000 to 1500 observations makes a big change, or not. To answer this more specific question, we can simplify the discussion a little bit. Let us assume that  $x_{1,2}$  are proportions, and denote them with  $p_{1,2}$ . The variance of proportions can be written as  $p(1 - p)$ , then. Let us further rewrite  $p_2$  it terms of  $p_1$  and  $d$ , the difference between  $p_1$  and  $p_2$ :

$$|p_1 - p_2| = d$$

$$p_2 = p_1 - d$$

---

<sup>1</sup> The smaller sample sizes for smaller countries might be also justified by their homogeneity or to not cause survey fatigue. Some analysis we have done indicate that countries' homogeneity varies with variables, while there is little evidence for a general higher homogeneity of small countries.

<sup>2</sup> For complex samples the 'threshold' values will change according to the design-effect.

We can now rewrite (2) to

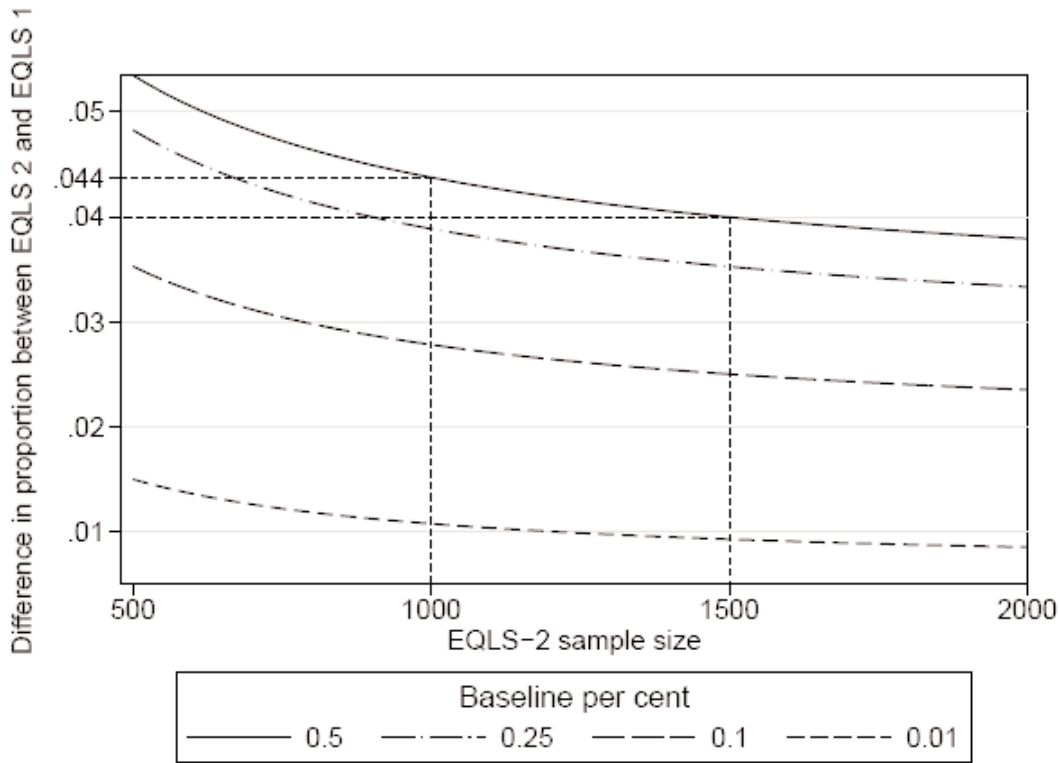
$$|d| > t_\alpha \times \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{(p_1-d)(1-(p_1-d))}{n_2}} \quad (3)$$

Setting (3) to zero and solving for d reveals the following function for the values of d, above which the difference between the two survey years becomes significant:

$$\frac{(2p_1-1)t_\alpha^2}{n_2} - \sqrt{\frac{t_\alpha^4(2p_1-1)^2}{n_1} + 4\left(\frac{p_1 \times (1-p_1)}{n_2} + \frac{p_1(1-p_1)}{n_1}\right)\left(\frac{t_\alpha^2}{n_2} + 1\right)t_\alpha^2} \quad (4)$$

$$\frac{2t_\alpha^2}{n_2} + 2$$

Figure 2: Threshold values for differences in proportions between EQLS-1 and EQLS-2



SRS, level = .05, EQLS-1 sample size = 1000

Figure 2 shows significance threshold values by possible EQLS-2 sample sizes, derived from equation (4). Thereby the 5% significance level has been used, and the sample size of EQLS-1 was set to 1000. The threshold values were calculated for observed EQLS-1 proportions of 1, 10, 25 and 50 per cent. The figure can be read as follows: assuming an EQLS-2 sample of 1000, and an arbitrary EQLS-1 proportion of 50%, the corresponding proportion in the EQLS-2 must differ by 4.4 percentage points to be significant. For an EQLS-2 sample of 1500 this threshold value is going to be 4 percentage points. The gain of the additional 500 observations is a reduction of the threshold value by 0.4 percentage points. Hence, the increase of the sample size does not change much.



The figure also reveals that the gain of additional observations decreases if the baseline proportion decreases, and increases when relative small sample sizes are compared. It should be also noted that similar gains can be also achieved by decreasing the design effect of the sample.

Finally remember that the curves in figure 2 are calculated under the assumption of an EQLS-1 sample size of 1000 observations. In practice, the number of observations in EQLS-1 tends to be somewhat lower due to item non-response. Moreover the small countries in EQLS-1 have sample sizes below 1000 observations. Lower sample sizes in EQLS-1 will shift the lines in the figure upwards. In practice the threshold value for significance will be often somewhat higher than that shown in the figure.

### **Estimation of group specific averages within countries**

Figure 1 can be also used to discuss the precision of estimators for subgroups within countries, for example the proportion of satisfied persons amongst unemployed Germans. Subgroup sizes of less than 100 observations are not rare for certain subgroups of EQLS-1. The confidence intervals around estimators for such rare subgroups are quite large. With 100 observations, the boundaries of the 95% confidence interval around a proportion of 50 percent will be 40% and 60% for a simple random sample, and 33% and 67% for a design effect of three.

