

Cross-National Comparison on Measurement Quality of Response Scales

2016 3MC

Pei-shan Liao, RCHSS, Academia Sinica

Willem E. Saris, RECSM, UPF

Diana Zavala-Rojas, RECSM, UPF

Introduction

- The quality of a survey measure has been an important issue, as it is closely associated with the reliability and validity of survey questions.
- The recently developed split-ballot multitrait-multimethod (SB-MTMM) approach has been used to evaluate the measurement quality of questions in surveys.
- The SB-MTMM approach has been applied to the European Social Survey (ESS) to examine the quality of questions across countries.

Purpose

- Whether the same quality of survey instruments can be achieved in both a different cultural context and in a logographic writing system, like the one in Taiwan, remains unknown.
- Using the data from the SB-MTMM design in the Taiwan Social Change Survey and corresponding data from 2002 ESS Round 1 in Denmark, this study aims to compare the measurement quality of different response designs across countries.

Split-Ballot Experiment

- The basic principle of this approach is to randomly divide the respondents into two or more equal-sized subsamples with equal representativeness of the total sample.
- The respondents of each subsample answer survey questions simultaneously and under the same conditions.
- Variations in the questionnaire for each of the subsamples are treated as experimental stimulus to examine questionnaire effects.

Multitrait-Multimethod (MTMM) Design

- Suggested by Campbell and Fiske (1959) to evaluate the validity of social science concepts based on the intercorrelations among measures of variables.
- The classic MTMM approach requires a respondent to answer questions about a minimum of three traits that are measured with three different methods, leading to nine different observed variables.
- It becomes a burden for the respondents to repeatedly answer similar questions and may cause memory bias or order effect of the questions.

Split-Ballot MTMM (SB-MTMM) Approach

- An approach developed by Saris et al. (2004) to reduce the response burden by means of using different combinations of two methods in multiple groups.
- The use of multiple groups is similar to the split-ballot design, while the MTMM approach provides information about the reliability and validity of the different questions.
- It enables researchers to evaluate measurement bias, reliability, and validity simultaneously, while reducing response burden.

The SB-MTMM Design

- This design uses three methods to measure three traits (i.e. concepts) as in the classical MTMM design.
- Random samples of the same population are also used as in the split-ballot experiments, but each respondent needs to answer the questions of the same trait only twice.
- In order to minimize the carry-over effect from previous answer, it is suggested that at least 20 minutes are required between the administrations of the related items.

Model Estimation

- The estimation model for the SB-MTMM design can be assumed to be the same as the one for the standard MTMM approach.
- The use of a minimum of three traits to be repeated using at least three methods serves the purpose of identification.

Model Estimation

- The total quality of a measure can be computed as $q_{ij}^2 = r_{ij}^2 \times v_{ij}^2$, where q_{ij}^2 represents the explained variance of the observed variable by the latent trait of interest.
- With respect to the multiple groups in the SB-MTMM design, estimates for the parameters of the model can be obtained using structural equation modeling for multiple-group analysis.

Research Design-TSCS

- A two-group SB-MTMM design is adopted in the 2015 Taiwan Social Change Survey (TSCS).
- The sample is randomly divided into two subsamples based on the respondent's number as odd (Sample 1) or even (Sample 2).

	Method 1	Method 2	Method 3
Method 1	Samples 1		
Method 2	NONE	Sample 2	
Method 3	Sample 1	Sample 2	Sample 1+2

Research Design-ESS1_DK

- The same two-group SB-MTMM design is adopted in 2002 ESS Round 1.
- Data from Denmark (ESS1_DK) that conducted face-to-face interviews using CAPI for both main questionnaire and supplementary questions for the SB-MTMM are used for the comparison.

Measures: Satisfaction with Politics

- On the whole how satisfied are you with the present state of the economy in [country]?
- Now thinking about the [country] government, how satisfied are you with the way it is doing its job?
- And on the whole, how satisfied are you with the way democracy works in [country]?
- M_1 : 4-point scale with full labels; M_2 and M_3 : from 0 to 5 and from 0 to 10, respectively, with end points labelled as “extremely dissatisfied” and “extremely satisfied.”

