Detection of misspecifications in Mplus

Daniel Oberski

Research and Expertise Center for Survey Methodology
Universitat Pompeu Fabra, Spain

Mini conference Measurement invariance: Methods, problems and further directions
Zürich, 16 July 2011
<table>
<thead>
<tr>
<th><strong>INTRODUCTION</strong></th>
<th><strong>EXAMPLE 1</strong></th>
<th><strong>THE PROBLEM AND A SOLUTION</strong></th>
<th><strong>EXAMPLE 2</strong></th>
<th><strong>CONCLUSIONS</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>INTRODUCTION</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>EXAMPLE 1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>THE PROBLEM AND A SOLUTION</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>EXAMPLE 2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>CONCLUSIONS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Example 1: Correlations between genetic polymorphisms**

- Rioux et al., Nature Genetics 29, 223 - 228 (2001)
**Example 1: Correlations between genetic polymorphisms**

- Rioux et al., Nature Genetics 29, 223 - 228 (2001)
- Crohn disease is an inflammatory disease of the intestines. Causes abdominal pain, diarrhea, and vomiting. The body’s immune system attacks the gastrointestinal tract, causing inflammation. 🍗
Example 1: Correlations between Genetic Polymorphisms

- Rioux et al., Nature Genetics 29, 223 - 228 (2001)
- Crohn disease is an inflammatory disease of the intestines. Causes abdominal pain, diarrhea, and vomiting. The body’s immune system attacks the gastrointestinal tract, causing inflammation.
- Gene “IBD5” is thought to increase risk of Crohn disease
Example 1: Correlations Between Genetic Polymorphisms

- Rioux et al., Nature Genetics 29, 223 - 228 (2001)
- Crohn disease is an inflammatory disease of the intestines. Causes abdominal pain, diarrhea, and vomiting. The body’s immune system attacks the gastrointestinal tract, causing inflammation.
- Gene “IBD5” is thought to increase risk of Crohn disease
- Unfortunately gene “IBD5” is difficult (expensive) to detect
Example 1: Correlations between genetic polymorphisms

- Rioux et al., Nature Genetics 29, 223 - 228 (2001)
- Crohn disease is an inflammatory disease of the intestines. Causes abdominal pain, diarrhea, and vomiting. The body’s immune system attacks the gastrointestinal tract, causing inflammation.
- Gene “IBD5” is thought to increase risk of Crohn disease
- Unfortunately gene “IBD5” is difficult (expensive) to detect
- Fortunately, however, there are lots of mutations that can easily be detected, which just happen to be strongly related to the occurrence of IBD5.
Example 1: Correlations between genetic polymorphisms

- Rioux et al., Nature Genetics 29, 223 - 228 (2001)
- Crohn disease is an inflammatory disease of the intestines. Causes abdominal pain, diarrhea, and vomiting. The body’s immune system attacks the gastrointestinal tract, causing inflammation.
- Gene “IBD5” is thought to increase risk of Crohn disease
- Unfortunately gene “IBD5” is difficult (expensive) to detect
- Fortunately, however, there are lots of mutations that can easily be detected, which just happen to be strongly related to the occurrence of IBD5.
- This is called “linkage disequilibrium” (don’t ask me why)
Example 1: Correlations between genetic polymorphisms

- Rioux et al., Nature Genetics 29, 223 - 228 (2001)
- Crohn disease is an inflammatory disease of the intestines. Causes abdominal pain, diarrhea, and vomiting. The body’s immune system attacks the gastrointestinal tract, causing inflammation.
- Gene “IBD5” is thought to increase risk of Crohn disease
- Unfortunately gene “IBD5” is difficult (expensive) to detect
- Fortunately, however, there are lots of mutations that can easily be detected, which just happen to be strongly related to the occurrence of IBD5.
- This is called “linkage disequilibrium” (don’t ask me why)
- Can we use those other mutations (“SNP’s”) as indicators of the presence of IBD5?
Example 1: Correlations between genetic polymorphisms

