

# Survey data quality in different countries

Daniel Oberski

Faculty of Social and Behavioural Sciences  
Tilburg University

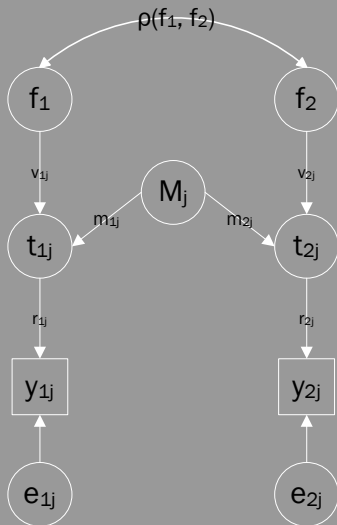
Survey Research Centre  
ESADE, Barcelona  
Universitat Ramon Llull



# 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?
  - Categorisation errors in the efficacy experiment
  - Meta analysis of the 4 different experiments
- 4 Conclusion

# 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  $y_{ij}$

# The basic response model

- The quality coefficient  $q$  is the product of the reliability and validity coefficients:
- $q = vr$
- The square  $q^2$  is called the 'total quality' of a measure.
- It is the percentage of variance in the observed variable that can be explained by the latent variable of interest.

# The basic response model

- The quality coefficient  $q$  is the product of the reliability and validity coefficients:
- $q = vr$
- The square  $q^2$  is called the 'total quality' of a measure.
- It is the percentage of variance in the observed variable that can be explained by the latent variable of interest.

# The basic response model

- The quality coefficient  $q$  is the product of the reliability and validity coefficients:
- $q = vr$
- The square  $q^2$  is called the 'total quality' of a measure.
- It is the percentage of variance in the observed variable that can be explained by the latent variable of interest.

# The basic response model

- The quality coefficient  $q$  is the product of the reliability and validity coefficients:
- $q = vr$
- The square  $q^2$  is called the 'total quality' of a measure.
- It is the percentage of variance in the observed variable that can be explained by the latent variable of interest.

# 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<br>all true | A little<br>true | Quite true | Very true | (Don't<br>know) |
|--|--------------------|------------------|------------|-----------|-----------------|
| <b>G64</b> 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.**

Not at all varied ☐ 1

A little varied ☐ 2

Quite varied ☐ 3

Very varied ☐ 4

**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                       |
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |



# 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<br>all true | A little<br>true | Quite true | Very true | (Don't<br>know) |
|------------|--|--------------------|------------------|------------|-----------|-----------------|
| <b>G64</b> | There is a lot of variety in my work.              | 1                  | 2                | 3          | 4         | 8               |
| ...        |  |                    |                  |            |           |                 |
| <b>G66</b> | My job is secure                                   | 1                  | 2                | 3          | 4         | 8               |
| ...        |  |                    |                  |            |           |                 |
| <b>G70</b> | 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.**

Not at all varied ☐ 1

A little varied ☐ 2

Quite varied ☐ 3

Very varied ☐ 4

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

**Please tick one box.**

Not at all secure ☐ 1

A little secure ☐ 2

Quite secure ☐ 3

Very secure ☐ 4

**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.**

Not at all at risk ☐ 1

A little at risk ☐ 2

Quite a lot at risk ☐ 3

Very much at risk ☐ 4

# 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                       |
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |

**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                       |
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |

**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                       |
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |

# 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.

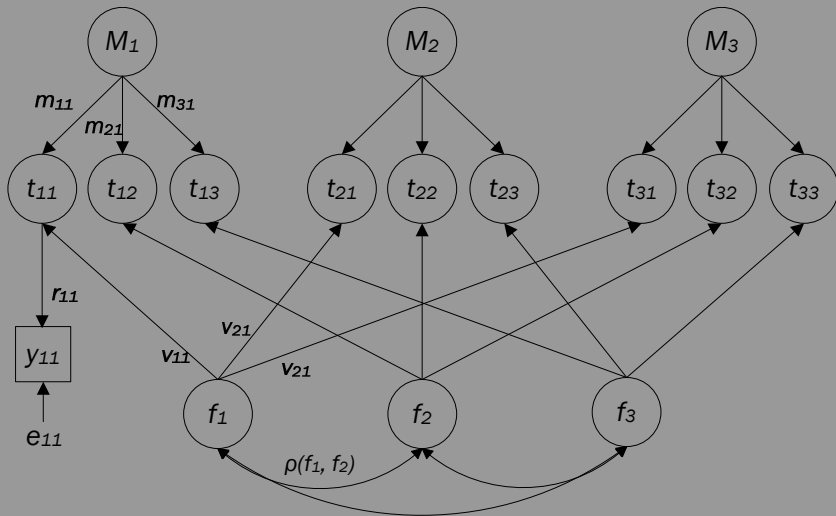
# Different models for MTMM experiments

- 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.

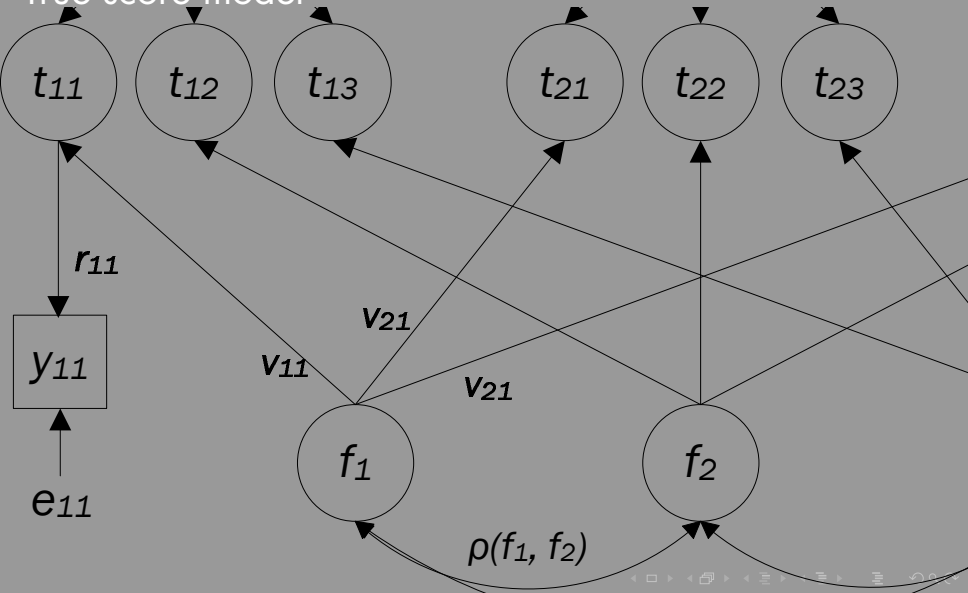
# Different models for MTMM experiments

- 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.

