



INSTITUTE FOR SOCIAL RESEARCH • SURVEY RESEARCH CENTER
SURVEY RESEARCH OPERATIONS
UNIVERSITY OF MICHIGAN

Assessing Quality of Survey Processes and Outputs: The 4th European Quality of Life Survey

Julie de Jong (U-M)

Kristen Cibelli Hibben (U-M)

Dan Zahs (U-M)

Beth-Ellen Pennell, University of Michigan (U-M)

Brita Dorer (GESIS-Leibniz Institute for the Social Sciences)

CSDI
Limerick, Ireland
March 26-28, 2018



The 4th EQLS Quality Assessment

- EQLS fieldwork completed in March 2017
- Tender invitations issued in April 2017
- Contract awarded to University of Michigan team in June 2017
 - Beth-Ellen Pennell
 - Kristen Cibelli Hibben
 - Julie de Jong
 - Dan Zahs
 - Brita Dorer (GESIS)
- Assessment to be completed by the end of March 2018



Survey Quality in 3MC

- Total survey error + Comparison error
 - Identifies error sources and estimates their relative magnitude; useful for evaluating tradeoffs (Groves et al., 2009; Smith, 2011; Pennell et al., 2017)
- Fitness for intended use
 - Study design meets user quality requirements (i.e., survey data accuracy, comparability, etc.) (Biemer & Lyberg, 2003)
- Survey process production quality
 - Quality of the processes that generate the product (Lyberg & Biemer, 2008; Lyberg & Stukel)

(Hansen et al., 2016)