- Can we use those other mutations ("SNP’s") as indicators of the presence of IBD5?
<table>
<thead>
<tr>
<th>Introduction</th>
<th>Example 1</th>
<th>The Problem and a Solution</th>
<th>Example 2</th>
<th>Conclusions</th>
</tr>
</thead>
</table>

**Alternative Models**
ALTERNATIVE MODELS
ALTERNATIVE MODELS
The restricted model estimated with Mplus
THE RESTRICTED MODEL ESTIMATED WITH Mplus

![Diagram of the restricted model estimated with Mplus]
THE RESTRICTED MODEL ESTIMATED WITH Mplus

\[
\begin{align*}
\text{IRG1143} & \quad 0.96 \\
\text{IRG1144} & \quad 0.98 \\
\text{IRG1218} & \quad 0.98 \\
\text{IRG1219} & \quad 0.93 \\
\end{align*}
\]
THE RESTRICTED MODEL ESTIMATED WITH Mplus

- Chisquare: 32.4, df: 2, p: 0.0000
THE RESTRICTED MODEL ESTIMATED WITH Mplus

- Chisquare: 32.4, df: 2, p: 0.0000
- CFI: 0.912
THE RESTRICTED MODEL ESTIMATED WITH Mplus

- Chisquare: 32.4, df: 2, p: 0.0000
- CFI: 0.912
- TLI: 0.735
THE RESTRICTED MODEL ESTIMATED WITH MPLUS

- Chisquare: 32.4, df: 2, p: 0.0000
- CFI: 0.912
- TLI: 0.735
- RMSEA: 0.616 (0.441, 0.812)
THE RESTRICTED MODEL ESTIMATED WITH MPlus

- Chisquare: 32.4, \( df: 2, p: 0.0000 \)
- CFI: 0.912
- TLI: 0.735
- RMSEA: 0.616 (0.441, 0.812)
- SRMR: 0.010
The restricted model estimated with Mplus

- Chisquare: 32.4, df: 2, p: 0.0000
- CFI: 0.912
- TLI: 0.735
- RMSEA: 0.616 (0.441, 0.812)
- SRMR: 0.010

Modification index (MI) of restricted-to-1-correlation:
THE RESTRICTED MODEL ESTIMATED WITH Mplus

- Chisquare: 32.4, df: 2, p: 0.0000
- CFI: 0.912
- TLI: 0.735
- RMSEA: 0.616 (0.441, 0.812)
- SRMR: 0.010

- Modification index (MI) of restricted-to-1-correlation:
  MI: 4.589 (p = 0.032)
The restricted model estimated with Mplus

- Chisquare: 32.4, $df$: 2, $p$: 0.0000
- CFI: 0.912
- TLI: 0.735
- RMSEA: 0.616 (0.441, 0.812)
- SRMR: 0.010
- Modification index (MI) of restricted-to-1-correlation:
  MI: 4.589 ($p = 0.032$)
  EPC: -0.014
THE RESTRICTED MODEL ESTIMATED WITH Mplus

- Chisquare: 32.4, df: 2, p: 0.0000
- CFI: 0.912
- TLI: 0.735
- RMSEA: 0.616 (0.441, 0.812)
- SRMR: 0.010

- Modification index (MI) of restricted-to-1-correlation:
  MI: 4.589 (p = 0.032)
  EPC: -0.014
  (this would imply a correlation of 1 - 0.014 = 0.986)
THE UNRESTRICTED MODEL ESTIMATED WITH Mplus
THE UNRESTRICTED MODEL ESTIMATED WITH Mplus
THE UNRESTRICTED MODEL ESTIMATED WITH Mplus

The model estimated with Mplus includes the following parameters:

- IRG1143: 0.96
- IRG1144: 0.99
- IRG1218: 0.98
- IRG1219: 0.99

Chi-square: 27.4, df: 1, p: 0.0000
ΔChi-square: 5.0, df: 1, p: 0.025
CFI: 0.923
TLI: 0.540
RMSEA: 0.812 (0.567, 1.087)
SRMR: 0.012
THE UNRESTRICTED MODEL ESTIMATED WITH Mplus

► Chisquare: 27.4, df: 1, p: 0.0000

![Diagram of correlations between variables IRG1143, IRG1144, IRG1218, and IRG1219, with correlation coefficients 0.96, 0.99, 0.98, and 0.99, respectively.]
The unrestricted model estimated with **Mplus**

- Chisquare: 27.4, \(df\): 1, \(p\): 0.0000
- \(\Delta\) Chisquare: 5.0, \(df\): 1, \(p\): 0.025
The unrestricted model estimated with Mplus

- Chisquare: 27.4, df: 1, p: 0.0000
- Δ Chisquare: 5.0, df: 1, p: 0.025
- CFI: 0.923
The unrestricted model estimated with Mplus

- Chisquare: 27.4, df: 1, p: 0.0000
- Δ Chisquare: 5.0, df: 1, p: 0.025
- CFI: 0.923
- TLI: 0.540
The unrestricted model estimated with Mplus

- Chisquare: 27.4, df: 1, p: 0.0000
- Δ Chisquare: 5.0, df: 1, p: 0.025
- CFI: 0.923
- TLI: 0.540
- RMSEA: 0.812 (0.567, 1.087)
THE UNRESTRICTED MODEL ESTIMATED WITH Mplus

- Chisquare: 27.4, df: 1, p: 0.0000
- Δ Chisquare: 5.0, df: 1, p: 0.025
- CFI: 0.923
- TLI: 0.540
- RMSEA: 0.812 (0.567, 1.087)
- SRMR: 0.012
Can we consider all four variables as indicators of the same thing?

▶ What shall we decide?

The correlation is significantly different from 1. When estimated it is 0.98. But according to all criteria, the model should be rejected. This is not what we wanted! Whether we used $\chi^2$, $\Delta \chi^2$, or any of the fit measures, we would make a wrong decision.
Can we consider all four variables as indicators of the same thing?

- What shall we decide?
- If the correlation between the two factors is sufficiently close to 1 (say > .9), we will be happy to save a couple of million euros in research money.
Can we consider all four variables as indicators of the same thing?

▶ What shall we decide?
▶ If the correlation between the two factors is sufficiently close to 1 (say $> .9$), we will be happy to save a couple of million euros in research money.
▶ The correlation is significantly different from 1.
Can we consider all four variables as indicators of the same thing?

- What shall we decide?
- If the correlation between the two factors is sufficiently close to 1 (say > .9), we will be happy to save a couple of million euros in research money.
- The correlation is significantly different from 1.
- When estimated it is 0.98.
Can we consider all four variables as indicators of the same thing?

▶ What shall we decide?
▶ If the correlation between the two factors is sufficiently close to 1 (say > .9), we will be happy to save a couple of million euros in research money.
▶ The correlation is significantly different from 1.
▶ When estimated it is 0.98.
▶ But according to all criteria, the model should be rejected.
Can we consider all four variables as indicators of the same thing?

- What shall we decide?
- If the correlation between the two factors is sufficiently close to 1 (say > .9), we will be happy to save a couple of million euros in research money.
- The correlation is significantly different from 1.
- When estimated it is 0.98.
- But according to all criteria, the model should be rejected.
- This is not what we wanted!
CAN WE CONSIDER ALL FOUR VARIABLES AS INDICATORS OF THE SAME THING?