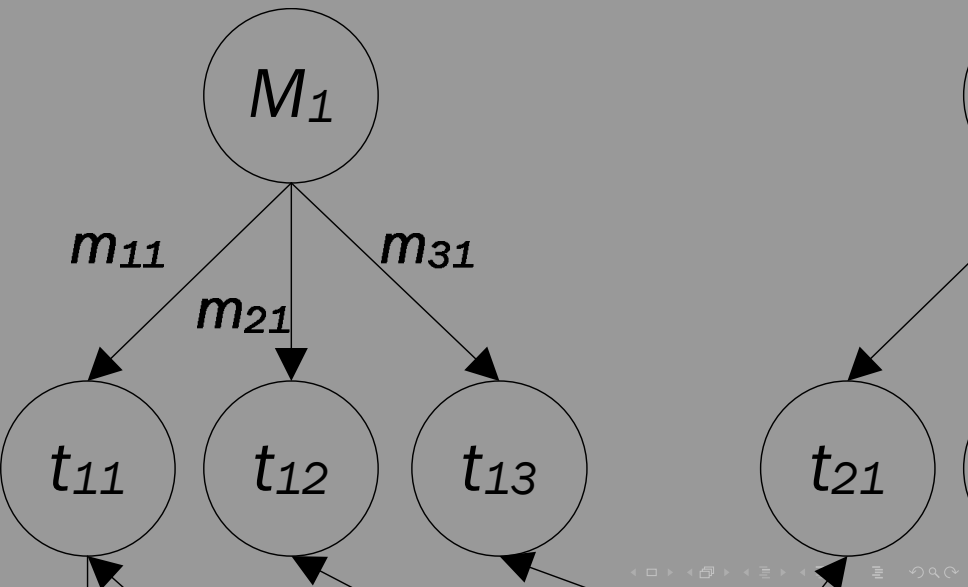
# True score model



## 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

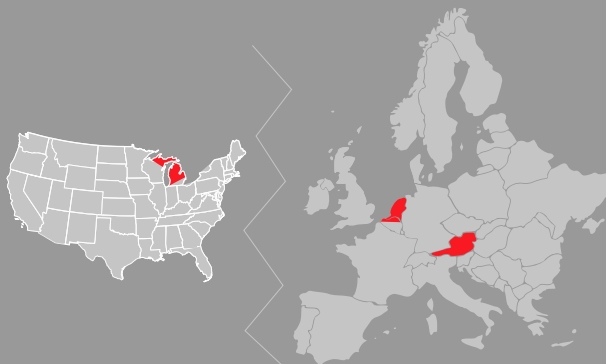
# True score model assumptions

- 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



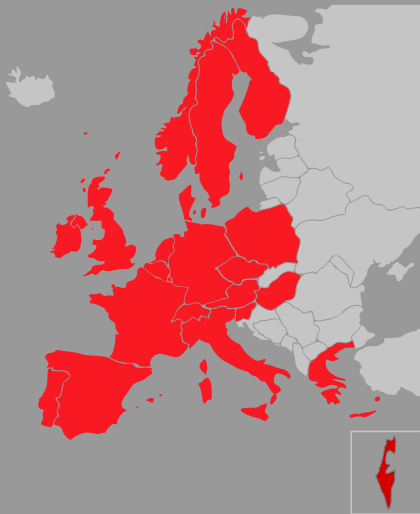
- 1 Austria
- 2 Belgium:  
Flanders
- 3 Netherlands
- 4 United States:  
Michigan

# The European Social Survey (ESS)



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

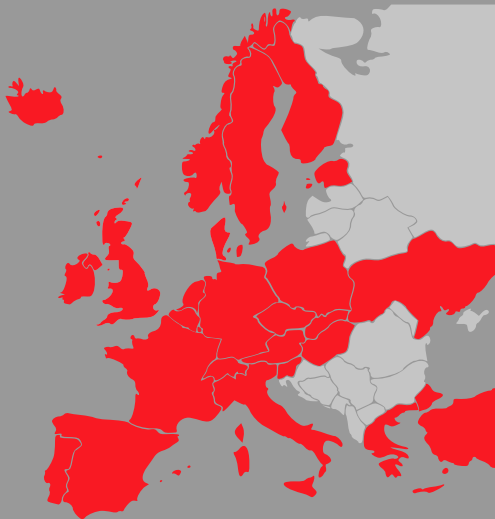
## Countries in round 1 of the ESS – 2002