Differences in Implementation

- One difference in the experimental design between 2015 TSCS and ESS1_DK is that all the respondents in the ESS1_DK got M1 first, and then M2 and M3 for samples 1 and 2, respectively.
- An unfolding technique, in which interviewers first asked about direction and then about the degree of attitudes (Schaeffer and Presser, 2003), is used for M₁ in Taiwan while in the ESS a direct question was used.

Results: Sample Description

- The distributions of demographic characteristics are similar in age and gender. Both have a majority of the respondents aged 60 years or older and equal proportions of males and females.
- The distributions of marital status and education differ significantly.
- Other characteristics including self-reported health status and interest in politics, as well as the items of satisfaction with politics, also show different distributions between 2015 TSCS and ESS1_DK.

Results: Correlation Matrices- Sample 1

ESS1_DK

	M ₁			M ₃		
	Q1	Q2	Q3	Q1	Q2	
M ₁						
Q1	1					
Q2	.407	1				
Q3	.277	.292	1			
M ₃						
Q1	.408	.257	.180	1		
Q2	.256	.670	.174	.448	1	
Q3	.282	.268	.465	.417	.403	1

	M ₁			M ₃		
	Q1	Q2	Q3	Q1	Q2	Q3
M ₁						
Q1	1					
Q2	.685	1				
Q3	.331	.398	1			
M ₃						
Q1	.513	.497	.258	1		
Q2	.505	.638	.323	.726	1	
Q3	.289	.350	.537	.428	.481	1

2015 TSCS

Results: Correlation Matrices- Sample 2

ESS1_DK

	M ₂			M ₃		
	Q1	Q2	Q3	Q1	Q2	
M ₂						
Q1	1					
Q2	.552	1				
Q3	.489	.432	1			
M ₃						
Q1	.548	.332	.265	1		
Q2	.362	.722	.180	.434	1	
Q3	.311	.267	.547	.386	.338	1

	M ₂			M ₃		
	Q1	Q2	Q3	Q1	Q2	Q3
M ₂						
Q1	1					
Q2	.687	1				
Q3	.391	.482	1			
M ₃						
Q1	.670	.628	.350	1		
Q2	.591	.747	.403	.781	1	
Q3	.285	.329	.631	.388	.413	1

2015 TSCS

Results: Reliability and Validity

	Reliability						Validity					
Method	M ₁		M ₂		M ₃		M ₁		M ₂		M ₃	
Country	D	T	D	T	D	T	D	T	D	T	D	T
Q1	.55	.55	.74	.72	.58	.72	.79	.86	.92	.98	.74	.81
Q2	.74	.69	.90	.87	.77	.79	.87	.90	.94	.98	.81	.83
Q3	.91	.62	.62	.85	.76	.62	.76	.88	.90	.98	.81	.81
Ave.	.73	.62	.75	.81	.70	.71	.81	.88	.92	.98	.79	.82

*: "D" denotes the ESS1_DK data and "T" the 2015 TSCS.

Results: Quality

	M ₁		M ₂		M ₃	
Country	D	T	D	T	D	T
Q1	.43	.47	.68	.71	.45	.58
Q2	.64	.62	.85	.85	.62	.66
Q3	.69	.54	.55	.83	.62	.50
Ave.	.59	.54	.69	.80	.56	.58

*: "D" denotes the ESS1_DK data and "T" the 2015 TSCS.

Concluding Remarks

- Using a two-group design, the results indicated questions measured by 6-point scales with labels at end points (M_2) have the best quality while the measures on either 4-point scale with full labels (M_1) or 11-point scale (M_3) perform equally acceptable.
- Although differences between Danish and Taiwanese data can be observed, the findings are comparable, despite that the order of applying methods differed.

Discussion

- One possibility for the relatively poor quality of M_1 in 2015 TSCS may be attributed to the unfolded inquiry procedure that are different from the other methods during the face-to-face interview.
- No general conclusions can be achieved about the effect of different aspects of the methods on the quality of questions. For the effects of the methods we refer to the results of meta-analysis (Saris and Gallhofer 2007/2014).
- Finally, a three-group design for the SB-MTMM approach is suggested for its better performance and efficiency.