

SURVEY QUALITY PREDICTOR 2.1

Diana Zavala-Rojas
Universitat Pompeu Fabra,
Barcelona

3MC Conference 2016
July 28th, 2016

Survey Quality Predictor (SQP 2.1)



Willem
Saris



Melanie
Revilla



Irmtraud
Gallhofer



Tom
Gruner



Daniel Oberski
(Tilburg
University)



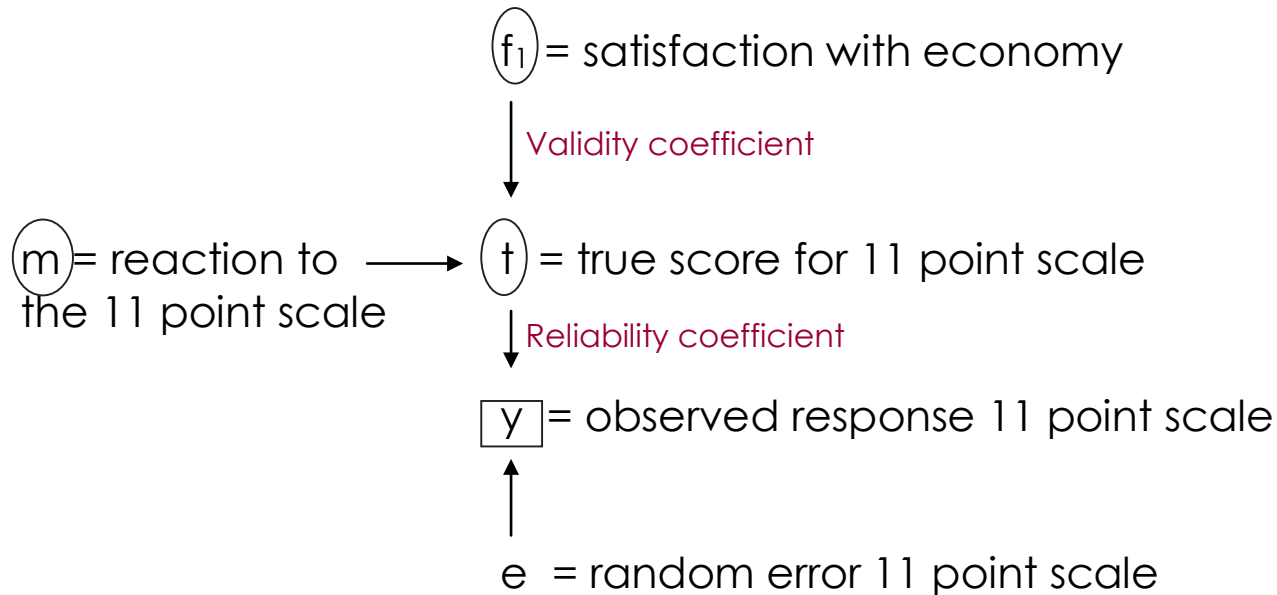
Diana
Zavala-Rojas



Laur Lilleoja
(Tallinn
University)

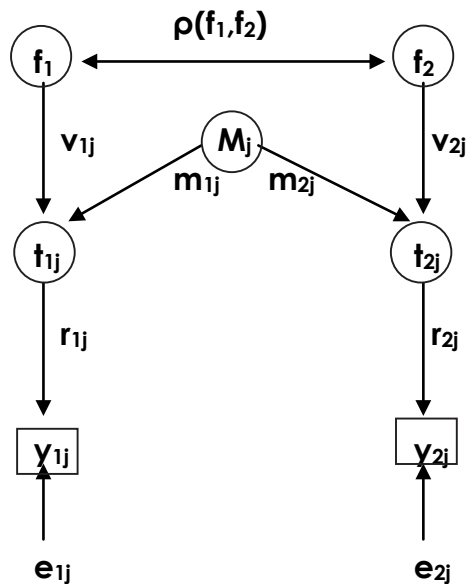
1. Definition of reliability, validity and measurement quality
2. Multi Trait-MultiMethod (MTMM) approach to estimation of reliability, validity and quality
3. Generalization: Survey Quality Prediction 2.1

Definition of measurement quality



- **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?

