Measurement errors in survey research SQP 2.0
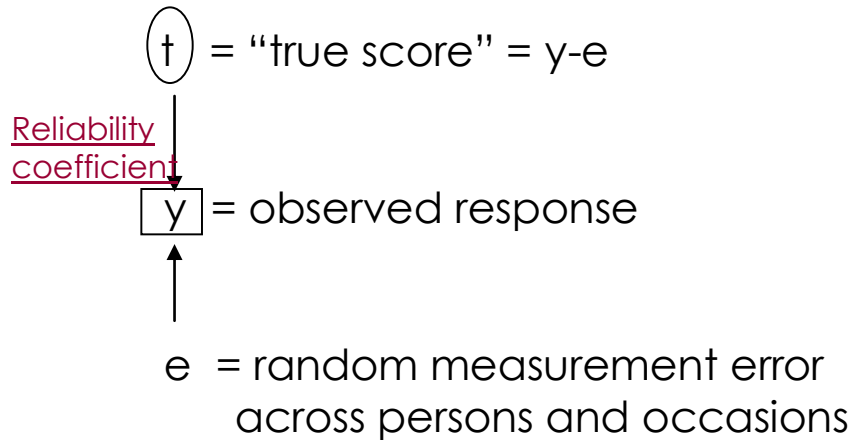
Willem Saris + ... + Diana Zavala

CSDI 2013

Content of this presentation

✓Definition of reliability, validity and quality in terms of random and systematic errors

✓MTMM approach to estimation of reliability, validity and quality
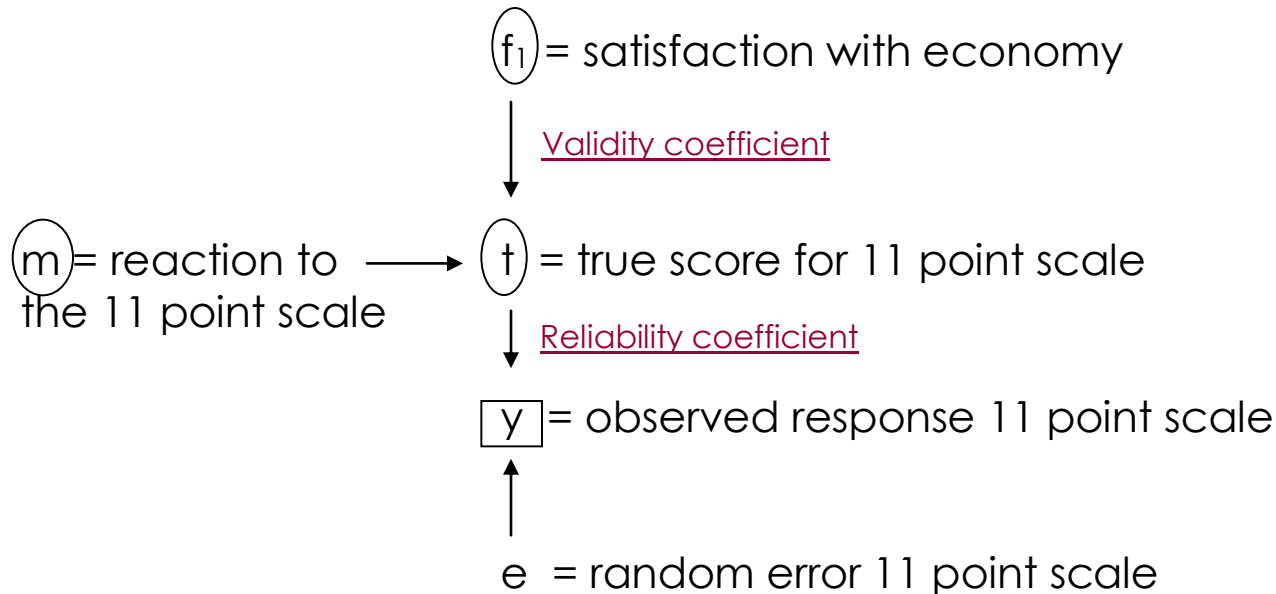
✓Application

✓Generalization: SQP 2.0

Definition of reliability, validity and quality in terms of random and systematic errors

