Survey data quality in different countries

Daniel Oberski

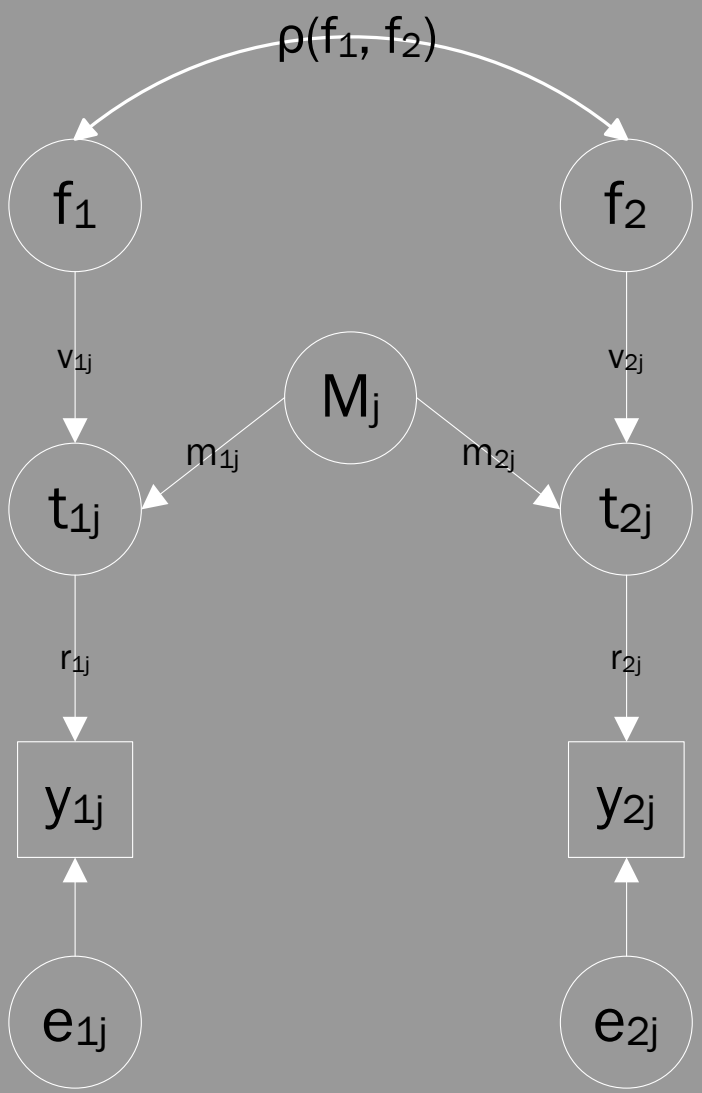
Faculty of Social and Behavioural Sciences
Tilburg University

Survey Research Centre
ESADE, Barcelona
Universitat Ramon Llull

Overview

1. Multitrait-multimethod experiments
   - An example experiment
   - Models
2. What has been done before
   - The international research project 1984–1996
   - Experiments in the European Social Survey
3. Why are there differences between countries?
4. The final goal: SQP

The basic response model



f_1,f_2 = variables 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}

First trait measured with three methods

CARD 73 Using this card, please tell me how true each of the following statements is about your current job.

		Not at all true	A little true	Quite true	Very true	(Don't know)
G64	There is a lot of variety in my work.	1	2	3	4	8

iS19 The next 3 questions are about your current job. Please choose one of the following to describe how varied your work is.

Please tick one box.

- ☐ 1 Not at all varied
- ☐ 2 A little varied
- ☐ 3 Quite varied
- ☐ 4 Very varied

iS32 Please indicate, on a scale of 0 to 10, how varied your work is, where 0 is not at all varied and 10 is very varied.

Please tick the box that is closest to your opinion

Not at all varied										**Very varied**
0	1	2	3	4	5	6	7	8	9	10
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

Three traits measured with first method

CARD 73 Using this card, please tell me how true each of the following statements is about your current job.

		Not at all true	A little true	Quite true	Very true	(Don't know)
G64	There is a lot of variety in my work.	1	2	3	4	8
...						
G66	My job is secure	1	2	3	4	8
...						
G70	My health or safety is at risk because of my work.	1	2	3	4	8

Three traits measured with second method

iS19 The next 3 questions are about your current job. Please choose one of the following to describe how varied your work is.
Please tick one box.

- ☐ 1 Not at all varied
- ☐ 2 A little varied
- ☐ 3 Quite varied
- ☐ 4 Very varied

iS20 Please choose one of the following to describe how secure your job is.
Please tick one box.

- ☐ 1 Not at all secure
- ☐ 2 A little secure
- ☐ 3 Quite secure
- ☐ 4 Very secure

iS21 Please choose one of the following to say how much, if at all, your work puts your health and safety at risk.
Please tick one box.

- ☐ 1 Not at all at risk
- ☐ 2 A little at risk
- ☐ 3 Quite a lot at risk
- ☐ 4 Very much at risk

Three traits measured with third method

Skip details of the model

iS32 Please indicate, on a scale of 0 to 10, how varied your work is, where 0 is not at all varied and 10 is very varied.
Please tick the box that is closest to your opinion

Not at all varied ***Very varied***

0 1 2 3 4 5 6 7 8 9 10

iS33 Now please indicate, on a scale of 0 to 10, how secure your job is, where 0 is not at all secure and 10 is very secure.
Please tick the box that is closest to your opinion

Not at all secure ***Very secure***

0 1 2 3 4 5 6 7 8 9 10

iS34 Please indicate, on a scale of 0 to 10, how much your health and safety is at risk from your work, where 0 is not at all at risk and 10 is very much at risk.
Please tick the box that is closest to your opinion