▶ What shall we decide?
▶ If the correlation between the two factors is sufficiently close to 1 (say > .9), we will be happy to save a couple of million euros in research money.
▶ The correlation is significantly different from 1.
▶ When estimated it is 0.98.
▶ But according to all criteria, the model should be rejected.
▶ This is not what we wanted!
▶ Whether we used $\chi^2$, $\Delta \chi^2$, or any of the fit measures, we would make a wrong decision.
Can we consider all four variables as indicators of the same thing?
CAN WE CONSIDER ALL FOUR VARIABLES AS INDICATORS OF THE SAME THING?

Restricted model:
Can we consider all four variables as indicators of the same thing?

**Restricted model:**

MI: 4.589 \( (p = 0.032) \)
CAN WE CONSIDER ALL FOUR VARIABLES AS INDICATORS OF THE SAME THING?

Restricted model:

MI: 4.589 ($p = 0.032$)

EPC: -0.014
Can we consider all four variables as indicators of the same thing?

**Restricted model:**

- **MI:** 4.589 \( (p = 0.032) \)
- **EPC:** -0.014

(this would imply a correlation of
1 \( - 0.014 = 0.986 \))
Can we consider all four variables as indicators of the same thing?

Restricted model:

MI: 4.589 \( (p = 0.032) \)

EPC: -0.014

(this would imply a correlation of 
\[ 1 - 0.014 = 0.986 \])

Unrestricted model:

\[ \Delta \chi^2 = 5.0 \ (p = 0.025) \]

\[ \text{Estimate of correlation in unrestricted model: 0.981} \]

The MI is approximately equal to the improvement in chi-square when correlation is freed

The EPC is approximately equal to the misspecification (here \( 0.981 - 1 = -0.019 \approx -0.014 \))

This is true in general! (Buse 1982; Satorra 1989)
Can we consider all four variables as indicators of the same thing?

**Restricted** model:
- MI: 4.589 ($p = 0.032$)
- EPC: -0.014

(this would imply a correlation of $1 - 0.014 = 0.986$)

**Unrestricted** model:
- Difference in model chi-squares: $\Delta \chi^2 = 5.0 (p = 0.025)$
Can we consider all four variables as indicators of the same thing?

**Restricted model:**

- MI: 4.589 ($p = 0.032$)
- EPC: -0.014

(this would imply a correlation of $1 - 0.014 = 0.986$)

**Unrestricted model:**

- Difference in model chi-squares: $\Delta \chi^2 = 5.0 (p = 0.025)$
- Estimate of correlation in unrestricted model: 0.981
CAN WE CONSIDER ALL FOUR VARIABLES AS INDICATORS OF THE SAME THING?

Restricted model:
MI: 4.589 ($p = 0.032$)
EPC: -0.014

(this would imply a correlation of $1 - 0.014 = 0.986$)

-Unrestricted model:

- Difference in model chi-squares: $\Delta \chi^2 = 5.0 (p = 0.025)$
- Estimate of correlation in unrestricted model: $0.981$

- The MI is approximately equal to the improvement in chi-square when correlation is freed
Can we consider all four variables as indicators of the same thing?

**Restricted model:**
- MI: 4.589 \( (p = 0.032) \)
- EPC: -0.014

(this would imply a correlation of 1 – 0.014 = 0.986)

**Unrestricted model:**
- Difference in model chi-squares: \( \Delta \chi^2 = 5.0 (p = 0.025) \)
- Estimate of correlation in unrestricted model: 0.981

- The MI is approximately equal to the improvement in chi-square when correlation is freed
- The EPC is approximately equal to the misspecification (here 0.981 – 1 = –0.019 ≈ –0.014)
Can we consider all four variables as indicators of the same thing?

Restricted model:

MI: 4.589 \( (p = 0.032) \)

EPC: -0.014

(this would imply a correlation of \( 1 - 0.014 = 0.986 \))

Unrestricted model:

Difference in model chi-squares:
\( \Delta \chi^2 = 5.0(p = 0.025) \)

Estimate of correlation in unrestricted model: 0.981

► The MI is approximately equal to the improvement in chi-square when correlation is freed

► The EPC is approximately equal to the misspecification (here \( 0.981 - 1 = -0.019 \approx -0.014 \))

► This is true in general! (Buse 1982; Satorra 1989)
Can we consider all four variables as indicators of the same thing?

