

---

# Final Exam

## Measurement models (Spring 2020)

MAE4101 Measurement Models

Centre for Educational Measurement at the University of Oslo (CEMO)

Date and time: 08 June 2020, 09:00-13:00

---

Welcome to the MAE4101 Measurement Models exam!

This exam covers the structural *and* the measurement model part of the course.

Before you begin, please make sure to consider the following:

- Read the questions carefully.
- Notice which task operators are used (e.g., name something vs. describe something).
- You may simplify subscripts or Greek symbols wherever appropriate (e.g.,  $Y_1$  instead of  $Y_1$ ,  $\lambda_1$  instead of  $\lambda_1$ ).
- Keep your explanations and descriptions brief.
- Partial credits will be given.

Starting with this home exam, you declare that you will work on the tasks without any help of others.

We wish you all the best for the exam and great success in working on the tasks!

Best regards,  
Denise, Jelena, Jarl, and Ronny

Name:	<b>SUGGESTED SOLUTIONS + GRADING</b>
-------	--------------------------------------

## Results

Task	Credits	Max. credits
Mediation		12
Complex path models		25
Confirmatory factor analysis		15
Family IQ Study		5
Measurement invariance		4
Monitoring and self-control		14
Genomics		7
<b>TOTAL:</b>		<b>82</b>

## Expected response time

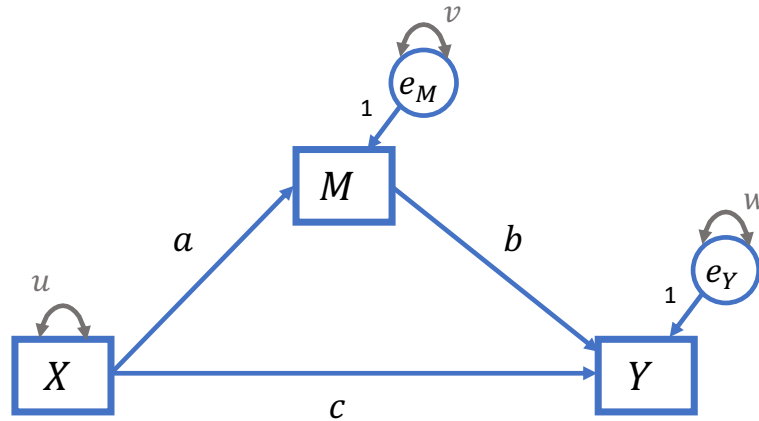
- Time the instructor needed to work on all tasks ( $t_I$ ):  $t_I = 50$  min
- Expected time for students ( $t_S$ ):  $t_S = 2.5t_I = 2.5 \cdot 50$  min = 125 min

## Grading

Grade	Credits threshold	% Correct threshold
A	74	90 %
B	65	80 %
C	57	70 %
D	49	60 %
E	41	50 %
F	< 41	< 50 %

## Mediation models (12 credits)

Mediation models have gained popularity in educational and psychological research. The figure below shows a typical mediation model with the three variables  $X$ ,  $M$ , and  $Y$ . Note that this model does *not* have a mean structure.



a) Provide the model equations for  $M$  and  $Y$ .

Fill in the equations here.

Variables	Model equation
$M =$	$aX + e_M$
$Y =$	$bM + cX + e_Y$

SCORING:

- 1 credit per correct equation
- **Total: 2 credits**

b) Using the notation of the mediation model, provide the formulas of the indirect and the total effect of  $X$  on  $Y$  via  $M$ .

Provide the formulas below.

Effect	Formula
Indirect effect	$ab$
Total effect	$ab + c$

SCORING:

- 1 credit per correct formula
- **Total: 2 credits**

c) Verify that the following elements of the model-implied covariance matrix can be expressed like this:

- $Var(X) = u$
- $Var(M) = a^2u + v$
- $Cov(X, M) = au$

Fill in the analytic steps here.

Component	Formula
$Var(X) =$	$Var(X) = u$ , because $X$ is an exogenous (independent) variable
$Var(M) =$	$Var(M) = Var(aX + e_M)$ $= a^2Var(X) + Var(e_M) + 2Cov(X, e_M)$ $= a^2u + v$
$Cov(X, M) =$	$Cov(X, M) = Cov(X, aX + e_M) = aCov(X, X) + Cov(X, e_M)$ $= aVar(X) + 0 = au$

SCORING:

- 2 credit per correct proof:
  - 1 credit for applying the variance and covariance laws
  - 1 credit for simplifying all terms
- **Total: 6 credits**

d) Suppose a mean structure is added to the mediation model. Provide the model equations for  $M$  and  $Y$  with a mean structure.

Fill in the equations here.