Not at all at risk ***Very much at risk***

0 1 2 3 4 5 6 7 8 9 10

Different models for MTMM experiments

- Classic MTMM model
- Correlated uniqueness (Kenny & Judd)
- Direct product (Browne)
- True score model
- MTM-1 (Eid 2000)

Different models for MTMM experiments

- Classic MTMM model
- Correlated uniqueness (Kenny & Judd)
- Direct product (Browne)
- True score model
- MTM-1 (Eid 2000)

Different models for MTMM experiments

- Classic MTMM model
- Correlated uniqueness (Kenny & Judd)
- Direct product (Browne)
- True score model
- MTM-1 (Eid 2000)

Different models for MTMM experiments

- Classic MTMM model
- Correlated uniqueness (Kenny & Judd)
- Direct product (Browne)
- True score model
- MTM-1 (Eid 2000)

Different models for MTMM experiments

- Classic MTMM model
- Correlated uniqueness (Kenny & Judd)
- Direct product (Browne)
- True score model
- MTM-1 (Eid 2000)

Different models for MTMM experiments

- We use the true score model
- Equivalent to the classic MTMM model
- Sometimes necessary to remove one method factor
- In that case our model is the equivalent to Eid's MTM-1 model.

Different models for MTMM experiments

- We use the true score model
- Equivalent to the classic MTMM model
- Sometimes necessary to remove one method factor
- In that case our model is the equivalent to Eid's MTM-1 model.

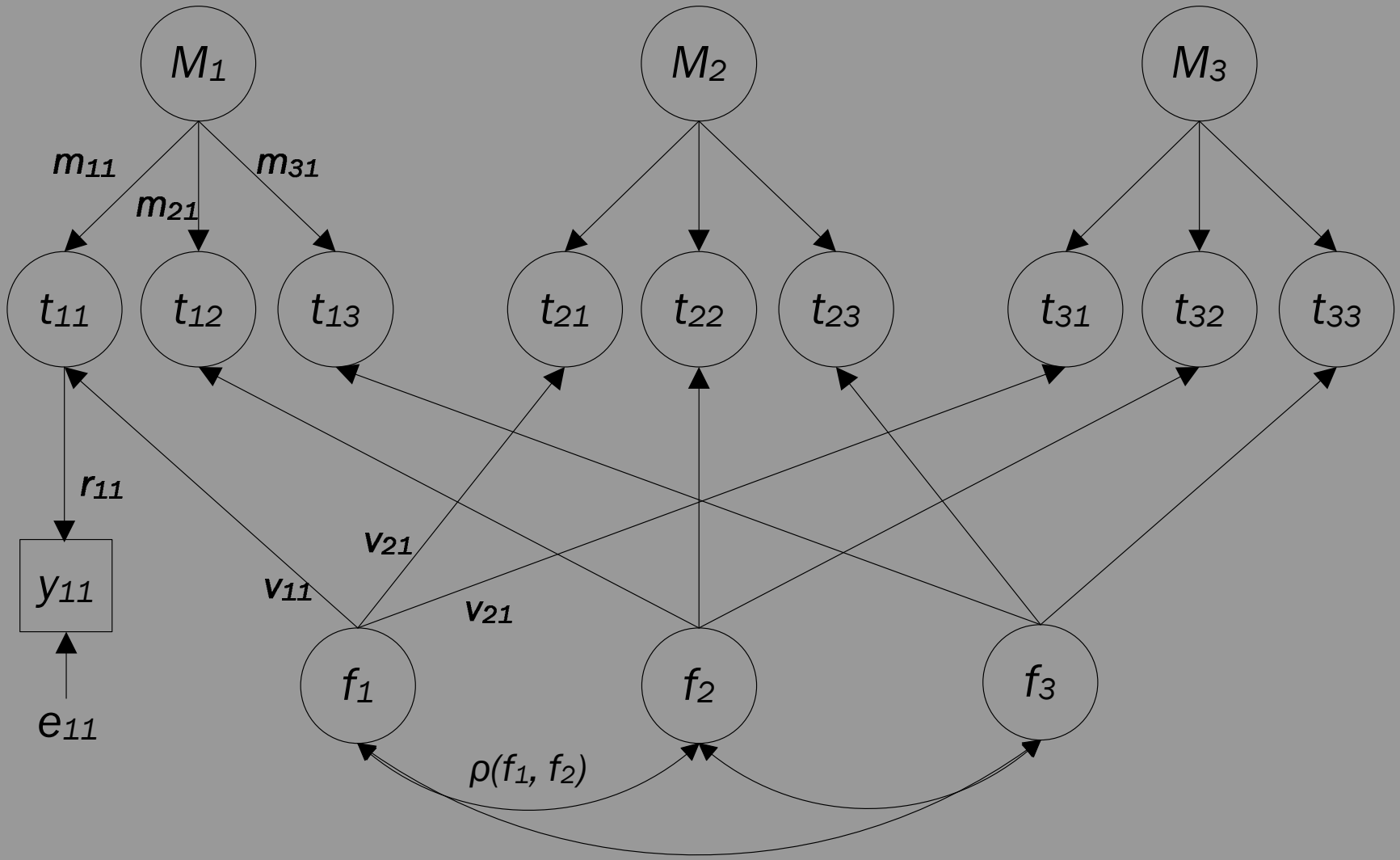
Different models for MTMM experiments

- We use the true score model
- Equivalent to the classic MTMM model
- Sometimes necessary to remove one method factor
- In that case our model is the equivalent to Eid's MTM-1 model.