The figure also reveals that even few additional observations will be very helpful for getting more precise estimators of these rare subgroups.

Assuming that the Foundation is interested in the quality of life of poor or marginalised persons it is necessary to strive for a sample size that allows precise estimations of quantities for such subgroups. However, increasing the general sample size helps little in this respect. Assuming equal sampling probabilities it requires 10 additional observations to get 1 additional observation of a subgroup, which holds a proportion of 10 per cent in the target population. Hence, increasing the general sample size means to invests most money in the observation of persons that are not from the subpopulation in question.

A different strategy to increase rare subpopulations is to oversample them by design. As it stands, this seems not feasible within the context of the EQLS.

However, there is also a third strategy: poor and marginalised subpopulations are not only rare in the target population, but they also tend to be under-represented in populations surveys due to unit non response. This means that an investment in the reduction of systematic unit non response can be also seen as an investment to increase the sample size of poor and marginalised persons. The Foundation might even consider a targeted strategy, meaning to increase call attempts in areas known to have higher numbers of specific subgroups (where available on or via the sampling frame). As decreasing the unit non response also positively affects the general representativeness of the sample, we believe that the small EQLS resources are better allocated here than in an increase of the general sample size.

### **Sampling design**

The recommendations in this section are based on the assumption that the Foundation follows our recommendation to invest in a better sampling strategy.

The sampling design should be chosen flexibly. We recommend the Foundation to ask for strict probability sampling in the tender (i.e. sampling from a frame). However, we also think that the strictness of the sampling can be as important as the general sampling method. We therefore recommend that the Foundation offers the possibility of random route as 'last resort' if a suitable frame cannot be found and/or some 'satellites' of the major contractor have no other experience than random route.

We remind the Foundation that even in countries where reliable frames exist, pitfalls must be expected. For example, it will be difficult to fully cover people with illegal status. Such systematic losses because of under coverage cannot be ruled out in practice. However, they must be documented carefully.

In the tender, the Foundation should ask for proposals that contain detailed information on the projected sampling designs. Based on this information the applicant that proposes the best overall package might be selected. We recommend to include outside expertise for the selection of the applicant.

The proposals should contain the following minimal information (if applicable):

- Target population
- Sampling design
  - Stratification: Definition of strata. Frame for stratification.
  - Multistage: Number of sampling stages. Definition of sampling units on each stage. Frames used for each stage, together with a short assessment of their over- and under-coverage problems. Application of PPS (probability proportional to size).
  - Justification of random route.

Based on this information, the Foundation might be able to assess the quality of the proposals. However this assumes that the proposals are very detailed. It is therefore necessary that some time is given to the applicants to document and justify their sampling designs. It also assumes that the Foundation spends some time assessing the proposals. The starting date of the fieldwork period should allow for this time.

In the following we give some comments on each point in order that the Foundation can assess the proposals.

### **Target population**

The target population of the Foundation is the resident population of each country aged 18 and above. The Foundation must be aware that there will be a substantial drop out of persons who cannot speak the languages offered by the questionnaire. These drop outs must be coded as such on the contact sheets.

We recommend to set a cut off value for the translation of the questionnaire to a minority language. If a sizeable proportion of people speak a language that is not the native one they should not be excluded. Otherwise the sample in some countries will be seriously compromised. The ESS says the questionnaire should be translated into first languages spoken by 5% of the population and at the very least the Foundation should set 8–10% as their target.

The applicants should indicate how far they are able to sample the given target population. It can be seen a sign of quality if an applicant clearly indicates that the definition of the target population is unsuitable in one country or another.

### **Stratification**

Stratification might or might not be used. If stratification is used, the definition of the strata should be made clear. In addition it should be made clear whether the frames used for each stratum are reliable.

Stratification is welcomed in cases where the stratification variable is linked to the aims of the survey (eg. urban/rural, region).

### Multistage sampling

We recommend a sampling design according to the best available practice of each country, and the experience of the respective fieldwork agency with it. The Foundation should allow SRS as well as sampling in one or more additional stages provided that the definition of the sampling units and the frames from which they have been selected are documented.

In case of sampling units of different size it must be documented what precautions are taken to guarantee an equal probability sampling mechanism (i.e. PPS, weighting).

It is the task of the fieldwork agencies to come up with a proposal of suitable sampling frames. For the assessment of the proposals it might be useful to know some of the frames used for sampling in the respective countries; therefore they are documented in table 1. It must be kept in mind that some of the frames might not be available any more for commercial fieldwork institutes.

Table 1: *Sampling approaches and anticipated outcomes for ESS round 3*

	Unit	Frame	n <sub>net</sub>	n <sub>gross</sub>	n <sub>eff</sub>
Austria	Individuals	Telephone book			
Belgium	Households	National register			
Bulgaria	Households	Electoral register	2,205	4,084	1,500
Cyprus	Households	2001 census	985	1,481	801
Denmark	Individuals	Central Person Register			
Estonia	Individuals	Population register	2,000	2,867	2,000
Finland	Individuals	Population register	2,000	3,000	2,000
France	Area based		1,749	3,600	1,190
Germany	Individuals	Local residents registers	3,056	4,843	1,364
Greece					
Hungary					
Ireland	Addresses	National electoral register	2,250	3,400	1,500
Latvia					
Netherlands	Addresses	Postal delivery points	1,892	3,000	1,580
Poland	Individuals	Personal records of population	1,621	2,574	1,500
Portugal	Area based		2,190	3,295	1,500
Slovakia	Individuals	Central register of citizens			
Slovenia	Individuals	Central register of citizens	1,470	2,250	1,081
Spain	Individuals	Continuous Census	1,904	3,290	1,500
Sweden	Individuals	Population register	1,963	3,000	1,963
Great Britain	Addresses	Postcode address files	2,332	4,608	1,458
Northern Ireland	Addresses	Postcode address files	73	144	46

### Justification for random route

We recommended asking for strict probability sampling in the tender, which by definition excludes the use of random route. However we also recommended that the Foundation should be open to fieldwork agencies that use random route, provided that they explain why they think the preferred sampling design is not suitable.