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



## Countries in round 2 of the ESS – 2004



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

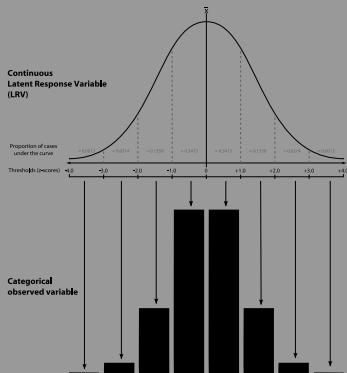
- Differences in complexity of language?
  - Not found
- Artifacts due to sending in the questionnaire later?
  - Only Sweden, Norway, Finland
- Artifacts due to mistakes in translation?
  - Two cases found for two experiments
- None of these findings suffice to explain the large differences we found!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

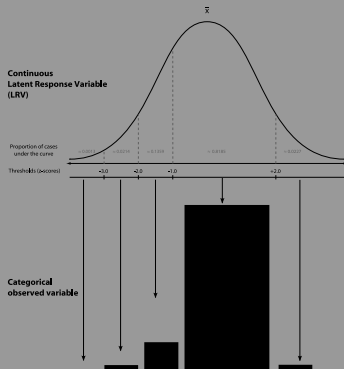
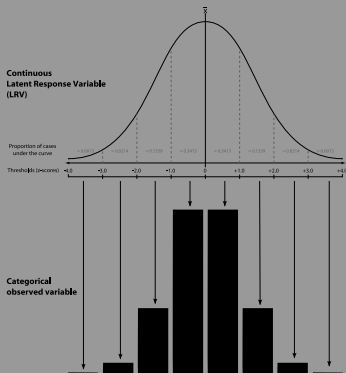
# Categorisation of continuous variables

Our model assumes that there are *unobserved* continuous latent response variables (LRV) that have been categorised into the *observed* categorical variables.



# 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.
- Method effects, quality coefficients, and thresholds can be estimated.
- Equivalent to a 2 parameter graded response model in IRT (Muthén & Asparouhov 2002).
- We call the ratio of the quality coefficient  $q = v.r$  from the categorical model to the same coefficient from the continuous model the 'categorisation factor'.

# Consequences of categorisation for the correlations between observed variables

- The fewer categories, the smaller the Pearson correlation
- The more skew in observed variables, the smaller the Pearson correlation
- The corrected ('polychoric') correlations are always higher than the Pearson correlations, but not necessarily equally so for all variables.



# Consequences of categorisation for the correlations between observed variables

- The fewer categories, the smaller the Pearson correlation
- The more skew in observed variables, the smaller the Pearson correlation
- The corrected ('polychoric') correlations are always higher than the Pearson correlations, but not necessarily equally so for all variables.

# Consequences of categorisation for the correlations between observed variables

- The fewer categories, the smaller the Pearson correlation
- The more skew in observed variables, the smaller the Pearson correlation
- The corrected ('polychoric') correlations are always higher than the Pearson correlations, but not necessarily equally so for all variables.

Therefore,

- If the skewness of observed variables is higher for variables measured by one particular method, then the corrected correlations between those variables will go up more than the others, and the method effects in the categorical model will be higher.
- Higher estimated method effects can lower the estimate of quality in the categorical model.
- If this differs across countries, the process can possibly explain differences in the quality.

Therefore,

- If the skewness of observed variables is higher for variables measured by one particular method, then the corrected correlations between those variables will go up more than the others, and the method effects in the categorical model will be higher.
- Higher estimated method effects can lower the estimate of quality in the categorical model.
- If this differs across countries, the process can possibly explain differences in the quality.

Therefore,

- If the skewness of observed variables is higher for variables measured by one particular method, then the corrected correlations between those variables will go up more than the others, and the method effects in the categorical model will be higher.
- Higher estimated method effects can lower the estimate of quality in the categorical model.
- If this differs across countries, the process can possibly explain differences in the quality.

# Analysis of the experiments

- We analysed the 4 experiments from the ESS which involved variables with 5 categories or less
- The topics: role of women, GP's, political efficacy, job.



- Compare the country with the highest quality to the country with the lowest quality for that experiment

# Analysis of the experiments

- We analysed the 4 experiments from the ESS which involved variables with 5 categories or less
- The topics: role of women, GP's, political efficacy, job.