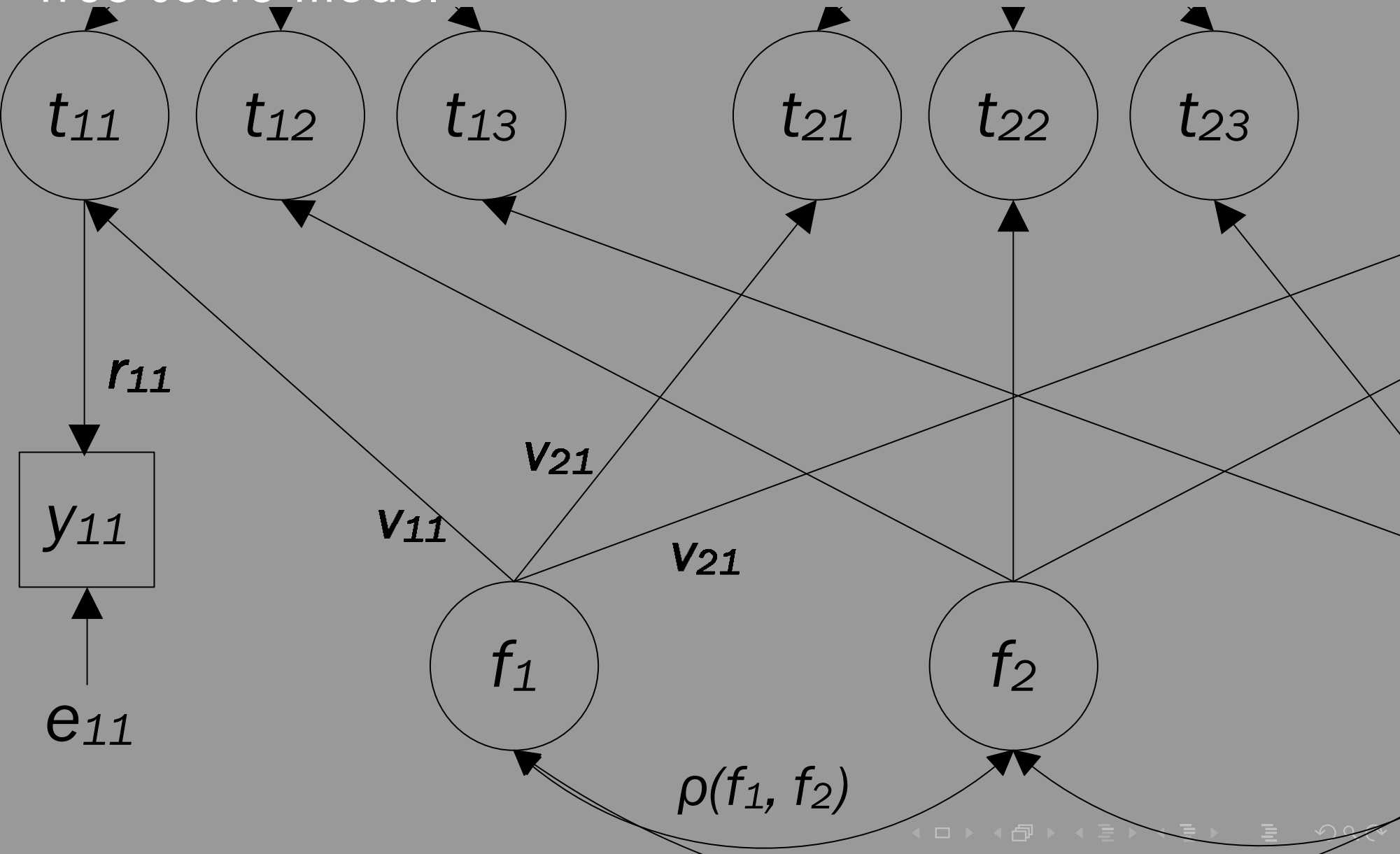
Different models for MTMM experiments

- We use the true score model
- Equivalent to the classic MTMM model
- Sometimes necessary to remove one method factor
- In that case our model is the equivalent to Eid's MTM-1 model.