Variables	Model equation
$M =$	$i_M + aX + e_M$
$Y =$	$i_Y + bM + cX + e_Y$

SCORING:

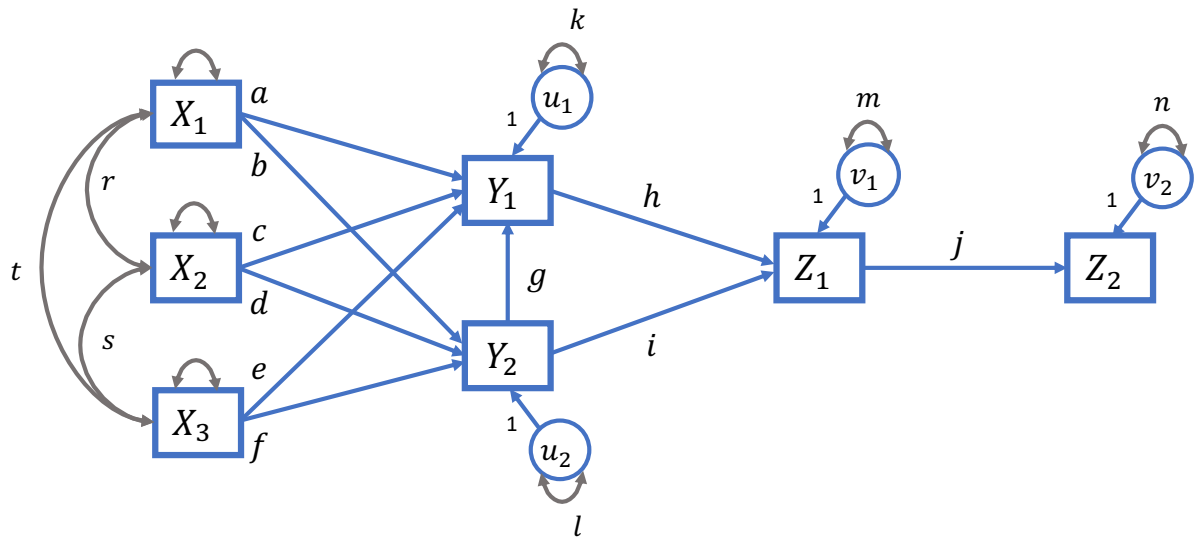
- 1 credit per correct equation
- **Total: 2 credits**

## Complex path models (25 credits)

Path models describe the structural relations among observed (manifest) variables and represent researchers' theories and hypotheses about these relations.

The following path diagram shows a path model including 7 variables (**Model 1**). Note that this model does *not* have a mean structure.

Model 1



- a) *Divide and conquer*: Identify the endogenous and exogenous manifest variables and the residual variables in the model.

Check the appropriate boxes below.

Variable	Exogenous variable	Endogenous variable	Residual
$X_2$	X	<input type="checkbox"/>	<input type="checkbox"/>
$Y_1$	<input type="checkbox"/>	X	<input type="checkbox"/>
$Y_2$	<input type="checkbox"/>	X	<input type="checkbox"/>
$u_1$	<input type="checkbox"/>	<input type="checkbox"/>	X
$Z_2$	<input type="checkbox"/>	X	<input type="checkbox"/>
$v_1$	<input type="checkbox"/>	<input type="checkbox"/>	X

### SCORING:

- 1 credit per correct column or 0.5 credits per correct response:
  - Independent variable column: 1 credit
  - Dependent variable column: 1 credit
  - Residual column: 1 credit
- Total: 3 credits**

b) Provide the model equations for the endogenous variables in Model 1.

Equations for the endogenous variables:
$Y_1 = 0 \cdot Y_1 + g \cdot Y_2 + 0 \cdot Z_1 + 0 \cdot Z_2 + a \cdot X_1 + c \cdot X_2 + e \cdot X_3 + u_1$ $= g \cdot Y_2 + a \cdot X_1 + c \cdot X_2 + e \cdot X_3 + u_1$
$Y_2 = 0 \cdot Y_1 + 0 \cdot Y_2 + 0 \cdot Z_1 + 0 \cdot Z_2 + b \cdot X_1 + d \cdot X_2 + f \cdot X_3 + u_2$ $= b \cdot X_1 + d \cdot X_2 + f \cdot X_3 + u_2$
$Z_1 = h \cdot Y_1 + i \cdot Y_2 + 0 \cdot Z_1 + 0 \cdot Z_2 + 0 \cdot X_1 + 0 \cdot X_2 + 0 \cdot X_3 + v_1 = h \cdot Y_1 + i \cdot Y_2 + v_1$
$Z_2 = 0 \cdot Y_1 + 0 \cdot Y_2 + j \cdot Z_1 + 0 \cdot Z_2 + 0 \cdot X_1 + 0 \cdot X_2 + 0 \cdot X_3 + v_2 = j \cdot Z_1 + v_2$

SCORING:

- 1 credit per correct equation
- **Total: 4 credits**

c) Determine the number of available pieces of information ( $p$ ), the number of parameters that need to be estimated in the model ( $q$ ), and the resultant degrees of freedom of the model ( $df_M$ ). Conclude whether or not the model is identified.