- Compare the country with the highest quality to the country with the lowest quality for that experiment

# Analysis of the experiments

- We analysed the 4 experiments from the ESS which involved variables with 5 categories or less
- The topics: role of women, GP's, political efficacy, job.



- Compare the country with the highest quality to the country with the lowest quality for that experiment



# Quality ( $q^2$ ) and method effects ( $m$ ) in the efficacy experiment

## Continuous MTMM model, main questionnaire (first method)

|       |             | 'Efficacy' |        |      |
|-------|-------------|------------|--------|------|
|       |             | Complex    | Active | Mind |
| $q^2$ | 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 |

# 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 |

# Efficacy experiment: Switzerland

## Polychoric correlations

|          |         | Method 1 |       |       | Method 2 |      |      |
|----------|---------|----------|-------|-------|----------|------|------|
| Method 1 | Complex | 1.00     |       |       |          |      |      |
|          | Active  | -0.37    | 1.00  |       |          |      |      |
|          | Mind    | -0.46    | 0.42  | 1.00  |          |      |      |
| Method 2 | Complex | 0.57     | -0.36 | -0.46 | 1.00     |      |      |
|          | Active  | -0.32    | 0.83  | 0.36  | -0.39    | 1.00 |      |
|          | Mind    | -0.36    | 0.44  | 0.69  | -0.49    | 0.43 | 1.00 |

## Pearson correlations

|          |         | Method 1 |       |       | Method 2 |      |      |
|----------|---------|----------|-------|-------|----------|------|------|
| Method 1 | Complex | 1.00     |       |       |          |      |      |
|          | Active  | -0.33    | 1.00  |       |          |      |      |
|          | Mind    | -0.34    | 0.36  | 1.00  |          |      |      |
| Method 2 | Complex | 0.55     | -0.35 | -0.45 | 1.00     |      |      |
|          | Active  | -0.30    | 0.82  | 0.33  | -0.34    | 1.00 |      |
|          | Mind    | -0.35    | 0.41  | 0.62  | -0.48    | 0.39 | 1.00 |

# % Increase in the correlations after correction for categorisation

## Efficacy experiment: Denmark

|  |         | Method 1  |            |     | Method 2   |            |
|--|---------|-----------|------------|-----|------------|------------|
| Method 1   | Complex |           |            |     |            |            |
|  | Active  | <b>8%</b> |            |     |            |            |
|  | Mind    | <b>8%</b> | <b>29%</b> |     |            |            |
| Method 2   | Complex | 10%       | 22%        | 16% |            |            |
|  | Active  | 13%       | 10%        | 16% | <b>19%</b> |            |
|  | Mind    | 13%       | 19%        | 10% | <b>15%</b> | <b>16%</b> |
| Mean percentage increase of the polychoric correlations: 11% |         |           |            |     |            |            |

# % Increase in the correlations after correction for categorisation

## Efficacy experiment: Switzerland

|   |         | Method 1   |            |     | Method 2   |            |
|---|---------|------------|------------|-----|------------|------------|
| Method 1  | Complex |            |            |     |            |            |
|   | Active  | <b>1%</b>  |            |     |            |            |
|   | Mind    | <b>36%</b> | <b>17%</b> |     |            |            |
| Method 2  | Complex | 3%         | 3%         | 2%  |            |            |
|   | Active  | 5%         | 1%         | 9%  | <b>17%</b> |            |
|   | Mind    | 1%         | 6%         | 13% | <b>3%</b>  | <b>11%</b> |
| Mean percentage increase of the polychoric correlations: 6.5% |         |            |            |     |            |            |