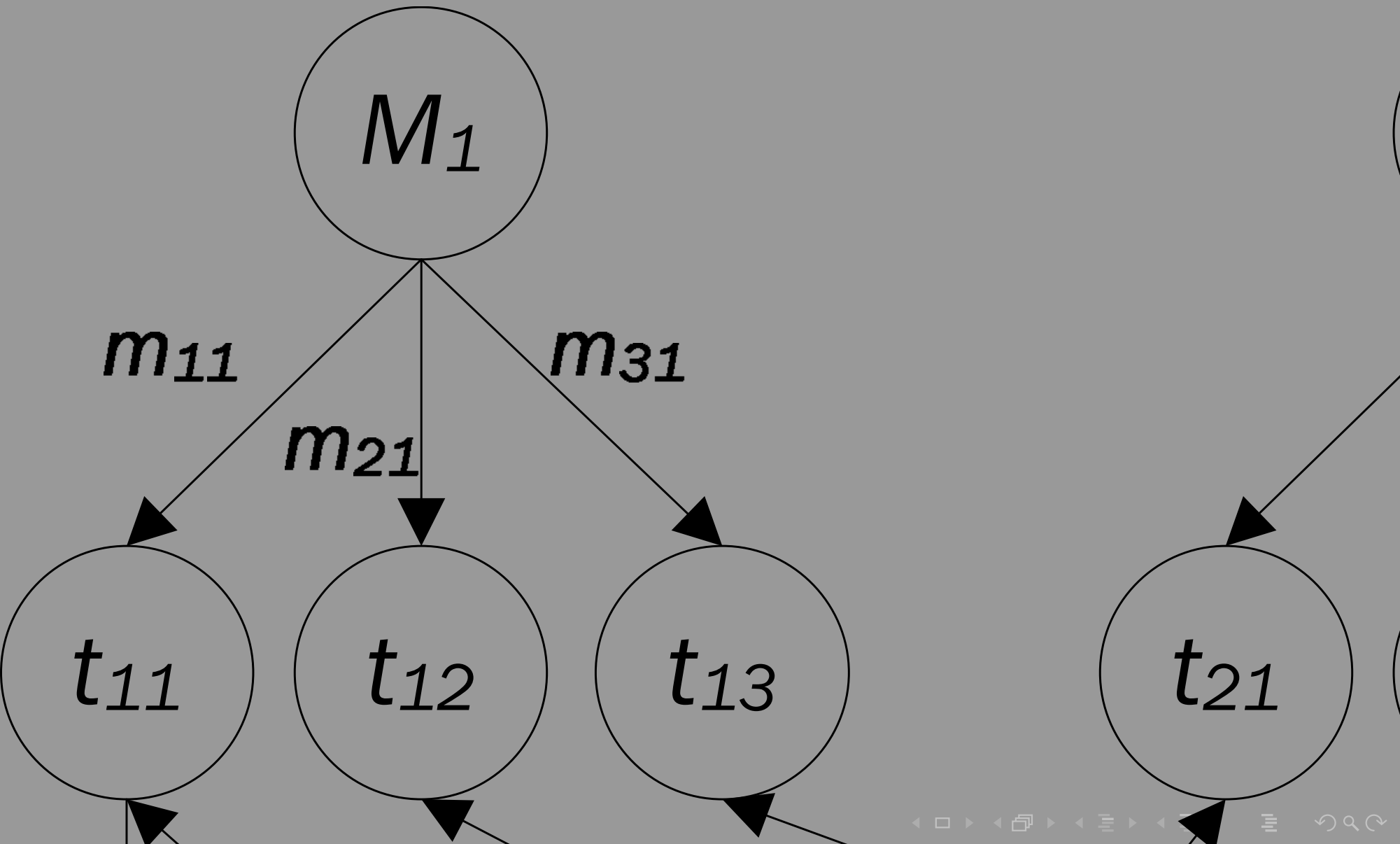
True score model



True score model



True score model



True score model assumptions

- No correlations among methods
- No correlations between traits and methods
- Equal method effects
- Linear and additive effects
- Normal errors, independent of all unobserved variables
- All variables are continuous

True score model assumptions

- No correlations among methods
- No correlations between traits and methods
- Equal method effects
- Linear and additive effects
- Normal errors, independent of all unobserved variables
- All variables are continuous

True score model assumptions

- No correlations among methods
- No correlations between traits and methods
- Equal method effects
- Linear and additive effects
- Normal errors, independent of all unobserved variables
- All variables are continuous

True score model assumptions

- No correlations among methods
- No correlations between traits and methods
- Equal method effects
- Linear and additive effects
- Normal errors, independent of all unobserved variables
- All variables are continuous

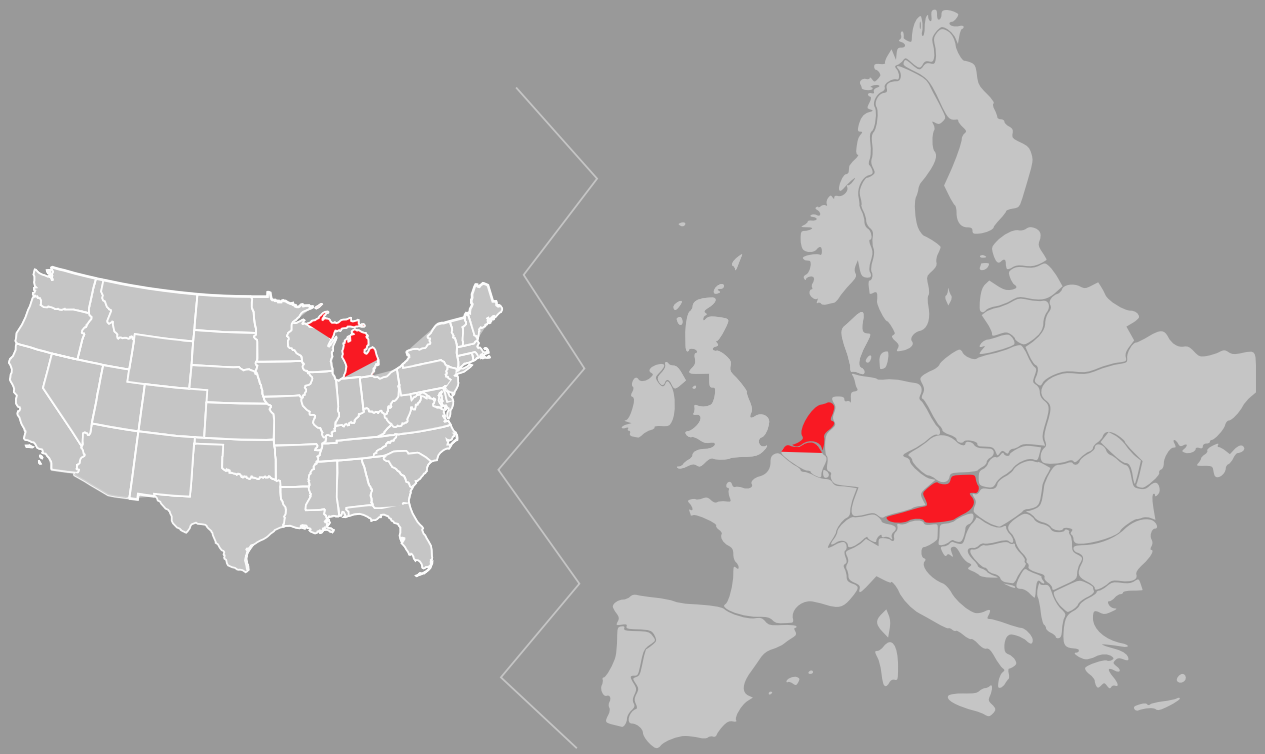
True score model assumptions

- No correlations among methods
- No correlations between traits and methods
- Equal method effects
- Linear and additive effects
- Normal errors, independent of all unobserved variables
- All variables are continuous