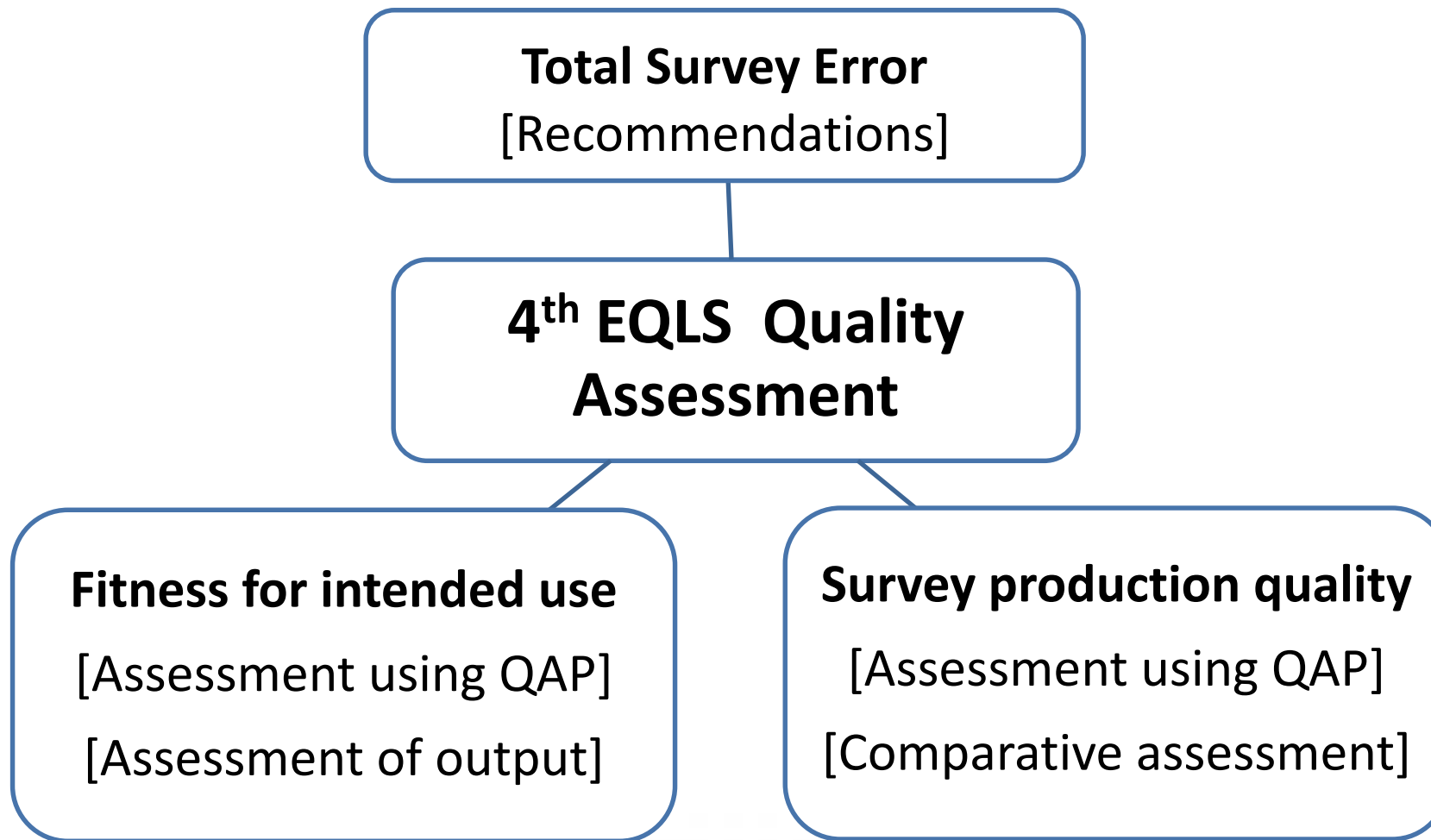


Assessment Components

- Achievement of targets in the Quality Assurance Plan
- Assessment of survey outputs
- Comparative assessment based on 3MC best practices
- Summary findings and recommendations



Integration of Quality Approaches





Quality Assurance Plan

145 total indicators organized by:

- Achievability
 - Ideal-world
 - Real-world
 - Required



Quality Assurance Plan

145 total indicators organized by:

- Achievability
- Quality dimension
 - Accuracy (73 indicators)
 - Accessibility (34 indicators)
 - Punctuality (31 indicators)
 - Coherence and comparability (3 indicators)
 - Relevance and timeliness (1 indicator)



Quality Assurance Plan

145 total indicators organized by:

- Achievability
- Quality dimension
- Survey lifecycle stage
 - Sampling
 - Questionnaire
 - Translation
 - Pilot
 - Data entry
 - Fieldforce training
 - Fieldwork
 - Fieldwork infrastructure
 - Data processing
 - Microdata
 - Paradata
 - Weighting

Quality Assurance Plan

Example Indicators

Theme	Sub-theme	Key quality dimension addressed	Indicator	Target
Sampling	Register vs. enumeration	Accuracy	Percentage of countries where a high quality register is used (up-to-date and approximating full coverage)	100%
	Sampling frame (country)	Accuracy	Percentage of countries where the sampling frame covers at least 95% of the populations	100%
	Reference statistics (overall)	Accessibility	Percentage of countries for which the characteristics of the reference statistics are documented in complete accordance with the template	100%
Translation	Initial translation	Punctuality	Initial translation delivered at agreed date	Y
Fieldforce training	Construction of interviewer training materials	Accuracy	Training materials cover strategies for convincing reluctant respondents	Y

Assessment Using Quality Assurance Plan

- Categorized indicators into four broad categories
 - Sampling frame development
 - Questionnaire development and advance translation, cognitive testing, and translation
 - Fieldwork (Implementation, monitoring, contact procedures, nonresponse, and paradata)
 - Weighting
- Assigned score based on compliance
 - Target met
 - Target not met
 - Outcome unknown
 - Target no longer applicable (N/A)



Assessments of Outputs

- Source questionnaire assessment
- Nonresponse statistics
- Assessment of sociodemographic sample composition
- Evaluation of other output statistics



Assessments of Outputs

- Source questionnaire assessment
 - Non-sensitive questions about behavior
 - Make questions as specific as possible
 - Use words that virtually all respondents will understand
 - Sensitive questions about behavior
 - Embed the sensitive question among other sensitive items to make it stand out less
 - Use familiar words in describing sensitive behavior
 - Attitudinal questions
 - Specify the attitude object clearly
 - Avoid double-barreled questions

Sources: Chapter 7 in Groves et al., 2009; Chapters 3 – 6 in Saris and Gallhofer, 2014.



Assessments of Outputs

- Source questionnaire assessment
- Nonresponse statistics
 - Comparison to 3rd EQLS
&
– Comparison to ESS Round 8 (Round 7 as applicable)
 - Response rates
 - Refusal rates
 - Cooperation rates
 - Contact rates



Assessments of Outputs

- Source questionnaire assessment
- Nonresponse statistics
- Assessment of sociodemographic sample composition
 - Comparison to Eurostat statistics
 - Age/gender distributions
 - Comparison to ESS Round 7 data
 - Age/gender distributions
 - Household size
 - Employment status



Assessments of Outputs

- Source questionnaire assessment
- Nonresponse statistics
- Assessment of sociodemographic sample composition
- Evaluation of other output statistics
 - Coefficient of variation
 - Design effects

