



*Workshop Cross-national Survey Harmonization and Analysis
Weights, Data Quality and Multi-level Modeling*

May 13, 2015

Assessing the Quality of Survey Documentation

Marta Kołczyńska

Matthew Schoene

Survey data are NOT created equal

- Cornerstones of Data Quality (Hox and de Leeuw 2008)
 - Coverage.
 - Sampling.
 - Specification.
 - Non-response.
 - Measurement.
- How to find out about this?

Edith D. de Leeuw
Joop J. Hox
Don A. Dillman



International
Handbook of Survey
Methodology



Survey Documentation

Survey Documentation provides insight into *how* the data were collected, including information on:

- The type of sampling.
- Size of the sample.
- Response rate.
- Preparation of questionnaire and other instruments.
- Pretesting.
- Fieldwork control.



Survey Documentation

Where can you find this documentation?

- Online text.
- Study descriptions.
- Methods reports.
- Technical reports.
- Codebooks.
- Random places.

Why is documentation important?

(1) Assumption that better documented surveys usually offer higher quality data – but problems exist.

- WAPOR offers guidelines, but they are rarely followed (Ingelhart 1997).
- And, fully documented surveys are very rare (Mohler, Pennell & Hubbard 2008).
- Inconsistencies exist in the ISSP and WVS programs (Medrano 2009; Scholz & Heller 2009).

Why is documentation important?

(2) Documentation is an integral part of data.

Data = “the physical representation of information **in a manner suitable for communication, interpretation, or processing** by human beings or by automatic means”

(UNSC 2000: 6)

Data

- For information to qualify as data, it needs to be usable.
- Usability of survey data depends on the availability and the high-quality of documentation
- Information about sampling, response rate, translation of the questionnaire, pretesting and fieldwork control cannot be found in the numerical data files (data sets)

Documentation Quality Measures

Measures of survey quality **as expressed in survey documentation.**

Information collected in the Harmonization project:

- Sampling,
- Response rate,
- Preparation of the questionnaire,
- Pretesting,
- Fieldwork control.

Sampling (ISSP 2011 SI)

Study description

Target population: adult residents of Slovenia, older than 18 years, living on permanent address. Excluded: Institutionalised people.

Sampling frame: Central Register of Population (a list of names and addresses constantly updated by public administration).

Sampling procedure: two-stage stratified random sample from Central Register of Population, where every population unit has equal probability of selection.

First stage: PSU selection is made by probability proportional to size of CEA (Clusters of Enumeration Areas) (150 PSUs). CEA are stratified according to 12 regions*6 type of settlement.

Second stage: systematic random selection inside CEA brings fixed numbers of persons (150x24) with name and address. Split-halves samples were used for parallel SJM surveys (2x1800).

Sampling (ISSP 2011 US)

Study Description

Multi-stage area probability sample.

Quality

Measures of survey quality as expressed in survey documentation

Set of 4 dichotomous indicators:

- Response rate
- Questionnaire translation method
- Pretesting/piloting
- Fieldwork control/back-checking

On the level of **national survey** (e.g. ISSP 2005 RU)

Response Rate

- 1 – documentation contains information about the response rate achieved, or information sufficient to compute the response rate.
- 0 – otherwise.

Response Rate

definition

- Response Rate = full interviews / full & partial interviews + non-interviews (refusal + break-off + non-contacts + others) + all cases of unknown eligibility.
 - 4 more definitions of response rates.
 - 4 definitions of cooperation rates.
 - 3 definitions of refusal rates.
 - 3 definitions of contact rates.

Response Rate

definition

“For non-probability samples, response rate calculations make little sense, given the broader inferential concerns. Further, for many of these surveys, the denominator is unknown, making the calculation of response rates impossible”

Source: Standard Definitions report (7th edition, 2011), p. 32, aapor.org

Response Rate (WVS 2005 CY)

study description

<i>Sample type:</i>	A sample of 600 people aged from 18 to 70+ years old
<i>Fieldwork Institute:</i>	Intercollege Survey Research Center KADEM
<i>Fieldwork methods:</i>	Face-to-face interviews
<i>Sample size:</i>	1,050
<i>Response rate:</i>	85%
<i>Language:</i>	Turkish, Greek, English

Response Rate (WVS 2005 CY)

Methodology Questionnaire

Total number of <u>starting</u> names/addresses	1	2	6	5
- addresses which could not be traced at all				0
- addresses established as empty, demolished or containing no private dwellings				0
- selected respondent too sick/incapacitated to participate				0
- selected respondent away during survey period				0
- selected respondent had inadequate understanding of language of survey				0
- no contact at selected address				0
- no contact with selected person				0
- refusal at selected address			6	5
- proxy refusal (on behalf of selected respondent)				0
- personal refusal by selected respondent				0
- other type of unproductive (please write in full details in the box below)				0
- full productive interview	1	0	5	0
- partial productive interview		1	5	0

$$1050 / 1265 = 0.83$$

$$1200 / 1265 = 0.949$$

Response Rate (ISSP 2010 IL)

study description

„Interviews: 1023

These figures pertain to interviews in Jewish and Mixed (Jewish-Arab) communities. In the case of additional 193 interviews conducted in small Arab communities there was no sampling list and we have no information on response rates”

Translation Method

1 – documentation contains information about the method of questionnaire translation (any documented method more sophisticated than translation by the survey team).