True score model assumptions

- No correlations among methods
- No correlations between traits and methods
- Equal method effects
- Linear and additive effects
- Normal errors, independent of all unobserved variables
- All variables are continuous

Countries in the international survey project 1984–1996 that have been included in SQP

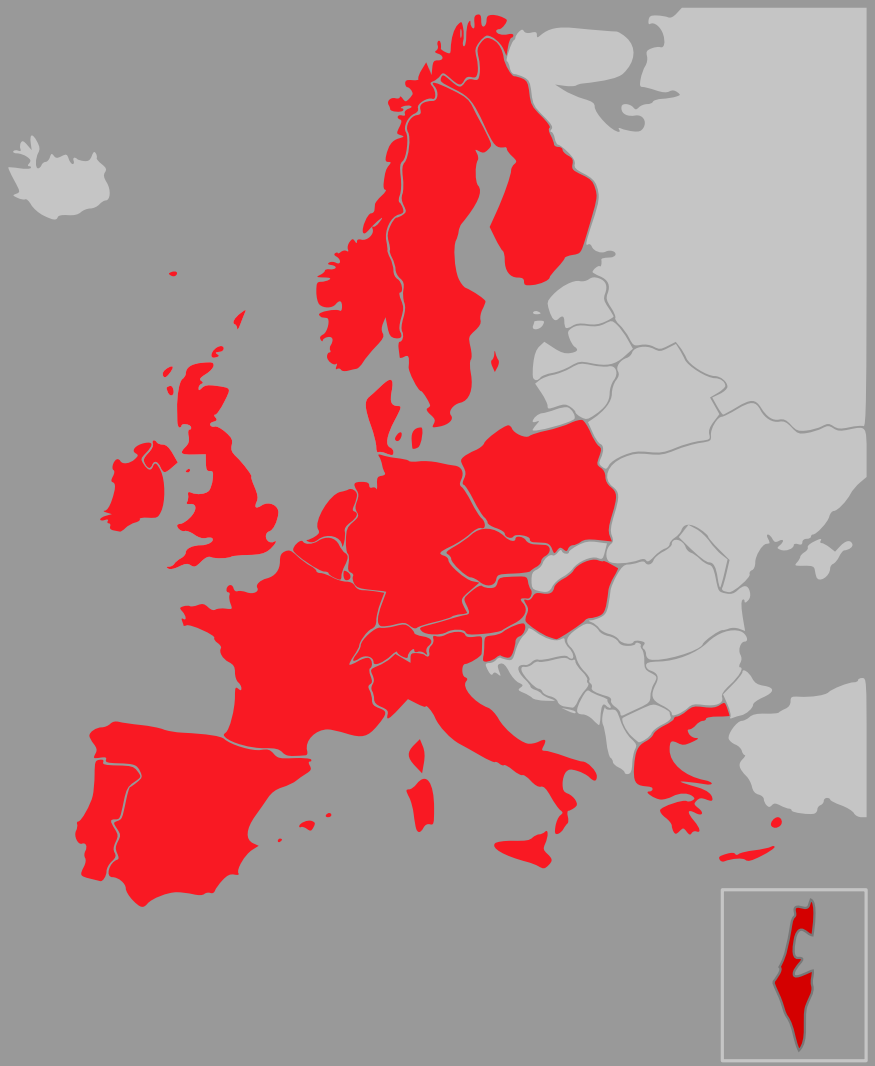


1. Austria
2. Belgium: Flanders
3. Netherlands
4. United States: Michigan

The European Social Survey (ESS)

- Three rounds, 4th coming up
- Six experiment in each round
- `http://www.europeansocialsurvey.org`

Countries in round 1 of the ESS – 2002



1 Austria
2 Belgium
3 Czech Republic
4 Denmark
5 Finland
6 France
7 Germany
8 Greece
9 Hungary
10 Ireland
11 Israel
12 Italy
13. Luxembourg
14. Netherlands
15. Norway
16. Poland
17. Portugal
18. Slovenia
19. Spain
20. Sweden
21. Switzerland
22. United Kingdom

Countries in round 2 of the ESS – 2004

1 Austria
2 Belgium
3 Czech Republic
4 Denmark
5 Estonia
6 Finland
7 France
8 Germany
9 Greece
10 Hungary
11 Iceland
12 Ireland
13 Italy
14. Luxembourg
15. Netherlands
16. Norway
17. Poland
18. Portugal
19. Slovakia
20. Slovenia
21. Spain
22. Sweden
23. Switzerland
24. Turkey
25. Ukraine
26. United Kingdom

Countries in round 3 of the ESS – 2006

1 Austria
2 Belgium
3 Bulgaria
4 Cyprus
5 Denmark
6 Estonia
7 Finland
8 France
9 Germany
10 Hungary
11 Ireland
12 Latvia
13. Netherlands
14. Norway
15. Poland
16. Portugal
17. Romania
18. Russian Federation
19. Slovakia
20. Slovenia
21. Spain
22. Sweden
23. Switzerland
24. Ukraine
25. United Kingdom

Some results from rounds 1 and 2

Country	Mean	Median	Minimum	Maximum
Portugal	0.79	0.81	0.63	0.91
Switzerland	0.79	0.84	0.56	0.90
Greece	0.78	0.79	0.64	0.90
Estonia	0.78	0.85	0.58	0.90
Poland	0.73	0.85	0.51	0.90
Luxembourg	0.72	0.73	0.53	0.88
United Kingdom	0.70	0.71	0.56	0.82
Denmark	0.70	0.70	0.52	0.80
Belgium	0.70	0.73	0.46	0.90
Germany	0.69	0.70	0.53	0.83
Spain	0.69	0.64	0.54	0.90
Austria	0.68	0.68	0.51	0.85
Czech Republic	0.65	0.60	0.52	0.87
Slovenia	0.63	0.60	0.46	0.82
Norway	0.59	0.59	0.35	0.83
Sweden	0.58	0.58	0.43	0.68
Finland	0.57	0.54	0.42	0.78

Differences between countries?