The justification of random route should also include a description how it is going to be performed. The question is the extent to which random routes can be judged to be 'strictly random'. At the very least, the following questions have to be answered in order to minimise the interviewer's influence on the selection of respondents:

- How are the rules for random routes defined in different countries (and is this comparable)
- What experience do interviewers have with random walks
- How can the whole random walk process be controlled

An appropriate method might involve the interviewer doing the complete walk, recording the sampled addresses and notifying these to the survey office before he/she begins contacting any addresses.

### Unequal sampling probabilities

We recommend abstaining from oversampling rare subpopulations. Such a sampling design would require either a separate sampling frame of the respective subpopulation, or a sampling frame that readily identifies the persons from the subpopulation. We doubt that such frames are available throughout all 29 target countries. Even if they were available we doubt that their suitability can be approved quickly enough before the fieldwork. Finally, disproportional sampling probabilities decrease the precision of the overall sample.

Unequal sampling probabilities are (practically) indispensable for household samples. To weight the data accordingly, it is necessary to ascertain the number of persons in the household that belong to the target population. This number can be used as the 'design weight'. Design weights should be applied before constructing the weights for 'redressment' (aka 'non response weights').

### Response rates

For the sake of comparability we recommend to use the formula of the ESS to calculate response rates, which is

$$\frac{\text{\# of achieved interviews}}{\text{\# of issued sample unit\# - of non eligible sampling units}}, \quad (5)$$

whereby the term 'non eligible sampling units' refer to non response due to

- Address not residential (institution, business/industrial purpose)
- Address not occupied (not occupied, demolished, not yet built)
- Other ineligible address
- Respondent moved abroad
- Respondent deceased

All non response due to

- Refusal by respondent
- Refusal by proxy (or household or address refusal)
- No contacts (after at least ... visits)
- Respondent mentally or physically unable to co-operate throughout the fieldwork period
- Language barrier
- Respondent unavailable throughout the fieldwork period for other reasons

should be not treated as 'non eligible sampling units'.

When comparing the results of a response rate defined like this with the reported response rates of EQLS-1, it should be kept in mind that the EQLS-1 response rates treated non contacts as non eligible sampling units. Hence, the reported response rates of EQLS-1 refer to what has been called the 'cooperation rate' (Groves, 2004), and not to what is normally considered as the response rate. Moreover the respective figures of EQLS-1 are suspected to be biased upwards because of the application of random route sampling (Kohler et al.,2006).

# Fieldwork

## Fieldwork time frame

In EQLS-1 the fieldwork for the survey started much later than originally envisaged and caused quite a few problems in a number of countries as well as a significant delay in delivering the data. The main problem was the summer break (Ahrendt, 2003, 4). For EQLS-2 it is very important that

- fieldwork period is long enough to ensure up to three call backs (i.e. four visits).
- holiday periods are avoided to ensure higher response rate, and
- data across countries are gathered in more or less the same period to ensure comparability among countries (to avoid ‘season effects’ like changes in the unemployment rate – and/or possible important historical events).

As for the length of the fieldwork period, it does not seem to pose any problem to allow for up to eight weeks. It must be noted that many fieldwork agencies are used to work only 2 weeks on one project, and may not be willing to change their habits. However, a period of 4 weeks should be the absolute minimum for any country especially if response rates are to be improved and call patterns are to be achieved at optimal times. The applicants should therefore clearly state how long they plan to actually work on the project.

As for the actual timing, we suggest not to start with fieldwork before mid April. This suggestion is based on the calendar of public and school holidays (see tables 2 and 3), which show that many countries will have school holidays until mid April.

Table 2: *Public holidays*

	AT	BE	CZ	DK	FI	FR	DE	IE	IT	NL	PT	SK	ES	SE	UK
<b>January</b>															
1	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
6					X				X						
<b>March</b>															
17								X							
<b>April</b>															
5				X											
6				X			X			X	X	X	X	X	X
9	X	X	X	X	X	X	X	X	X	X		X		X	X
25									X		X				
30										X					
<b>May</b>															
1	X	X	X		X	X	X		X		X	X	X	X	
4				X											
7								X							X
8			X			X						X			
17	X	X		X	X	X	X			X				X	
21		X													
27	X														
28				X			X			X					X

**Second European Quality of Life Survey: Recommendations for the methodology**

 Table 2: *Public holidays (cont'd)*

	AT	BE	CZ	DK	FI	FR	DE	IE	IT	NL	PT	SK	ES	SE	UK
<b>June</b>															
4								X							
5				X											
6														X	
7	X										X				
23					X										
<b>July</b>															
5			X									X			
6			X												
<b>August</b>															
6								X							
15	X	X				X			X		X				
29												X			
<b>September</b>															
28			X												
<b>October</b>															
5											X				
12													X		
26	X														
28			X												
29								X							
<b>November</b>															
1	X	X				X			X		X	X	X		
3							X								
17			X												
<b>December</b>															
1						X					X				
6													X		
8	X										X				
24			X									X			
25	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
26	X		X	X	X		X	X	X	X		X		X	X

Note: *Holidays are not yet available for Cyprus, Estonia, Greece, Hungary, Latvia, Lithuania, Luxembourg, Malta, Poland, Slovenia, Bulgaria, Romania, Turkey and Croatia.*