The Classical test model

$\widehat{t}$ = "true score" = y-e

<u>Reliability coefficient</u>

$\boxed{y}$ = observed response
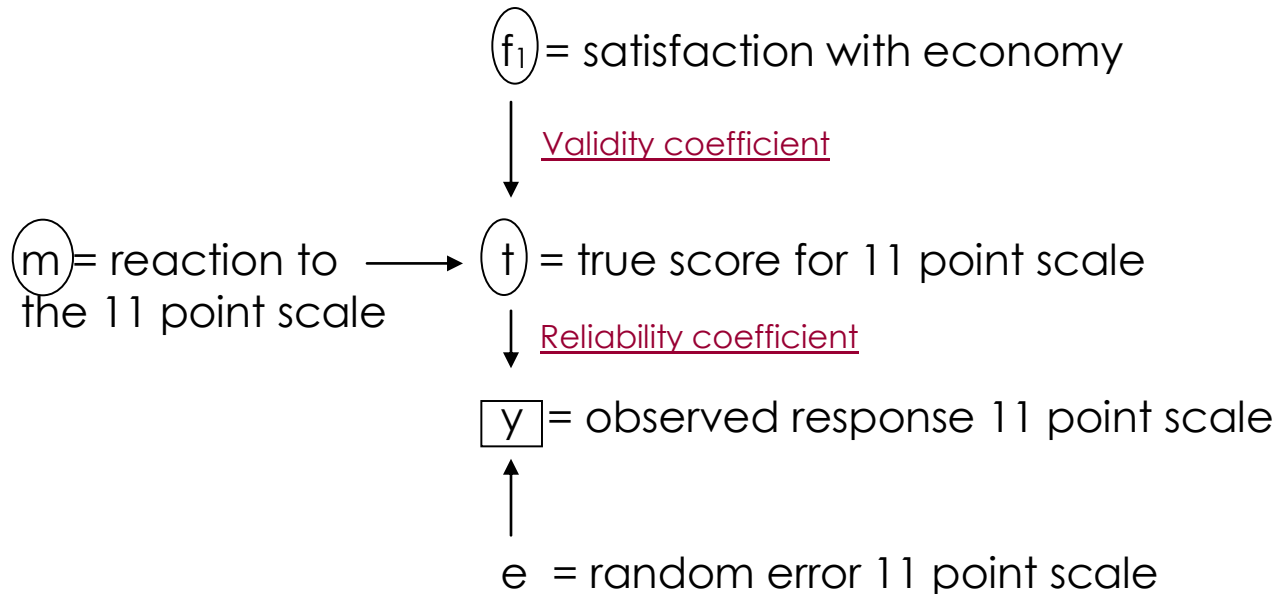
e = random measurement error across persons and occasions

- **Reliability =** (reliability coefficient)^2 = strength of the relationship between t and y **= 1 - var(e)**

Definition of quality

f_1 = satisfaction with economy

↓ Validity coefficient

m = reaction to → t = true score for 11 point scale
the 11 point scale

↓ Reliability coefficient

y = observed response 11 point scale

↑

e = random error 11 point scale

- **Quality =** strength of the relationship between y and f **= reliability*validity**

Definition of quality

f_1 = satisfaction with economy

Validity coefficient

m = reaction to the 11 point scale → t = true score for 11 point scale

Reliability coefficient

y = observed response 11 point scale

e = random error 11 point scale

- **Quality =** strength of the relationship between y and f **= reliability*validity**

Measurement model for two traits, same method
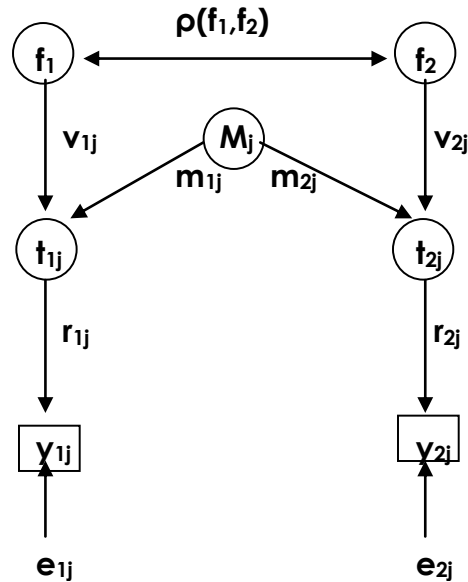
f_i = i^{th} variable of interest
v_{ij} = validity coefficient for variable i
M_j = method factor for both variables
m_{ij} = method effect on variable i