0 – otherwise; includes:

Documentation includes information that no translation method was used.

Documentation does not include information about translation method at all.

Response Rate (ISSP 2011)

Methods Report

	BE-F	BE-W	CL	HR	CZ	DK	FI	FR	DE	GB	IL	JP	KR
<i>Language(s) of the fielded module</i>													
Language 1 (L1)	Dutch	French	Spanish	Croatian	Czech	Danish	Finnish	French	German	English	Hebrew	Japanese	Korean
Language 2 (L2)		Dutch					Swedish				Arabic		
Language 3 (L3)											Russian		
<i>Was the questionnaire translated?</i>													
Yes, translated:													
- by member(s) of research team	X	L1-L2	X	X	X	X		X			L1-L3	X	X
- by translation bureau	X	L1-L2											
- by specially trained translator(s)							L1-L2		X			X	
- other						X ^{DK}							
No, not translated										X			

Response Rate (ISSP 2011)

Methods Report

	BE-F	BE-W	CL	HR	CZ	DK	FI	FR	DE	GB	IL	JP	KR	LT
<i>Language(s) of the fielded module</i>														
Language 1 (L1)	Dutch	French	Spanish	Croatian	Czech	Danish	Finnish	French	German	English	Hebrew	Japanese	Korean	Lithuanian
Language 2 (L2)		Dutch					Swedish				Arabic			
Language 3 (L3)											Russian			
<i>Was the translated questionnaire assessed/checked or evaluated?</i>														
Yes:														
- group discussion			X	X	X			X	X		L1-L3		X	
- expert checked it	X ^{BE-F}	X ^{BE-W}					L1-L2		X		L1-L3 ^{IL}	X		
- back translation				X										
- other	X ^{BE-F}	X ^{BE-W}				X								X
No											L2 ^{IL}			
Not applicable										X				

Pretesting

- 1 – documentation contains information about pretesting/piloting
- 0 – otherwise; includes:
 - documentation contains information about no pretesting having been carried out
 - documentation does not contain information about pretesting

Pretesting EQLS 2012)

Technical Report

Table 11a EQLS Questionnaire test phases in the EU27 countries

COUNTRY	PRE-TEST COGNITIVE	PRE-TEST LIVE EQLS INTERVIEWS	PILOT (FIELD TRIAL)
AT			X
BE	X	X	X
BG			X
CY			X
CZ			X
DE			X
DK			X
EE			X
EL			X
ES			X
...			...

Pretesting (EQLS 2012)

Technical Report

Table 11b EQLS Questionnaire test phases in the non-EU countries

COUNTRY	COGNITIVE INTERVIEWS (PREPARATORY PHASE)	PILOT INTERVIEWS (FIELD TRIAL)
TR	X	X
HR	X	X
MK	X	X
KO	X	X
RS	X	X
ME	X	X
IS		X

Fieldwork Control

- 1 – documentation contains information about fieldwork control/backchecking
- 0 – otherwise; includes:
 - documentation contains information about no fieldwork control having been carried out
 - documentation does not contain information about fieldwork control

Fieldwork Control (ESS 2010 PL)

Data Documentation Report

44 Control operation

44.1 Interviews

Number of units selected for back-check:	320
Number of back-checks achieved:	205
Number of units where outcome was confirmed:	205
Type of back-check: Personal or Telephone:	Telephone

44.2 Refusals

Number of units selected for back-check:	70
Number of back-checks achieved:	55
Number of units where outcome was confirmed:	41
Type of back-check: Personal, Telephone or Mail:	Personal and telephone

44.3 Non-contacts

Number of units selected for back-check:	50
Number of back-checks achieved:	34
Number of units where outcome was confirmed:	29
Type of back-check: Personal, Telephone or Mail:	Personal and telephone

Fieldwork Control (EVS 2000 BY)

source book

28e. Were any interviews back-checked?

1. Yes, + approximate proportion
2. No

	at	be	bu	by	cz	de	dk	ee	es	fi	fr	gr
1	30	16	5			20	5	10	5	5	na	na
2				x	x							na

To Repeat

4 indicators of documentation quality:

1. Response rate
2. Translation method
3. Pretesting
4. Fieldwork control

0/1