Table 3: School holiday periods

	Autumn 2006	Spring/Easter 2007	Summer 2007
<b>AT</b>		31/MAR-10/APR	30/JUN-08/SEP
<b>BE</b>	30/OCT-03/NOV	02-13/APR	01/JUL-31/AUG
<b>BG</b>	09-12/NOV	01-09/APR	23/MAI-15/SEP
<b>CY</b>		31/MAR-15/APR	23/JUN-02/SEP
<b>CZ</b>	26+27/OCT	05/FEB-18/MAR; 05-06/APR	02/JUL- 31/AUG
<b>DE</b>	02/OCT-04/NOV	05/MAR-16/APR	21/JUN-10/SEP
<b>DK</b>	16-22/OCT	02-08/APR	30/JUN -13/AUG
<b>EE</b>	28/OCT-05/NOV	17-25/MAR	05/JUN-31/AUG
<b>Eng.</b>	23-27/OCT	02-13/APR	23/JUL - first week of SEP
<b>NoIr</b>	30-31/OCT	06-13/APR	b 01/JUL - first week of SEP
<b>Scot.</b>	02-27/OCT	31/MAR-13/APR	end of JUN to mid-AUG
<b>ES</b>		30/MAR-16/APR	late/JUN- mid SEP
<b>FI</b>	13-30/OCT	06-09/APR	02/JUN to mid-AUG
<b>FR</b>	25/OCT- 06/NOV	31/MAR-02/MAY	4/JUL- 03/SEP
<b>GR</b>	30/MAR-16/APR		22/JUN-31/AUG
<b>HU</b>	30/OCT-03/NOV	31/MAR-15/APR	15/JUN-31/AUG
<b>IE</b>	30/OCT- 03/NOV	30/MAR-15/APR	late JUN to end of AUG
<b>IT</b>	04-14/APR		12/JUN - SEP
<b>LT</b>	31/OCT - 06/NOV	10-17/APR	JUL-AUG
<b>LU</b>	28/OCT - 05/NOV	31/MAR-15/APR	16/JUL- 14/SEP
<b>LV</b>	23-27/OCT	30/MAR-05/APR	01/JUN-31/AUG
<b>MT</b>	01-3/NOV	04-11/APR	28/JUN- 14/SEP
<b>NL</b>	14-29/OCT		30/JUN-02/SEP
<b>PL</b>		05-10/APR	23/JUN-31/AUG
<b>PT</b>		26/MAR-09/APR	25/JUN+10/11 weeks
<b>RO</b>	30/OCT - 05/NOV	07-15/APR	16/JUN-31/AUG
<b>SE</b>	10/OCT-10/NOV	19/MAR-20/APR	08/JUN - end of AUG
<b>SK</b>	02-03/NOV	19/FEB-09/MAR; 05-10/APR	01/JUL-31/AUG
<b>SL</b>	30/OCT - 03/NOV	27/APR-02/MAY	25/JUN-31/AUG
<b>TR</b>	22-25/OCT	23-24/APR	19/JUN - 18/SEP

Source: *Organisation of school time in Europe. Primary and general secondary education – 2006/07 school year Brussels: Eurydice, 2006, adapted.* Notes: School holidays for 2007/2008 are not yet available; Table list the entire holiday period. Actual holidays can be shorter depending on region and/or age of pupil.

It is also clear that in some countries (like Latvia, Estonia, or Italy) summer school holidays start already at the beginning of June, and in many other countries they start in mid-June. In all countries, school year starts no later than in mid-September. In general there seems to be two possible periods for fieldwork: March-May or mid-September-November. It is also clear from the tables that there are more public holidays in April and May than in October and November. Similarly, spring (Easter) school holidays tend to be longer than autumn holidays (if there are any at all). This has



important consequences for fieldwork planning. Though we support aiming at earliest possible period (March-May), it seems reasonable to postpone fieldwork in all countries to mid-September, especially if during preparation it becomes clear that in some countries the spring period is not realistic (because of questionnaire development delay etc.). In this case, we recommend using the spring time for elaboration of sampling designs, pilots etc. In any event the time-frame should be realistic and should respect national particularities in public and schools holidays.

### Central organisation

We strongly recommend that the Foundation assumes a central and direct role in the commissioning of fieldwork and its monitoring. The Foundation should insist on on-going documentation and transparency of every survey subprocess from translation-backchecking to mapping the documentation of PSUs, the daily or weekly status of interviews and revisits on the PSU levels, etc. The Foundation should insist on the right to send a representative or delegate to attend the local translation, sampling, interviewing, data entry or data processing without prior notice.

On the commissioning side the Foundation should consider developing a short questionnaire on how fieldwork will be organised in each country. As it is likely that the agency selected will 'subcontract' to 'partner' firms this will enable the Foundation to get a clearer idea of what is proposed in each country. In addition they should seek very clear information in the proposals on:

- The type of relationship between the central firm and the partner organisations who conduct the work in each country (including the contracts between them)
- The level and type of supervision the central body will conduct before, during and after fieldwork
- The track record of the central coordinating firm in working with the satellite organisation. We recommend to reject applicants that work together for the very first time.

In terms of fieldwork itself the Foundation should insist on reports demonstrating that the local fieldwork organisations work on the survey according to the contract. This might include weekly reports on

- the number of achieved interviews per country. This demonstrates that the fieldwork actually has started and makes progress. If the number of interviews is behind the schedule the Foundation might check out with the fieldwork organisation the number of interviewers currently working on the project.
- the number of non contacted sampling units. A high number of non contacted addresses might show that not all addresses have been allocated to interviewers, or that some interviewers are unable to start working on the project. The Foundation can insist on immediate allocation of the sampling units to the interviewers. It might be sensible to ask for using more interviewers on the project.
- the response rate, i.e.

$$\text{Resp. Rate} = \frac{\text{\# of completed interviews}}{\text{\# of sampling units selected - ineligible}} \quad (6)$$

If response rates are considerably low response rate maximation, refusal conversion strategies or timing of call backs might be discussed.

To be able to spot uncommon figures on these quantities the Foundation should ask for projections of the target number of interviews to be achieved by each week of fieldwork.

Further guidelines on ways to monitor fieldwork can be found on the ESS Web-site<sup>3</sup>. Whilst a balance needs to be struck between burden on the survey organisation and monitoring, a failure to have such procedures in place will severely limit the ability of the Foundation to identify problems whilst there is still an opportunity to remedy them. Even 20 minutes per country per week should be sufficient to ensure a minimal level of checks.

## **Field work regulations**

We recommend the Foundation to ask in the tender about what the applicants intend to do to decrease survey errors. With this information it will be possible to select the applicant, who offers the best overall package. We recommend to include outside expertise for this selection.