t_{ij} = true score for y_{ij}

r_{ij} = reliability coefficient

y_{ij} = the observed variable

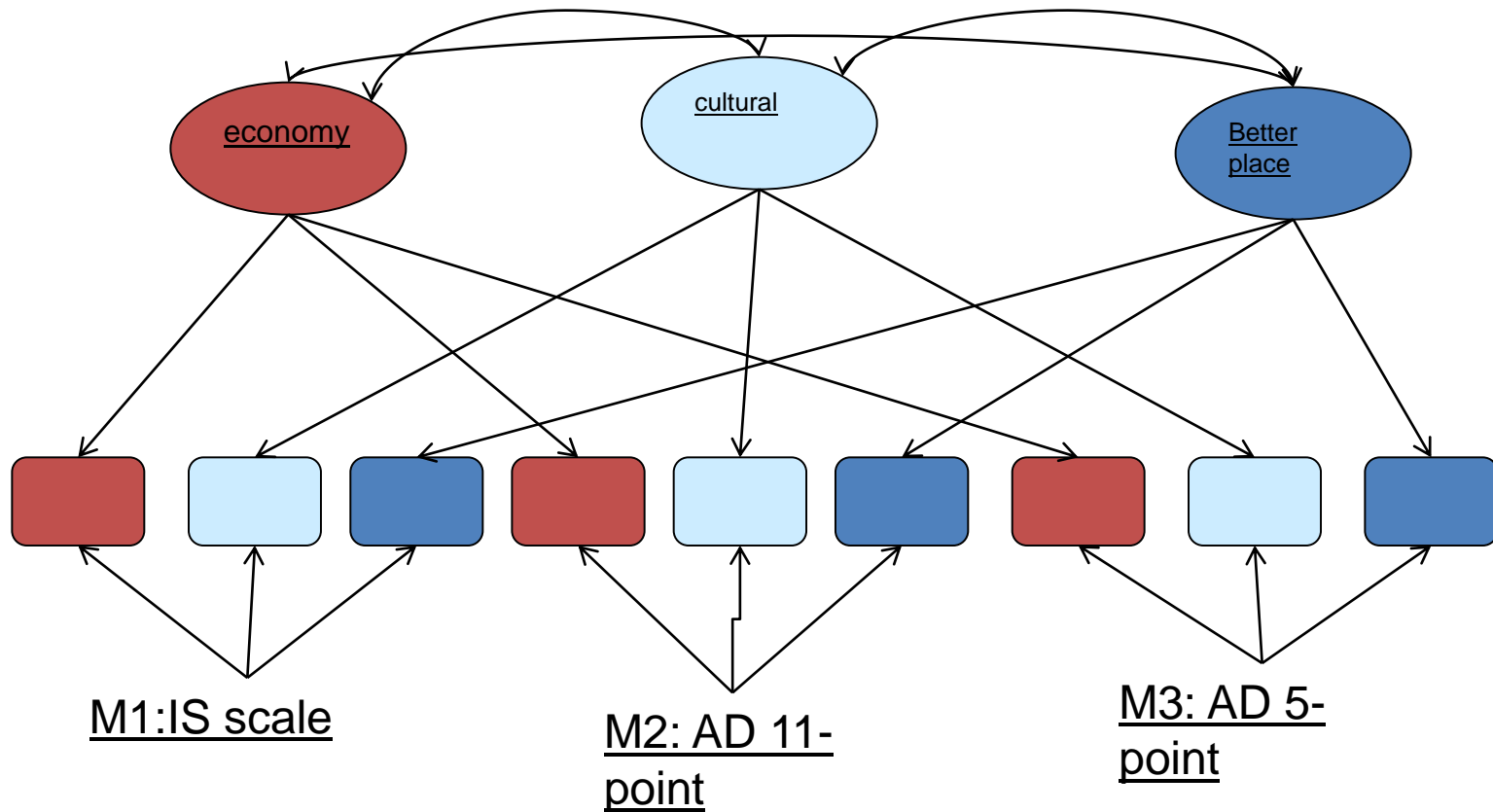e_{ij} = the random error in variable y_{ij}

When are they equal?

$$\rho(y_{1j}, y_{2j}) = r_{1j}v_{1j}\rho(f_1, f_2)v_{2j}r_{2j} + r_{1j}m_{1j}m_{2j}r_{2j}$$

Corr between observed variables

True corr

MultiTrait-MultiMethod (MTMM) approach to estimation of reliability, validity and quality

MTMM on attitudes to immigrants

Campbell and Fiske (1959), Andrews (1984), Saris and Andrews (1991), Saris et al. (2005)

Application

Item specific versus Agree/disagree scales
(IS)   (A/D)

The experiment

- MTMM experiment: attitudes toward immigrants

- 3 traits measured with 4 methods:
  – IS scale
  – 5-point A/D scale
  – 7-point A/D scale
  – 11-point A/D scale

- IS scale: in main questionnaire

- The 3 A/D scales
  – presented to 3 random subgroups of the sample in the supplementary questionnaire of the ESS

The IS scales

B38 CARD 15 Would you say it is generally bad or good for [country]'s economy that people come to live here from other countries? Please use this card.