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

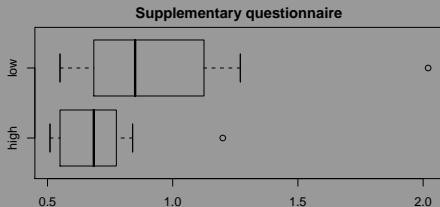
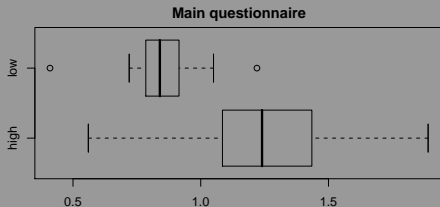
|                       |             | 'Efficacy' |        |      |
|-----------------------|-------------|------------|--------|------|
|                       |             | Complex    | Active | Mind |
| Continuous analysis   |             |            |        |      |
| $q^2$                 | 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

The categorisation factor,  $q_{cat}/q_{cont}$ :





# A meta-analysis of the categorisation error studies

|                               | Estimate             | S.E. | 95% C.I. |       |
|-------------------------------|----------------------|------|----------|-------|
|                               |                      |      | lower    | upper |
| (Intercept)                   | 1.04                 | 0.36 | 0.31     | 1.77  |
| <i>Topic</i>                  |                      |      |          |       |
| Doctors                       | (reference category) |      |          |       |
| Efficacy                      | 0.06                 | 0.10 | -0.14    | 0.27  |
| Job                           | 0.04                 | 0.40 | -0.71    | 0.78  |
| Women                         | 0.38                 | 0.26 | -0.14    | 0.90  |
| <i>Scale</i>                  |                      |      |          |       |
| Direct                        | (reference category) |      |          |       |
| Agree-disagree                | -0.11                | 0.35 | -0.81    | 0.59  |
| True-false                    | 0.17                 | 0.32 | -0.48    | 0.81  |
| Negative                      | -0.50                | 0.23 | -0.96    | -0.02 |
| Main questionnaire            | -0.30                | 0.29 | -0.88    | 0.29  |
| Highest quality               | -0.19                | 0.09 | -0.37    | -0.01 |
| Highest quality $\times$ main | 0.66                 | 0.15 | 0.35     | 0.96  |

Multiple R-Squared: 0.45; Adjusted R-squared: 0.35

# Conclusions

- Differences in data quality between countries are reduced after the categorisation is taken into account.
- Differences in use of the scale seems to play a large role in causing differences between countries.

# Conclusions

- Differences in data quality between countries are reduced after the categorisation is taken into account.
- Differences in use of the scale seems to play a large role in causing differences between countries.

# Recommendations

- One way to prevent categorisation errors is to use continuous or near-continuous scales;
- Don't use categories which are difficult to choose in some countries but not in others (e.g. disagree that ŠMen should take as much responsibility as women for the home and children.Š in Slovenia vs. Greece.)
- High-quality questions may not be as affected by categorisation errors as low-quality ones!

# Recommendations

- One way to prevent categorisation errors is to use continuous or near-continuous scales;
- Don't use categories which are difficult to choose in some countries but not in others (e.g. disagree that 'Men should take as much responsibility as women for the home and children.' in Slovenia vs. Greece.)
- High-quality questions may not be as affected by categorisation errors as low-quality ones!

# Recommendations

- One way to prevent categorisation errors is to use continuous or near-continuous scales;
- Don't use categories which are difficult to choose in some countries but not in others (e.g. disagree that ŠMen should take as much responsibility as women for the home and children.Š in Slovenia vs. Greece.)
- High-quality questions may not be as affected by categorisation errors as low-quality ones!

## Further study, problems

- Investigate normality assumption, linearity;
- Unobserved heterogeneity;
- Prediction of the data quality based on characteristics of the question.

## Further study, problems

- Investigate normality assumption, linearity;
- **Unobserved heterogeneity;**
- Prediction of the data quality based on characteristics of the question.



## Further study, problems

- Investigate normality assumption, linearity;
- Unobserved heterogeneity;
- Prediction of the data quality based on characteristics of the question.

That's it for now. Moltes gràcies per a la seva atenció!