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

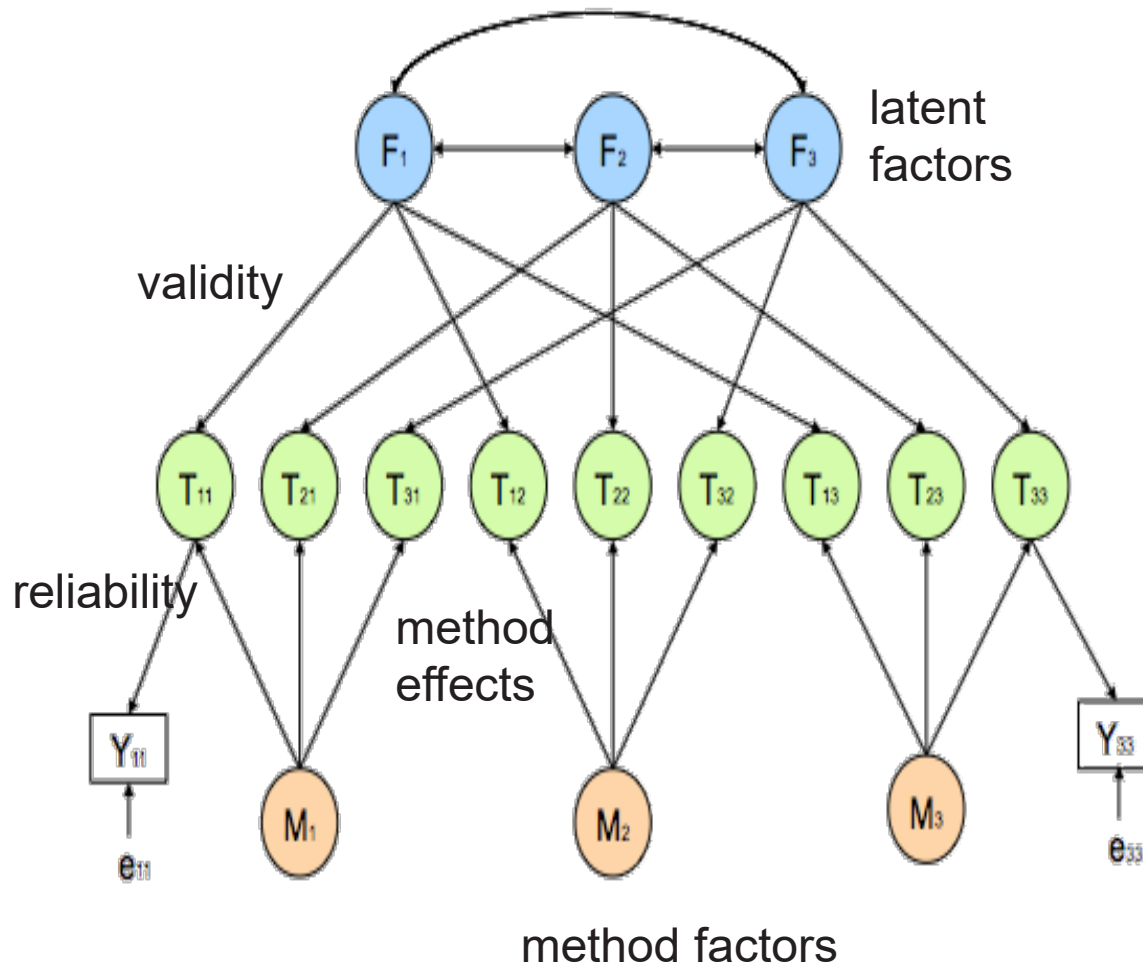
Correlation between observed variables

True correlation

**MULTITRAIT-
MULTIMETHOD (MTMM)
APPROACH TO
ESTIMATION OF
RELIABILITY, VALIDITY
AND QUALITY**

- Campbell and Fiske (1959):
 - Idea of repeating several traits using several methods
 - But no real model
 - Direct analysis of the MTMM matrix
- Andrews (1984)
 - Classical factor model for 3 traits and 3 methods
 - reliability and validity coefficients for all 9 questions and the correlations between the traits can be estimated
- Saris and Andrews (1991)
 - True Score MTMM model
- Saris, Satorra and Coenders (2004)
 - Split Ballot True Score MTMM model

MTMM True Score Model



Pros:

- allow estimation of reliability and validity of question
- Allow separating random and systematic errors → can correct for them
- Good approach for subjective variables

Cons:

- They cannot be done for all types of questions
- Results for specific questions: generalization problem
- Repeat 2 times the questions: burdens and costs

GENERALIZATION: SURVEY QUALITY PREDICTION 2.1

- Many MTMM experiments have been done already
- In the ESS:
 - In the supplementary questionnaire
 - In each round, 2 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 2.1
- Some experiments done before the ESS began too

- 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
 - For of the response scale
 - Salience
 - Presence of interviewer
 - Topic
 - Social desirability
 - Etc..

- From the MTMM experiments:
 - estimates of reliability and validity for 4000 questions
 - For these questions:
 - About 60 question characteristics are coded
- Measurement quality of questions and their characteristics
- This information can be used to get an optimal prediction of the measurement quality on the basis of the questions characteristics

- SQP 2.1 makes predictions of the quality of **new** items not part of MTMM experiments and not part of the ESS
- Many different languages available
- Not only provides point estimates but also confidence intervals for the predictions
- User friendly interface

- SQP 2.1 is growing database of survey questions with information about their measurement quality
- The users together build up this library of questions
- If one does not trust the prediction by another user one can always recode the question to get a new prediction

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

The Survey Quality Predictor (SQP) is:

- **an extensive open-source database of survey questions and quality estimates** built up through the collaboration of the users. The SQP database contains a wide range of survey questions concerning many different topics in many different forms and languages.
- **a coding system of formal and linguistic characteristics of survey questions which allows a prediction of their reliability, validity and quality to be obtained.** This prediction is based on a meta-analysis of the relationships between the quality estimates of survey questions obtained through Multitrait-Multimethod (MTMM) experiments and the formal and linguistic characteristics of the questions in those experiments.
- **a tool for improving questions.** By providing information about the quality of different question formats, the software can help design better questions.



The **quality** of a survey question is defined, in this context, as the strength of the relationship between the latent variable of interest and the response to the survey question, the observed variable. The quality can be computed by taking the product of reliability and validity.

The information provided by SQP is particularly useful to:

- **consult, compare and evaluate questions** which are part of the SQP database,
- **design new questionnaires**, since SQP provides suggestions of changes in question format for improving the quality of questions.
- **correct for measurement error** in the substantive analyses.

To know much more about SQP, check the **SQP tutorials**:



Thank you

sqp@upf.edu
<http://sqp.upf.edu>

diana.zavala@upf.edu