Bad for the economy										Good for the economy	(Don't know)
00	01	02	03	04	05	06	07	08	09	10	88

B39 CARD 16 And, using this card, would you say that [country]'s cultural life is generally undermined or enriched by people coming to live here from other countries?

Cultural life undermined										Cultural life enriched	(Don't know)
00	01	02	03	04	05	06	07	08	09	10	88

B40 CARD 17 Is [country] made a worse or a better place to live by people coming to live here from other countries? Please use this card.

Worse place to live										Better place to live	(Don't know)
00	01	02	03	04	05	06	07	08	09	10	88

Results of the MTMM analyses

- The mean quality of 3 questions over 23 countries for different methods

Method	Q1	Q2	Q3
IS(11)	.76	.81	.76
	(.07)	(.05)	(.08)
A/D(5)	.44	.45	.47
	(.11)	(.11)	(.12)
A/D(11)	21	28	32
	(.13)	(.15)	(.13)
A/D(7)	.27	.31	.35
	(.11)	(.12)	(.13)

- IS scales much better quality than A/D scales
- Also, IS scales have lower standard deviations

The more general results

- 4 similar experiments
- more than 20 countries
- more than 23 languages

In general:

– **Quality IS scale is .2 better than quality A/D scale**

– **Quality of the A/D scales does not improve by adding more than 5 categories**

* [Saris, Revilla, Krosnick and Schäfer (2009) SRM
And: Revilla, Saris and Krosnick (forthcoming) SMR]

RECSM
Research and Expertise Centre for **Survey** Methodology

MTMM: Conclusions

Pros:

- allow estimation of reliability and validity of question

- Allow separating random and systematic errors → can correct for them

- Best approach for subjective variables

Cons:

    - They cannot be done for all types of questions

    - Results for specific questions: generalization problem

    - Repeat 3 times the questions: burdens and costs

Therefore an alternative has been developed: SQP

Generalization:
Survey Quality Prediction 2.0
sqp.upf.edu

Data

- Quite some MTMM experiments have been done already

- In the ESS:
– In the supplementary questionnaire
– In each round, 4 to 6 MTMM experiments have been done
– In around 25 countries
– Each experiment: 3 or 4 methods * 3 traits

➔ Thousands of quality estimates available in the program SQP 2.0

- Some experiments done before the ESS began too

General idea

- For each of the questions being part of an MTMM experiment, we can code a certain number of characteristics

  – Number of words of the question
  – Language
  – Labels
  – Number of response categories
  – Presence of interviewer
  – Topic
  – Social desirability
  – etc

[Saris and Gallhofer, 2007]

General idea

- From the MTMM experiments:
  – estimates of reliability and validity for 4000 questions

- For these 4000 questions:
  – 60 question characteristics are coded

$\rightarrow$ file with the quality of 4000 questions and their characteristics

- This information can be used to get an optimal prediction of the quality criteria on the basis of the questions characteristics

[Saris and Gallhofer, 2007]

Results: SQP 2.0

- SQP 2.0 makes predictions of the quality of **new** items not involved in MTMM experiments and not involved in ESS

- Many different languages available

- Not only provides point estimates but also confidence intervals for the predictions

- Provides also in a simple way suggestions for improvements of the questions

- User friendly interface

4 ways of using the new program

1. For ESS questions involved in MTMM experiments
– Quality estimate from the MTMM, questions codes and prediction

2. Questions of the ESS not involved in MTMM experiments
- Text available: users have to code the questions to get a prediction

3. Questions that are formulated for new studies
– The text of the question should be introduced
– User has to code the question to get a prediction

4. Support translation processes detecting deviations in the formal characteristics of questions

A library of survey questions

- SQP 2.0 will contain a growing database of survey questions with quality information

- The users together built up this question library

- If one does not trust the prediction by another user one can always recode the question to get a new prediction

- Let us have a look at the program

- sqp.upf.edu

More information

- Saris and Gallhofer (2007). *Design, evaluation and analysis of questionnaires for survey research*. Wiley.

- Saris, Oberski, Revilla, Zavala, Lilleoja, Gallhofer, and Gruner (2011). "Final report about the project JRA3 as part of ESS Infrastructure". RECSM working paper 24.

RECSM

RECSM

Research and Expertise Centre for Survey Methodology

www.upf.edu/survey
recsm@upf.edu