Comparative Assessment Based on Current Best Practices in 3MC Research

- Developed list of best practices for each phase
 - Consulted survey methodology literature
 - Consulted other 3MC survey processes
- Identified process followed by ESS
- Assessed process followed by 4th EQLS

Best practices: Sampling frame development

1. Comparable target and survey populations must be defined and documented for each country.
2. Sampling frames in each country must be identified and evaluated with consideration given to the accuracy of frame.
3. In the absence of an existing sampling frame meeting accuracy criteria, a sampling frame best covering the target population, given budget constraints, must be developed.
4. If the sampling frame is at the level of a household, then a procedure to randomly select elements from the sampling frame must be determined.
5. The sample size necessary to meet the desired level of precision must be determined for each country.

Best practices: Questionnaire development

1. Research question(s) or objective(s) must be clearly defined.
2. Subject-area experts, area/cultural specialists, linguistic experts, and survey research experts should be a part of the questionnaire development team or process.
3. A translatability assessment or, ideally, an advance translation process should be carried out to make the source questionnaire as easy as possible to translate into other languages and to implement in other cultures.
4. An analysis plan should be produced relating each survey question to one or more of the research questions.

Best practices: Questionnaire development

5. A team translation approach, ideally following the TRAPD model, should be followed to translate the source questionnaire into target languages.
6. An appropriate set of pretesting and/or post-hoc evaluation methods should be used to assess the quality of questions, based on available resources.
7. Develop a documentation scheme for questionnaire design decisions and changes to the source questionnaire across time for repeat surveys.
8. Show cards should be produced for those survey items as needed, for use by interviewers in all participating countries following a standard protocol.

Best practices: Fieldwork

1. A standard CAPI instrument, with components for both data collection and sample management should be used in all countries.
2. A standard CAPI instrument should be developed centrally and then thoroughly evaluated in all countries.
3. A checklist of minimum interviewer candidate requirements should be established and a comprehensive, standardized interviewer training must be conducted.
4. The use of incentives for participation should be determined and documented in each country.
5. A standard pretest protocol should be developed and implemented in each country.

Best practices: Fieldwork

6. Mode of first contact should be standardized across all countries.
7. At minimum, a partially-interpenetrated interviewer field assignment plan should be developed and implemented to permit estimation of interviewer effects.
8. Identify both computer-generated and interviewer-generated paradata to be collected and develop clear analysis procedures for the different types of paradata.
9. A data-driven assessment protocol for the selection and verification of cases should be established and include thorough documentation for both selection rationale and verification outcome.
10. A nonresponse bias analysis should be conducted for all participating countries.



Best practices: Weighting

1. The following survey weights should be constructed, as needed, and fully documented:
 - Design or base weights to correct for different probabilities of selection.
 - Weights to adjust for under coverage, nonresponse, and to make weighted sample estimates conform to external values.
 - Supranational or population size weights to adjust for different national population sizes.

Best practices: Weighting

2. Consider weight trimming or other methods for addressing widely varying survey weights.
3. A guide should be provided to assist end users with the correct use of survey weights.
4. If comparison over time is an important goal, the weighting strategy should be kept as consistent as possible for multi-wave surveys or modified weights should be produced and made available based on the most recent weighting methodology.

Summary Findings and Recommendations

- Summary findings
 - Strengths and challenges
- Recommendations and opportunities for improvement
 - Recommendations for each survey lifecycle phase

Table 1. Prioritization of Recommendations for Sampling Frame Development Activities

	Low cost		High cost		Source of error
	High impact	Low impact	High impact	Low impact	
Recommendation 1	X				Nonresponse error
Recommendation 2	X				Sampling error
Recommendation 3		X			Coverage & nonresponse error
Recommendation 4		X			Comparison error
Recommendation 5			X		Sampling error



Summary Findings and Recommendations

- Summary findings
 - Strengths and challenges
- Recommendations and opportunities for improvement
 - Recommendations for each survey lifecycle phase
 - Recommendations for web add-on
 - Recommendations for disclosure and dissemination issues
 - Recommendations for approach to quality assurance



Broader implications for survey quality or survey quality assessment for 3MC surveys

- Possible consensus around set of best practices
- Starting point for indicators to assess survey quality in 3MC research



**INSTITUTE FOR SOCIAL RESEARCH • SURVEY RESEARCH CENTER
SURVEY RESEARCH OPERATIONS**

UNIVERSITY OF MICHIGAN