What we studied already:

- Differences in complexity of language?
- Artifacts due to sending in the questionnaire later?
- Artifacts due to mistakes in translation?

Differences between countries?

What we studied already:

- Differences in complexity of language?
  - Not found
- Artifacts due to sending in the questionnaire later?

- Artifacts due to mistakes in translation?

Differences between countries?

What we studied already:

- Differences in complexity of language?
  - Not found
- Artifacts due to sending in the questionnaire later?

- Artifacts due to mistakes in translation?

Differences between countries?

What we studied already:

- Differences in complexity of language?
  - Not found
- Artifacts due to sending in the questionnaire later?
  - Only Sweden, Norway, Finland
- Artifacts due to mistakes in translation?

Differences between countries?

What we studied already:

- Differences in complexity of language?
  - Not found
- Artifacts due to sending in the questionnaire later?
  - Only Sweden, Norway, Finland
- Artifacts due to mistakes in translation?

Differences between countries?

What we studied already:

- Differences in complexity of language?
  - Not found
- Artifacts due to sending in the questionnaire later?
  - Only Sweden, Norway, Finland
- Artifacts due to mistakes in translation?
  - Two cases found for two experiments

Differences between countries?

What we studied already:

- Differences in complexity of language?
  - Not found
- Artifacts due to sending in the questionnaire later?
  - Only Sweden, Norway, Finland
- Artifacts due to mistakes in translation?
  - Two cases found for two experiments
- None of these findings suffice to explain the large differences we found!

Differences between countries?

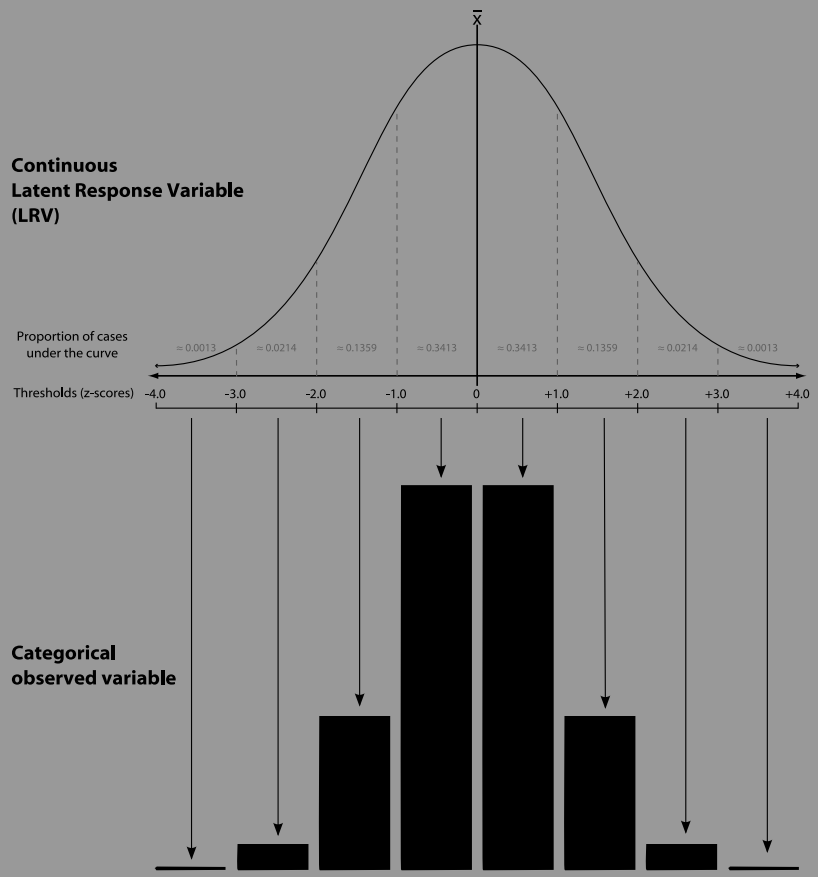
What we studied already:

- Differences in complexity of language?
  - Not found
- Artifacts due to sending in the questionnaire later?
  - Only Sweden, Norway, Finland
- Artifacts due to mistakes in translation?
  - Two cases found for two experiments
- None of these findings suffice to explain the large differences we found!
- Differences in use of the scale?