**Restricted model:**
- MI: 4.589 ($p = 0.032$)
- EPC: -0.014

  (this would imply a correlation of $1 - 0.014 = 0.986$)

**Unrestricted model:**
- Difference in model chi-squares: $\Delta \chi^2 = 5.0 (p = 0.025)$
- Estimate of correlation in unrestricted model: 0.981

  - The MI is approximately equal to the improvement in chi-square when correlation is freed
  - The EPC is approximately equal to the misspecification (here $0.981 - 1 = -0.019 \approx -0.014$)
  - This is true in general! (Buse 1982; Satorra 1989)
Can we consider all four variables as indicators of the same thing?

**Restricted model:**

MI: 4.589 ($p = 0.032$)

EPC: -0.014

(this would imply a correlation of $1 - 0.014 = 0.986$)

- The MI is approximately equal to the improvement in chi-square when correlation is freed
- The EPC is approximately equal to the misspecification (here $0.981 - 1 = -0.019 \approx -0.014$)
- This is true in general! (Buse 1982; Satorra 1989)
Can we consider all four variables as indicators of the same thing?

Restricted model:
MI: 4.589 \( (p = 0.032) \)
EPC: -0.014

(this would imply a correlation of 
\[ 1 - 0.014 = 0.986 \]

- The MI is approximately equal to the improvement in chi-square when correlation is freed
- The EPC is approximately equal to the misspecification (here \( 0.981 - 1 = -0.019 \approx -0.014 \))
- This is true in general! (Buse 1982; Satorra 1989)
WHAT IS THE PROBLEM?

- The misspecification of -0.014 is not substantively relevant
What is the problem?

- The misspecification of -0.014 is not substantively relevant
- The tests are very sensitive to this very small misspecification
WHAT IS THE PROBLEM?

- The misspecification of -0.014 is not substantively relevant
- The tests are very sensitive to this very small misspecification
- It appears the power of the test is very high
**What is the problem?**

- The misspecification of -0.014 is not substantively relevant
- The tests are very sensitive to this very small misspecification
- It appears the **power of the test** is very high
- Conclusion: when the power of the test is high and the test statistics indicate the model should be rejected, the EPC must be inspected.
WHAT IS THE PROBLEM?

▶ The misspecification of -0.014 is not substantively relevant
▶ The tests are very sensitive to this very small misspecification
▶ It appears the power of the test is very high
▶ Conclusion: when the power of the test is high and the test statistics indicate the model should be rejected, the EPC must be inspected.
▶ If the misspecification (EPC) does not exceed some threshold of acceptability, the model is not misspecified
What is the Problem?

- The misspecification of -0.014 is not substantively relevant
- The tests are very sensitive to this very small misspecification
- It appears the **power of the test** is very high

- Conclusion: when the power of the test is high and the test statistics indicate the model should be rejected, the EPC must be inspected.
- If the misspecification (EPC) does not exceed some threshold of acceptability, the model is **not misspecified**
- On the other hand, if the EPC *does* exceed the threshold, the model is **misspecified**
# Decision Rules

<table>
<thead>
<tr>
<th></th>
<th>High power</th>
<th>Low power</th>
</tr>
</thead>
<tbody>
<tr>
<td>Significant MI</td>
<td>Inspect EPC</td>
<td>Misspecification</td>
</tr>
<tr>
<td>Nonsignificant MI</td>
<td>No misspecification</td>
<td>Inconclusive</td>
</tr>
</tbody>
</table>

How do we obtain the power?
**DECISION RULES**

<table>
<thead>
<tr>
<th>Significant MI</th>
<th>Nonsignificant MI</th>
<th>High power</th>
<th>Low power</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Inspect EPC</td>
<td>Misspecification</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>No misspecification</strong></td>
<td>Inconclusive</td>
</tr>
</tbody>
</table>