The questionnaire should collect the following information for each participating country:

- Data collector
- Mode of data collection (PAPI vs. CAPI)
- Questionnaire language(es)
- Number of interviewers (experienced/not experienced)
- Mode of interviewer briefing (personal vs. written)
- Payment of interviewers (hourly vs. per interview, Bonus arrangements)
- An estimation of realisable response rates
- Proposed strategies for refusal conversion
- Proposed strategies for increasing the response rate
- Type and amount of control back-checks

Table 4: *Methods to decrease unit non-response*

<b>Method</b>	<b>Utility</b>	<b>Costs</b>	<b>Practicability</b>
Advance letters	++	-	++
Incentives	-	0	++
Converting refusals	++	++	+
Refusal Avoidance Training	+	+	0
Hourly payment rates	++	++	--
Moderate assignment rates	++	+	--
Selection of experienced interviewers	++	++	--
Progress monitoring	++	--	+

---

<sup>3</sup> <http://naticent02.uuhost.uk.uu.net/fieldwork/progress>

Table 5: *Methods to increase data quality*

Method	Utility	Costs	Practicability
Personal briefing of interviewers	++	+	-
Back-checking	++	+	-
Hourly payments	++	++	--
<b>Storing interviewer numbers in the dataset</b>	--	--	+

The applicants might be encouraged to check out the ESS document ‘Field Procedures in the European Social Survey Round 3: Enhancing Response Rates’<sup>4</sup> for developing their strategies for increasing the response rate. We recommend to prescribe the use of advance letters (if applicable), a minimum of 3 call backs (i.e. 4 visits) on different times and weekdays (including weekends) and the storing of interviewer numbers in the dataset as an absolute minimum requirement.

In order to assess the proposals of the fieldwork agencies we document some possible methods with an assessment of their scientific gain, their costs, and their practicability (tables 4 and 5). The meanings of the various approaches are described in some detail in the methodological review (Kohler et al., 2006). The term ‘practicability’ in the table refers to the likely willingness of a fieldwork agency to apply it, even if they do not apply it regularly. In both tables ‘--’ was used for very low, ‘-’ for low, ‘0’ for intermediate, ‘+’ for high and ‘++’ for very high. As mentioned above we also recommend to include outside expertise to assess the proposals of the applicants.

It should be noted that utility and costs of all the methods mentioned in the table depend on the specific way in which they are implemented.

## Field Work Report

The contractor should deliver an extensive fieldwork report. This fieldwork report is an indispensable mean to assess the data quality for each country and to further improve the methodology of future rounds. The fieldwork report should contain information on the following items:

1. Instruments
  - Copy of all questionnaires
  - Copy of all lists, show-cards, etc.
  - Copy of instructions to the interviewers
  - Description of pretest with results
2. Sampling procedure
  - Definition of the target population
  - Description of the frame population

---

<sup>4</sup> <http://naticent02.uuhost.uk.uu.net/fieldwork/response/%20rate/%20enhancement.htm>

- Detailed description of the sampling method, including primary and secondary sampling units, stratification variables, cluster sizes, etc.
- Description of the selection of target persons

### 3. Field work procedures

- Total number of interviewers
- Advance information
- Description of interviewers
- Training of interviewers
- Assignment rate per interviewer
- Call schedules
- Respondent incentives
- Break down of non response (depends on sampling design)
  - Refusal by respondent
  - Refusal by proxy (or household or address refusal)
  - No contacts
  - Address not residential
  - Address not occupied
  - Other ineligible address
  - Respondent moved abroad
  - Respondent deceased
  - Respondent mentally or physically unable to co-operate throughout the fieldwork period
  - Language barrier
  - Respondent unavailable throughout the fieldwork period for other reasons
  - Interviewer failures
- Back-Checking regularities
- Description of other problems during field work

### 4. Data preparation

- Data-entry
- Description of data adjustments
- Copy of validation scripts
- Formulas for design weights and redressment

This standardised report must be given for each country separately.

# Questionnaire

The development of a survey questionnaire primarily follows the research interests, which cannot be commented on methodological grounds. Therefore, methodological recommendations on the questionnaire can be only of general nature. With the exception of socio-demographic variables these general remarks primarily concern the procedures of questionnaire development, and not the questionnaire itself. With respect to questionnaire development we need to distinguish between existing high quality questions and newly developed questions. Generally, the inclusion of existing questions is easier than that of newly developed ones. It should be however noted that there should be at least some validation study available to describe an existing question as a high quality one.

## Old and new questions

For the second round of the EQLS survey the Foundation have indicated their preference that the majority of questions be repeated in identical format to the first round of the study. Some caution should be expressed here. Whilst the desire for repeated measures is clearly understandable some work should be undertaken to evaluate the existing questions. This could include assessment in terms of item non response in round 1 of the survey, reliability and validity testing (Saris et al., 2004) and the extent to which individual items scale and measure the required underlying concepts for attitudinal variables. Background variables should be compared to existing official indicators where known. Some of this assessment has already been conducted in Nauenburg and Mertel (2004). Moreover, the WZB has invited Rainer Schnell, a leading German survey specialist, to comment upon the EQLS-1 questionnaire. Important results of these attempts have been compiled in the appendix of the methodological review (Kohler et al., 2006), and we recommend the questionnaire development team to draw upon these materials.

A number of new areas have been proposed for inclusion in the second Round of EQLS. These are:

- Quality of governance and inclusiveness of the state
- Europeanisation of aspiration levels and concepts of citizenship
- Employment and social capital (need to re-frame existing questions rather than add new ones)
- Migration and ethnicity.

In most of these areas it might be possible to draw from existing surveys that have been conducted cross-nationally. The ESS and ISSP should be considered amongst other sources. Whilst questions should never be transported from one survey to another without careful consideration of their quality, where existing high quality questions can be identified it could reduce the need for extensive piloting.