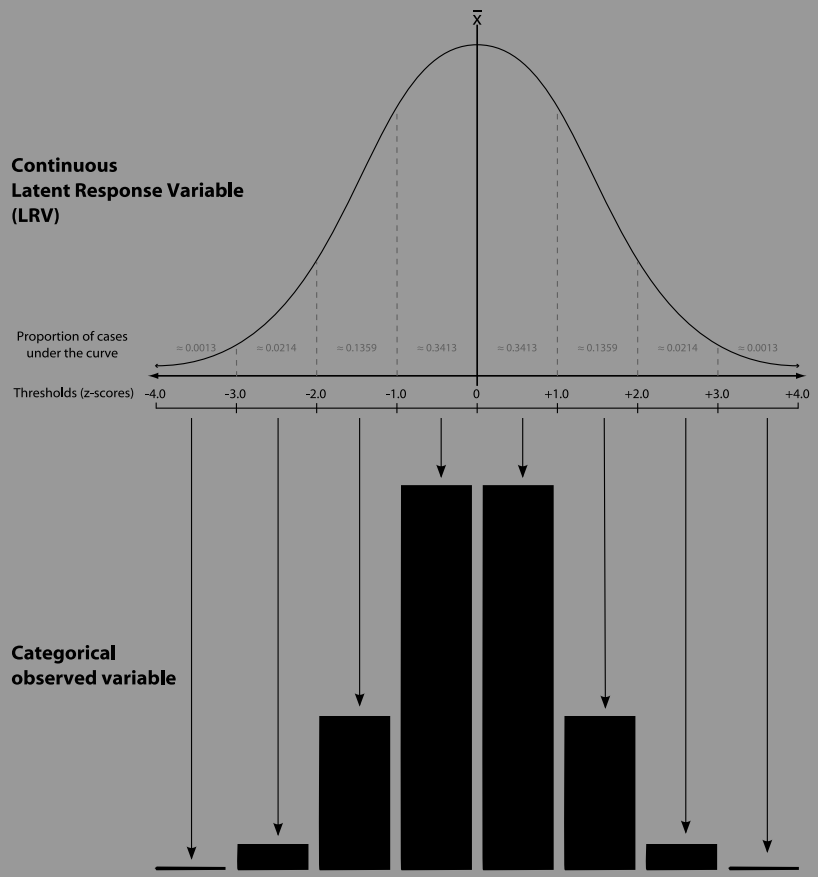
Categorisation of continuous variables

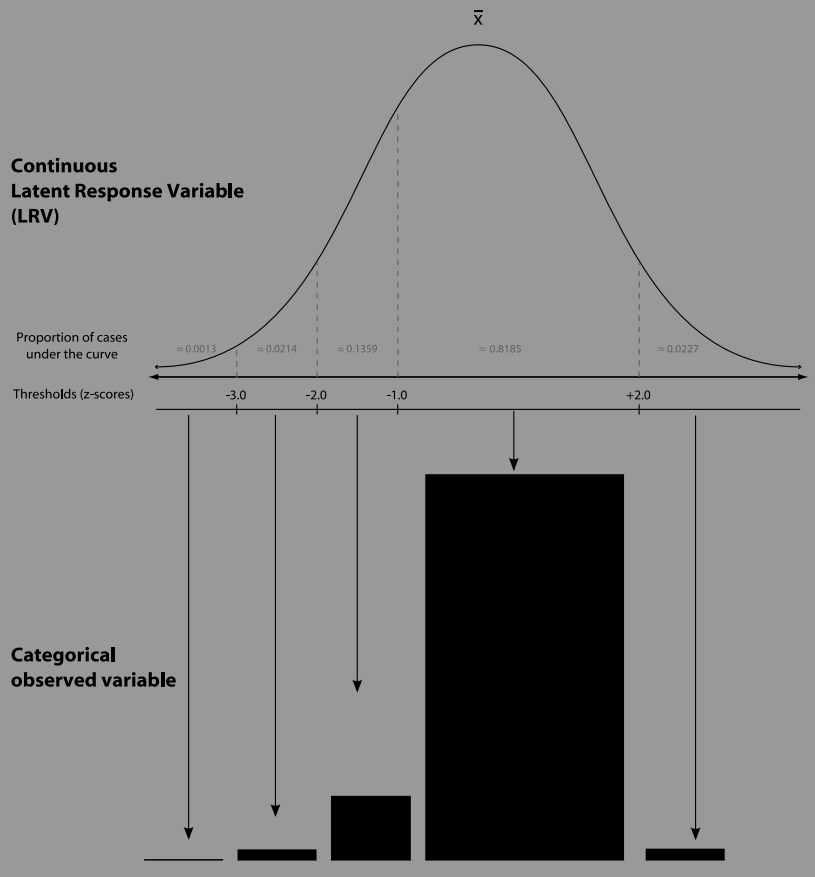
- Our model assumes that there are *unobserved* continuous latent response variables (LRV) that have been categorised into the *observed* categorical variables.
- These continuous latent response variables are related to each other according to the MTMM model.
- Equivalent to a 2 parameter graded response model in IRT (Muthén & Asparouhov).

Categorisation of continuous variables



Categorisation of continuous variables





Consequences of categorisation for the MTMM model

Efficacy experiment: Denmark

Polychoric correlations

		Method 1			Method 2		
Method 1	Complex	1.00					
	Active	-0.44	1.00				
	Mind	-0.51	0.47	1.00			
Method 2	Complex	0.66	-0.45	-0.51	1.00		
	Active	-0.44	0.74	0.46	-0.51	1.00	
	Mind	-0.52	0.51	0.67	-0.56	0.56	1.00

Pearson correlations

		Method 1			Method 2		
Method 1	Complex	1.00					
	Active	-0.40	1.00				
	Mind	-0.47	0.37	1.00			
Method 2	Complex	0.60	-0.37	-0.44	1.00		
	Active	-0.39	0.67	0.40	-0.43	1.00	
	Mind	-0.46	0.43	0.62	-0.49	0.48	1.00

% Increase in the correlations after correction for categorisation

Efficacy experiment: Denmark

		Method 1			Method 2	
Method 1	Complex					
	Active	**8**%				
	Mind	**8**%	**29**%			
Method 2	Complex	10%	22%	16%		
	Active	13%	10%	16%	**19**%	
	Mind	13%	19%	10%	**15**%	**16**%

Mean percentage increase of the polychoric correlations: 11%

Quality (q^2)and method effects (m) according to the continuous and categorical models, with categorisation factors

			'Efficacy'	
		Complex	Active	Mind
Continuous analysis				
q^2	Denmark	0.77	0.83	0.79
	Switzerland	0.49	0.81	0.50
m	Denmark	0.00	0.00	0.00
	Switzerland	0.00	0.00	0.00
Categorical analysis				
q^2	Denmark	0.63	0.70	0.63
	Switzerland	0.62	0.94	0.62
m	Denmark	0.11	0.08	0.11
	Switzerland	0.00	0.00	0.00
Categorisation factor				
	Denmark	1.23	1.18	1.25
	Switzerland	0.79	0.86	0.81