How do we obtain the power?
Power

- Exercise: guess the sample size in the previous example …
POWER

- Exercise: guess the sample size in the previous example ...
- Hint: the power of the test was very high (close to 1.0)
Power

- Exercise: guess the sample size in the previous example …
- Hint: the power of the test was very high (close to 1.0)
- The correct answer is:

$n = 40$

The high power is due to the very large loadings

So power does not just depend on sample size. Things are not so simple.
Exercise: guess the sample size in the previous example …

Hint: the power of the test was very high (close to 1.0)

The correct answer is:

\[ n = 40 \]
Exercise: guess the sample size in the previous example …

Hint: the power of the test was very high (close to 1.0)

The correct answer is:

\[ n = 40 \]

The high power is due to the very large loadings
Exercise: guess the sample size in the previous example …

Hint: the power of the test was very high (close to 1.0)

The correct answer is:

\[ n = 40 \]

The high power is due to the very large loadings

So power does not *just* depend on sample size. Things are not so simple.
The power of the modification index test to detect a certain misspecification (say, \( \delta \)) can be determined just from the value of the MI and the EPC.

The program Jrule for Mplus (Oberski 2010) helps you make decisions about misspecifications.
The program Jrule for Mplus (Oberski 2010) helps you make decisions about misspecifications.

It reads in your Mplus output file and gives information about MI, EPC, the power of the MI test, and the recommended decision based on your own criteria.
The program Jrule for Mplus (Oberski 2010) helps you make decisions about misspecifications. It reads in your Mplus output file and gives information about MI, EPC, the power of the MI test, and the recommended decision based on your own criteria. It can be downloaded for free from http://wiki.github.com/daob/JruleMplus/
### Example 1: The Problem and a Solution

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Decision</th>
<th>Group</th>
<th>MI</th>
<th>EPC</th>
<th>Power</th>
<th>NCP</th>
</tr>
</thead>
<tbody>
<tr>
<td>GENE1</td>
<td>Not misspecified (EPC &lt; delta)</td>
<td>1</td>
<td>4.589</td>
<td>0.029</td>
<td>1.000</td>
<td>54.566</td>
</tr>
<tr>
<td>GENE2</td>
<td>Not misspecified (EPC &lt; delta)</td>
<td>1</td>
<td>4.589</td>
<td>0.029</td>
<td>1.000</td>
<td>54.566</td>
</tr>
<tr>
<td>GENE2 WITH GENE1</td>
<td>Not misspecified (EPC &lt; delta)</td>
<td>1</td>
<td>4.589</td>
<td>-0.014</td>
<td>1.000</td>
<td>234.133</td>
</tr>
</tbody>
</table>

### Example 2: Conclusions

The current output file is `Z:\home\daob\work\Presentations\Jrule\id-1fac.out`. 
EXAMPLE 2: PERSONALITY TRAITS AND VOTING

▶ “Big Five” personality traits: Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism
Example 2: Personality traits and voting

▶ “Big Five” personality traits: Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism

▶ Correlated with voting
Example 2: Personality Traits and Voting

- “Big Five” personality traits: Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism
- Correlated with voting
- Hypothesized to affect voting only indirectly, through things like “a sense that voting is a duty”, “political efficacy” (Gallego & Oberski, frth)
The problem and a solution

(All regression equations are also controlled for age, sex, and education -- not shown)
HIGHLY SIMPLIFIED PATH MODEL

Extraversion

Openness

Conscientiousness

Political efficacy

Sense of duty

Vote in EU elections

Should we introduce a path from Openness/Conscientiousness/Extraversion directly to Voting?

(all regression equations are also controlled for age, sex, and education -- not shown)
HIGHLY SIMPLIFIED PATH MODEL

Should we introduce a path from Openness/Conscientiousness/Extraversion directly to Voting?