## Translation

The Foundation should ask for a translation strategy with an emphasis on committee translation and committee assessment (Harkness, 2003, 2006). The committee should consist of (at least) two professional bilinguals, and one survey specialist of the respective country (with English language skills). Back translation might be requested as information for the quality assessment of the translation.

We strongly recommend to place the translated questionnaires on the Internet.

For new items proposed for the questionnaire notes should be included to assist translators to understand the purpose of the questions and the underlying concepts and dimensions which they aim to measure. Annotations are used to elaborate words or phrases which may have specific meanings in English which might not be immediately obvious to a translator.

An example from the ESS Round 3 personal and social well-being module demonstrates this well:

*‘Using this card, please tell me how much of the time during the past week you could not get going’*

The annotation is: ‘not get going’ the sense of ‘felt lethargic and lacked motivation’.

The annotation is only a guide for translators and should not form part of the question wording or be an interviewer note.

In addition it is essential that those involved in developing the new questionnaire items include experts in questionnaire design. Including only substantive experts is likely to result in less than optimally designed questions. Experts on translation should also be consulted to avoid questions that are likely to cause major translation difficulties.

## **Pretest**

Each question in a comparative survey must pass at least three pretests before it can be actually used:

1. A test of whether each question itself is capable to ascertain the information for which the question was intended
2. A test of whether the question can be asked in a survey together with all the other questions in that survey.
3. A test of whether the survey with the specific question can be implemented in each country selected for the entire study.

In what follows we will refer the first type as ‘pretest of new questions’, to the second type as ‘1st pilot study’, and to the third type as ‘2nd pilot study’. The three types should be done one at a time. Only after a question has passed one type, the next step should be performed. However, if a specific question has been used with success in other surveys it is sensible to abstain from the first type of pretesting.

### **Pretests of new questions**

We recommend cognitive testing of new items in the source language.

Around 20 cognitive interviews could be conducted in either the UK or Ireland. This could include a mixture of ‘protocol analysis’ which will involve having respondents ‘think aloud’ as they answer the survey questions and the use of specific probes to test whether envisaged problems with the new questions are causing respondents difficulty. This enables the following types of cases to be identified:

- cases where respondents misunderstood survey questions or key concepts
- cases where respondents did not know the information needed to answer the question from memory
- cases where the respondent could not recall the information needed from memory in a survey interview situation
- cases where respondents used an inappropriate strategy for making a judgement
- cases where respondents chose to hide certain information and provide a socially desirable answer.

Cognitive testing can be fairly expensive (1200 Euro per interview for testing 20 question items and performing high quality analysis) and so drawing on existing questions that have been validated and tested on other high quality surveys would be ideal if they meet the information requirements of the Foundation.

### **1st pilot study**

As long as a fairly limited number of new questions are added to EQLS-2 the first pilot study could be limited to the newly added questions and those where a relationship is expected with existing items (eg. core demographic items). We recommend to conduct the first pilot in either the UK or Ireland (with costs likely to be significantly lower in Ireland) and one other country. The sample size should be a minimum of 200 achieved cases in each country. This will allow analysis of basic frequencies to determine item non response and some descriptive statistics to look at the relationship between the new items and the intended concepts that need to be measured. However with such small samples the power of the estimates will be fairly limited. An interviewer debrief should be conducted and a small respondent debrief could also be considered. Testing in a non English speaking country will enable the source questionnaire to be amended at an early stage and should reduce problems when translations occur at a later stage. If items are drawn from existing sources it might be possible to exclude this stage of piloting and include all countries in the second stage.

### **2nd pilot study**

We recommend to conduct a small tests of the entire translated questionnaire in all countries. Since the number of new items is likely to be small it would be tempting to recommend that only the new items are pretested. However it is likely that new fieldwork partners will be involved (even if the Round 1 agency is re-appointed) and that the structure of the questionnaire will change. And each agency will want to transport the questionnaire into their own 'house style'. So as a minimum it is recommended that each country tests the questionnaire on a sample of 30 interviews using the same sample source as envisaged for the main stage interview.

Feedback on the questions should be focused on existing questions that have caused concern and the new items. However it is essential to test the routing and interviewer instructions in each country. Failure to do this could result in serious data errors that would prove extremely costly in analytical terms. And even if stage 1 has taken place the two countries that participated should also test the full questionnaire before it is fielded but could conduct a smaller number of interviews.

The 2nd pilot study should also be used as an opportunity to test the methodology of contacting respondents and general survey administration. So advance letters, incentives, refusal conversion techniques, contact form records etc. should also form part of this exercise.

## **Background variables**

In making recommendations for the background variables for the EQLS, it is necessary to take account of the fact that the main purpose of the survey is to measure 'quality of life' issues. Further, the budgetary and time constraints under which the survey is conducted also need consideration. Thus, since the survey is not primarily about issues such as income or social class, measurement of these variables should not be to the detriment of the resources available for the quality of life measures or to the detriment of response rates.

Background variables of comparative surveys must have two properties:

- They must measure the same theoretical concept in each country.
- They must measure a concept that is comparative to other sources available for the countries.

Two strategies could lead to background variables that have these properties. One can either ask the same questions (ASQ) or equivalent, but not identical questions (Braun and Mohler, 2003, 104). Which of the two is better is different for each background variable. In the following we propose questions for the several background variables:

### **Household and household size**

Questions regarding the household should make clear the concept of household. This can be either done by a directive to the interviewer, or as a part of the first question on household.

The project instructions of EQLS1 gave a very detailed and clear definition of what has to be considered as a household. We recommend to stick with this definition.

### **Income**

Income should be net income (after tax and other compulsory deductions) – income which is available to the household. There may be some difficulties in collecting data on net income in France, because of the way tax is collected (net pay is usually reported as gross minus compulsory social insurance).

The major problem of ascertaining income is the number of missing values. Asking for household income often results in a high proportion of missing values because it is (a) a sensitive issue, and (b) a question that is difficult to answer.

Different strategies exist to raise the number of valid answers to the question on household income. One strategy is to decrease the number of income categories in order to make the question easier and somewhat less sensitive. The ESS goes in this direction as it uses only 12 categories, while the EQLS-1 uses 19.