Consequences of correction for categorisation: conclusions

- The monomethod correlations for the second method increase more than those of the first method
- The method effects

Does categorisation explain differences across countries

- h

Consequences of categorisation for the correlations between observed variables

- The fewer categories, the smaller the Pearson correlation
- The more skew, the smaller the Pearson correlation

Therefore,

- b

Consequences of categorisation for the correlations between observed variables

- The fewer categories, the smaller the Pearson correlation
- The more skew, the smaller the Pearson correlation

Therefore,

- b

Consequences of categorisation for the correlations between observed variables

- The fewer categories, the smaller the Pearson correlation
- The more skew, the smaller the Pearson correlation

Therefore,

- b

The final goal: Survey Quality Predictor (SQP)

1. Estimate the model for all experiments
2. Save the reliability, validity, and method effect coefficients
3. Relate the coefficients to different aspects of the question
4. Predict the quality of survey questions from their characteristics (SQP)
5. Improve survey questions
6. `http://www.sqp.nl`

The final goal: Survey Quality Predictor (SQP)

1. Estimate the model for all experiments
2. Save the reliability, validity, and method effect coefficients
3. Relate the coefficients to different aspects of the question
4. Predict the quality of survey questions from their characteristics (SQP)
5. Improve survey questions
6. `http://www.sqp.nl`

The final goal: Survey Quality Predictor (SQP)

1. Estimate the model for all experiments
2. Save the reliability, validity, and method effect coefficients
3. Relate the coefficients to different aspects of the question
4. Predict the quality of survey questions from their characteristics (SQP)
5. Improve survey questions
6. `http://www.sqp.nl`

The final goal: Survey Quality Predictor (SQP)

1. Estimate the model for all experiments
2. Save the reliability, validity, and method effect coefficients
3. Relate the coefficients to different aspects of the question
   - Complexity of the sentence: no. words/sentence, avg. no. syllables, ...
   - Response scale: type, no. categories, ...
   - Formulation of the request: agree-disagree, extra information, ...
   - Data collection method: computer assisted, interviewer present, ...
4. Predict the quality of survey questions from their characteristics (SQP)
5. Improve survey questions
6. `http://www.sqp.nl`

The final goal: Survey Quality Predictor (SQP)

1. Estimate the model for all experiments
2. Save the reliability, validity, and method effect coefficients
3. Relate the coefficients to different aspects of the question
   - Complexity of the sentence: no. words/sentence, avg. no. syllables, ...
   - Response scale: type, no. categories, ...
   - Formulation of the request: agree-disagree, extra information, ...
   - Data collection method: computer assisted, interviewer present, ...
4. Predict the quality of survey questions from their characteristics (SQP)
5. Improve survey questions
6. `http://www.sqp.nl`

The final goal: Survey Quality Predictor (SQP)

1. Estimate the model for all experiments
2. Save the reliability, validity, and method effect coefficients
3. Relate the coefficients to different aspects of the question
   - Complexity of the sentence: no. words/sentence, avg. no. syllables, ...
   - Response scale: type, no. categories, ...
   - Formulation of the request: agree-disagree, extra information, ...
   - Data collection method: computer assisted, interviewer present, ...
4. Predict the quality of survey questions from their characteristics (SQP)
5. Improve survey questions
6. `http://www.sqp.nl`

The final goal: Survey Quality Predictor (SQP)

1. Estimate the model for all experiments
2. Save the reliability, validity, and method effect coefficients
3. Relate the coefficients to different aspects of the question
   - Complexity of the sentence: no. words/sentence, avg. no. syllables, ...
   - Response scale: type, no. categories, ...
   - Formulation of the request: agree-disagree, extra information, ...
   - Data collection method: computer assisted, interviewer present, ...
4. Predict the quality of survey questions from their characteristics (SQP)
5. Improve survey questions
6. `http://www.sqp.nl`

The final goal: Survey Quality Predictor (SQP)

1. Estimate the model for all experiments
2. Save the reliability, validity, and method effect coefficients
3. Relate the coefficients to different aspects of the question
   - Complexity of the sentence: no. words/sentence, avg. no. syllables, ...
   - Response scale: type, no. categories, ...
   - Formulation of the request: agree-disagree, extra information, ...
   - Data collection method: computer assisted, interviewer present, ...
4. Predict the quality of survey questions from their characteristics (SQP)
5. Improve survey questions
6. `http://www.sqp.nl`

The final goal: Survey Quality Predictor (SQP)

1. Estimate the model for all experiments
2. Save the reliability, validity, and method effect coefficients
3. Relate the coefficients to different aspects of the question
   - Complexity of the sentence: no. words/sentence, avg. no. syllables, ...
   - Response scale: type, no. categories, ...
   - Formulation of the request: agree-disagree, extra information, ...
   - Data collection method: computer assisted, interviewer present, ...
4. Predict the quality of survey questions from their characteristics (SQP)
5. Improve survey questions
6. `http://www.sqp.nl`

The final goal: Survey Quality Predictor (SQP)

1. Estimate the model for all experiments
2. Save the reliability, validity, and method effect coefficients
3. Relate the coefficients to different aspects of the question
   - Complexity of the sentence: no. words/sentence, avg. no. syllables, ...
   - Response scale: type, no. categories, ...
   - Formulation of the request: agree-disagree, extra information, ...
   - Data collection method: computer assisted, interviewer present, ...
4. Predict the quality of survey questions from their characteristics (SQP)
5. Improve survey questions
6. http://www.sqp.nl