I will conclude we should if the effect is bigger than 0.05.
THE MEDIATION MODEL ESTIMATED WITH MPLUS
The mediation model estimated with Mplus

Extraversion

Openness

Conscientiousness

Political efficacy

Sense of duty

Vote in EU elections

(all regression equations are also controlled for age, sex, and education -- not shown)
The mediation model estimated with Mplus

Extraversion → Openness → Political efficacy → Vote in EU elections

Conscientiousness → Sense of duty → Vote in EU elections

(all regression equations are also controlled for age, sex, and education -- not shown)
THE MEDIATION MODEL ESTIMATED WITH MPLUS

- Chisquare: 12.3, $df = 4^*$,
  $p = 0.0152$

(all regression equations are also controlled for age, sex, and education -- not shown)
THE MEDIATION MODEL ESTIMATED WITH Mplus

- Chisquare: 12.3, $df = 4^*$, $p = 0.0152$
- CFI: 0.965

(all regression equations are also controlled for age, sex, and education -- not shown)
The mediation model estimated with Mplus

- Chisquare: 12.3, \( df = 4^* \), 
  \( p = 0.0152 \)
- CFI: 0.965
- TLI: 0.948

(all regression equations are also controlled for age, sex, and education -- not shown)
The mediation model estimated with Mplus

- Chisquare: $12.3, df = 4^*$, $p = 0.0152$
- CFI: 0.965
- TLI: 0.948
- RMSEA: 0.026

(all regression equations are also controlled for age, sex, and education -- not shown)
The mediation model estimated with Mplus

- Chisquare: 12.3, $df = 4^*$, $p = 0.0152$
- CFI: 0.965
- TLI: 0.948
- RMSEA: 0.026
- WRMR: 0.885

(all regression equations are also controlled for age, sex, and education -- not shown)
The mediation model estimated with **Mplus**

- Chisquare: 12.3, \( df = 4^* \), \( p = 0.0152 \)
- CFI: 0.965
- TLI: 0.948
- RMSEA: 0.026
- WRMR: 0.885

- MI’s and EPC’s:
THE MEDIATION MODEL ESTIMATED WITH MPLUS

- Chisquare: 12.3, $df = 4^*$, $p = 0.0152$
- CFI: 0.965
- TLI: 0.948
- RMSEA: 0.026
- WRMR: 0.885

- MI’s and EPC’s:
  VOTE ON CONS  MI: 1.349, EPC: 0.062

(all regression equations are also controlled for age, sex, and education -- not shown)
The mediation model estimated with Mplus

- Chisquare: 12.3, $df = 4^*$, $p = 0.0152$
- CFI: 0.965
- TLI: 0.948
- RMSEA: 0.026
- WRMR: 0.885

- MI’s and EPC’s:
  VOTE ON CONS MI: 1.349, EPC: 0.062
  VOTE ON EXTR MI: 7.259**, EPC: 0.072

(all regression equations are also controlled for age, sex, and education -- not shown)
THE MEDIATION MODEL ESTIMATED WITH MPLUS

- Chisquare: 12.3, df = 4*, p = 0.0152
- CFI: 0.965
- TLI: 0.948
- RMSEA: 0.026
- WRMR: 0.885

- MI’s and EPC’s:
  VOTE ON CONS MI: 1.349 , EPC: 0.062
  VOTE ON EXTR MI: 7.259**, EPC: 0.072
  VOTE ON OPEN MI: 1.349 , EPC: -0.041

*df calculated for model with categorical variables (WLSMV estimator)
The model seems to fit well, except for the regression parameter “VOTE ON Extraversion”.

▶ So it seems our hypothesis that personality traits affect voting only indirectly is not rejected.

▶ Hooray?
The model seems to fit well, except for the regression parameter “VOTE ON Extraversion”.