Another strategy is to first ask for an exact figure of the income, and to use income categories if this figure is not given. Often this second question is introduced with an explicit reference to the anonymity of the survey. This strategy is used in most countries participating in the International Social Survey Programme (ISSP).

A comparison of the missing values of the ESS and the EQLS has shown that the proportions of missing values in EQLS-1 and ESS are quite similar (Kohler et al., 2006). Moreover, the proportions of missing values in the ISSP are also similar to that of the EQLS-1 or the ESS. A comparison of different German surveys shows that a survey with a simple show-card has the fewest missing values.

In essence we think that the Foundation can stick with the income question of EQLS-1.

### **Employment status**

For the purpose of the Labour Force Surveys, the focus is on the person's economic activity in a specific reference period (the last seven days). The ILO approach to economic status prioritises employment/self-employment over other activities such as home duties or education/training. In other words, if the person is a student who worked a few hours in the previous week, they are regarded as 'at work'. The ILO measure of unemployment is quite demanding in a survey, as it requires items on job search activity and availability for work as well.

An alternative approach is to ask the persons 'usual situation' with regard to employment, and an additional item on whether the person worked more than a certain number of hours (15 was used in the ECHP) in the previous week. This approach was adopted by the EQLS-1, but without the additional inquiry on the working hours. We recommend to stick with it for the sake of comparability.



The EQLS asked the respondent to choose his 'principal economic status'. The wording 'economic status' is a technical term that does not fit the answer categories. It should be substituted by a better term from common speech, like, 'the usual situation of employment'.

### Education

In EQLS-1 question no 47 which asked for the highest level of education the respondent completed produced some irregular results which seriously hamper data comparability. Only four countries produced consistent results by the yardstick that at least 90% of those who named primary education as their highest level of schooling also report to have left full-time education below age 16. In eight countries less than half of those with primary education reported to have left the school system under age 16. This caused analysts of the EQLS-1 data to work with the terminal age of education as the more reliable measure of formal educational training.

Given the differences within Europe in the age of starting school, the use of the terminal education age as a measure for education has been discouraged.

We recommend a combination of the 'years of full-time education' (or full-time equivalent) and a collection of nationally specific categories. These nationally specific categories should be collected in such a way as to allow conversion into ISCED centrally. However, this imposes additional work on the central survey agency, and it presupposes that an adequate mapping of the national qualifications onto ISCED is available (Hoffmeyer-Zlotnik, 2003). For most countries it is possible to clone to mappings of the European Social Survey. The ESS also provides countries with details on how to map specific qualifications to ISCED based on information from Eurostat. This could be useful for EQLS too.

National categories and how they are converted into ISCED should be specified in technical report.

Here are the respective questions of the ESS Questionnaire (3rd round, F6, F7)

What is the highest level of education you have achieved? Please use this card. [Card 48; Card show country specific educational certificates]

About how many years of education have you completed, whether full-time or part-time? Please report these in full-time equivalents and include compulsory years of schooling.

We like to emphasise that while the ISCED is fairly standard as a classification of education in international surveys, it is not without problems. These problems relate to disagreements (often a matter of national pride) regarding the mapping of national-level qualifications onto ISCED, not to the data collection itself.

An advantage of using country-specific categories for education is that it allows the construction of better non-response weights.

### Occupation/Social class

For occupation we recommend to adapt the question the ESS uses for parent's occupation (ESS-3, F54). It provides a sensible, well tested alternative to the 'gold standard' European Socio-economic Classification (ESeC) and the EQLS-1 operationalisation, which is not easily comparable to the ESeC. An adapted version of this question for the respondent might be:

Which of the descriptions on this card best describes the sort of work you do [show card]?

- Traditional professional occupations such as: accountant; solicitor; medical practitioner; scientist; civil/mechanical engineer
- Modern professional occupations such as: teacher; nurse; physiotherapist; social worker; welfare officer; artist; musician; police officer (sergeant or above); software designer
- Clerical and intermediate occupations such as: secretary; personal assistant; clerical worker; office clerk; call centre agent; nursing auxiliary; nursery nurse
- Senior manager or administrators (usually responsible for planning, organising and co-ordinating work and for finance) such as: finance manager, chief executive
- Technical and craft occupations such as: motor mechanic; fitter; inspector plumber; printer; tool maker; electrician; farmer; gardener; train driver
- Semi-routine manual and service occupations such as: postal worker; machine operative; security guard; caretaker; farm worker; catering assistant; receptionist; sales assistant
- Routine manual and service occupations such as: HGV driver; van driver; cleaner; porter; packer; sewing machinist; messenger; labourer; waiter/waitress; bar staff
- Middle or junior managers such as: office manager; retail manager; bank manager; farm manager; restaurant manager; warehouse manager; publican
- (Don't know)

It should be noted that a similar question seems to have been asked in the pilot study for EQLS-1. As a result of the pilot study this original question (W10) has been changed to the question Q15/16 of the final questionnaire. Without the exact question wording of W10 we cannot comment on what has happened in the pilot.

### Urban/Rural

We recommend storing low level geographic identifiers (NUTS-3) in the dataset to ensure comparability with the EUROSTAT data. The identifiers might be used to merge population densities from official statistics, which are available at the NUTS-3 level.

The inclusion of NUTS-3 will also allow merging several other official statistics available on the NUTS-3 level, like for example infant mortality. Inclusion of NUTS-1,2 and 3 will also allow comparisons of statistical European regions, which have direct policy implications (eligibility of regions for objective 1 under the structural funds is set at NUTS-2 level).

If the Foundation follows our recommendation, it is essential that the tender specification mentions that the NUTS information is required.

### **Migration background/Ethnicity**

With respect to the operationalisation of migration background it has to be kept in mind, that there will be substantial selective drop out due to language problems of immigrants. We therefore only propose a very rough indicator. Essentially this is a shortened version of the questions C26-C36 in the ESS Round 3 questionnaire (European Social Survey, 2006).