So it seems our hypothesis that personality traits affect voting only indirectly is not rejected.
CONCLUSIONS

- The model seems to fit well, except for the regression parameter "VOTE ON Extraversion".
- So it seems our hypothesis that personality traits affect voting only indirectly is not rejected.
- Hooray?
### Example 1: The Problem and a Solution

#### Example 2: Conclusions

![Image of Mplus output]

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Decision</th>
<th>Group</th>
<th>MI</th>
<th>EPC</th>
<th>Power</th>
<th>NCP</th>
</tr>
</thead>
<tbody>
<tr>
<td>VOTEU09 ON OPEN</td>
<td>Inconclusive</td>
<td>1</td>
<td>1.419</td>
<td>-0.025</td>
<td>0.479</td>
<td>3.633</td>
</tr>
<tr>
<td>VOTEU09 ON EXTR</td>
<td>Misspecified (EPC &gt;= delta)</td>
<td>1</td>
<td>7.259</td>
<td>0.040</td>
<td>0.769</td>
<td>7.259</td>
</tr>
<tr>
<td>VOTEU09 ON CONS</td>
<td>Inconclusive</td>
<td>1</td>
<td>1.349</td>
<td>0.043</td>
<td>0.191</td>
<td>1.167</td>
</tr>
</tbody>
</table>
CONCLUSIONS

- The power to detect a misspecification for Openness and Conscientiousness is very low (0.479 and 0.191 respectively)
CONCLUSIONS

- The power to detect a misspecification for Openness and Conscientiousness is very low (0.479 and 0.191 respectively)
- Guess the sample size…
CONCLUSIONS

- The power to detect a misspecification for Openness and Conscientiousness is very low (0.479 and 0.191 respectively)
- Guess the sample size…
- $n = 3121$ (you probably saw that coming)
CONCLUSIONS

- The power to detect a misspecification for Openness and Conscientiousness is very low (0.479 and 0.191 respectively)
- Guess the sample size…
- $n = 3121$ (you probably saw that coming)
- The low power is due to small effects and the sampling design
CONCLUSIONS

- The power to detect a misspecification for Openness and Conscientiousness is very low (0.479 and 0.191 respectively)
- Guess the sample size…
- $n = 3121$ (you probably saw that coming)
- The low power is due to small effects and the sampling design

- The results on the possible presence of direct effects on voting from Openness and Conscientiousness can only be called inconclusive
Conclusions

- The power to detect a misspecification for Openness and Conscientiousness is very low (0.479 and 0.191 respectively)
- Guess the sample size…
- \( n = 3121 \) (you probably saw that coming)
- The low power is due to small effects and the sampling design

- The results on the possible presence of direct effects on voting from Openness and Conscientiousness can only be called inconclusive
- This means we need better measures or a better model or a bigger sample or a combination
Overall conclusions

- Chi square, fit measures, and MI are all affected by the power of the test
Overall conclusions

- Chi square, fit measures, and MI are all affected by the power of the test
- The power is not only a function of the sample size but can surprise you
OVERALL CONCLUSIONS

- Chi square, fit measures, and MI are *all* affected by the power of the test
- The power is not only a function of the sample size but can surprise you
- To make a correct decision, one must take into account the power of the test
Overall conclusions

- Chi square, fit measures, and MI are all affected by the power of the test
- The power is not only a function of the sample size but can surprise you
- To make a correct decision, one must take into account the power of the test
- Saris & a. (2009) suggest one method for doing this
OVERALL CONCLUSIONS

- Chi square, fit measures, and MI are *all* affected by the power of the test
- The power is not only a function of the sample size but can surprise you
- To make a correct decision, one must take into account the power of the test
- Saris & a. (2009) suggest one method for doing this
- That method is implemented in the free software Jrule for Mplus (Oberski 2010)
Thank you very much for your attention!

http://wiki.github.com/daob/JruleMplus/

daniel.oberski@upf.edu

This presentation: http://daob.org/