1. Are you a citizen of [country]?
2. Were you born in [country]?
3. Do you belong to a minority ethnic group in [country]?
4. Was your father born in [country]?
5. Was your mother born in [country]?

See [http://naticent02.uuhost.uk.uu.net/questionnaire/questionnaire\\_development/chapter\\_10.doc](http://naticent02.uuhost.uk.uu.net/questionnaire/questionnaire_development/chapter_10.doc) for details on these questions.

### **Religion**

We propose a shortened version of the ESS battery on religion (European Social Survey, 2006, C17-C24).

1. Do you consider yourself as belonging to any particular religion or denomination?
2. Which one? [Country-specific codes for recoding into standard coding frame]
3. Regardless of whether you belong to a particular religion, how religious would you say you are? Please use this card.

See the development report [http://naticent02.uuhost.uk.uu.net/questionnaire/questionnaire\\_development/chapter\\_09.doc](http://naticent02.uuhost.uk.uu.net/questionnaire/questionnaire_development/chapter_09.doc) for details on these questions.

# Data

## Data entry

The Foundation should leave it up to the fieldwork agency to use their preferred way of data entry. The applicants should clearly indicate what sort of data entry they plan to use. Advantages and disadvantages of data entry are given below.

- Transferring data from the questionnaire with a data entry system. The main disadvantage is the risk of transmission errors. Professional data entry institutes often use double entry to control the inputted data. Depending on the software used to input the data, error and consistence checks can be applied. A recommended system for data entry is the freeware program EpiData (<http://www.epidata.dk>). An alternative are self programmed data entry masks in a language like PHP, which allows to input the data in a SQL database (i.e. MySQL) over the Internet. It will require a substantial amount of programming to implement consistency checks of this way of data entry, however.
- Machine readable questionnaires can be used to automate the data entry process completely. The advantage is that the data entry will be finished rather quickly and error-free. The main disadvantage is the high price of hard and software. Moreover, machine readable questionnaires requires accurate work of the interviewers when marking the answers. Finally, the necessity to include some specific markers in the questionnaire can make respondents suspicious.
- CAPI. In computer aided personal interviews the data entry step is bypassed by entering the answers of the respondent directly into the computer. The main advantage of this technique is that the data sets can be finalised rather quickly. However, it is not a feasible strategy in some countries due to insufficient technical equipment. Moreover entry errors can result in a serious disturbance of the interview. However, a well programmed CAPI instrument is very unlikely to cause major disturbances during an interview.

## Consistency checks

We recommend that the Foundation keeps some resources back for consistency checks after the delivery of the data by the fieldwork agency. These consistency checks should be done independently from the field-work agency by scientific data analysts. The results of the consistency checks, and the decisions taken should be documented on the Internet.

## Methodology datasets

The fieldwork agency should provide an additional data set for doing methodological research. This data set should include

- Variables for the correct calculation of standard errors (e.g. strata identifiers, sampling unit identifiers, design weights)
- Variable to control the data quality (e.g. variables from an interviewer questionnaire)

**Ulrich Kohler**, Wissenschaftszentrum Berlin für Sozialforschung (WZB); **Rory Fitzgerald**, Centre for Comparative Social Surveys, City University, London; **Arnošt Veselý**, Centre for Social and Economic Strategies, Charles University, Prague; **Dorothy Watson** and **Toney Fahey**, Economic and Social Research Institute, Dublin; **Femke DeKeulenaer**, Gallup Organisation Europe, Brussels; with the collaboration of Robert Manchin, Gallup Europe.

# References

- Ahrendt, D., The Quality of Life Survey. On behalf of the European Foundation for the Improvement of Living and Working Conditions. Field-work Technical Report of Intomart GfK, Hilversum, 2003
- Braun, M. and Mohler, P. P., *Background Variables*. In Harkness, J. A., Van de Vijver, F. J. R., and Mohler, P. P., editors, *Cross-Cultural Survey Methods*, Wiley Series in Survey Methodology, pages 101–115. Hoboken, Wiley, 2003
- European Social Survey (2006). Final Source Questionnaire. Round 3, 2006/7. Amendment 03.  
<http://naticent02.uuhost.uk.uu.net/questionnaire/mainquestionnaire.htm>.
- Groves, R. M., *Survey Errors and Survey Costs*. Wiley Series in Survey Methodology. New York, Wiley, 2004
- Harkness, J. 2003. *Questionnaire Translation*. In Harkness, J., Van de Vijver, F., and Mohler, P. P., editors, *Questionnaire Translation*, pages 35–56. Hoboken, Wiley, 2003
- Harkness, J. A., Round 3 ESS Translation Guidelines, 2006  
<http://naticent02.uuhost.uk.uu.net/essdocs/R3/Methodology/r3translationguidelines.pdf>.
- Hoffmeyer-Zlotnik, J. H., *Education as a Sociological Background Characteristic*. In Hoffmeyer-Zlotnik, J. H. and Wolf, C., editors, *Advances in Cross-National Comparison. A European Working Book for Demographic and Socio-Economic Variables*, pages 245–256. New York, Kluwer Academic, 2003
- Kohler, U., *Internal Criteria for Representativeness for European Comparative Surveys*. Submitted to. *Survey Research Methods*, 2006
- Kohler, U., *Quality Assessment of European Surveys. Towards an Open Method of Coordination for Survey Data*. In Alber, J., Fahey, T., and Saraceno, C., editors, *Social Conditions in the Enlarged Version*. Routledge, 2007
- Kohler, U., Fitzgerald, R., Vesely, A., Watson, D., and De Keulenaer, F. Technical assistance in preparing second European quality of live survey. Methodological review. Paper prepared for the European Foundation for the Improvement of Living and Working Conditions, 2006
- Nauenburg, R. and Mertel, B. European Quality of Life Survey. Technical Report. Draft version. Paper distributed by the Wissenschaftszentrum Berlin, 2004
- Saris, W. E., Satorra, A., and Coenders, G. A New Approach to Evaluating the Quality of Measurement Instruments: The Split-Ballot MTMM Design. *Sociological Methodology*, 37:311–347, 2004