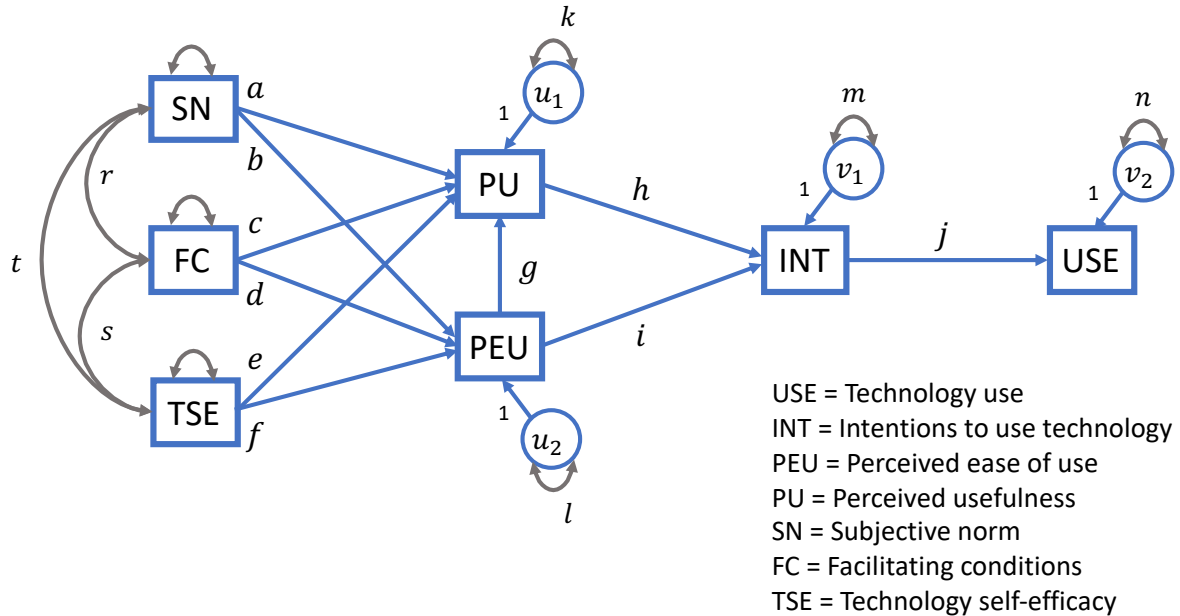
Indices	Numbers
# Observed pieces of information ( $p$ )	$p = \frac{7(7+1)}{2} = 28$
# Parameters to be estimated ( $q$ )	$q = 10 + 4 + 6 = 20$
Degrees of freedom of the model ( $df_M$ )	$df_M = 28 - 20 = 8$
Conclusion	The model is over-identified, $df_M > 0$ .

SCORING:

- 1 credit per correct number ( $p$ ,  $q$ , and  $df_M$ )
- 1 credit for the correct conclusion that the model is identified or over-identified.
- Note:  $q = 10$  path coefficients + 4 residual variances + 6 variances and covariances of the exogenous variables
- **Total: 4 credits**

The famous Technology Acceptance Model (TAM)—a model describing the mechanisms behind a person’s intentions to use (INT) and use of technology (USE)—can be represented by Model 1.

TAM (Model 1):



d) Using the syntax of the R package lavaan, specify the TAM (Model 1) by providing the command lines for the structural relations below. The model is labelled “tam”.

Lavaan code for model specification:

```
tam <- '
  # Structural relations
  USE ~ INT
  INT ~ PEU + PU
  PEU ~ TSE + FC + SN
  PU ~ PEU + TSE + FC + SN

  # Variances and covariances of exogenous variables
  SN ~~ SN + FC + TSE
  FC ~~ FC + TSE
  TSE ~~ TSE
'
```

SCORING:

- 1 credit per correct line for the endogenous variables USE, INT, PEU, and PU
- 1 credit for specifying correctly the variances of the exogenous variables SN, FC, and TSE
- 1 credit for specifying correctly the covariances among the exogenous variables SN, FC, and TSE
- **Total: 6 credits**

e) Two indirect effects of PEU on USE via PU and INT exist. How can these two effects be estimated? Use the labels of the path coefficients and provide the formulas below.

Effect	Formula
Indirect effect 1	$ghj$
Indirect effect 2	$ij$

**SCORING:**

- 1 credit per correct formula
- **Total: 2 credits**

f) A researcher estimated the TAM using a sample of  $N = 576$  in-service teachers to explain their use of digital teaching (USE). In light of the path coefficients shown below, examine which of the following statements are true.

Path coefficients:

Coefficient	Estimate	SE	p-value
$a$	0.249	0.046	.000
$b$	0.197	0.043	.000
$c$	0.268	0.045	.000
$d$	0.304	0.042	.000
$e$	0.149	0.045	.001
$f$	0.440	0.040	.000
$g$	0.605	0.043	.000
$h$	0.436	0.038	.000
$i$	0.426	0.045	.000
$j$	0.634	0.032	.000

Statement	True	False
(1) Perceived ease of use (PEU) does not show a direct effect on the intentions to use technology (INT).	<input type="checkbox"/>	<input checked="" type="checkbox"/>
(2) There is evidence for the positive and significant relation between the intentions to use technology (INT) and the actual use of technology (USE).	<input checked="" type="checkbox"/>	<input type="checkbox"/>
(3) The effects of the exogenous variables on PU are the same as those on PEU.	<input type="checkbox"/>	<input checked="" type="checkbox"/>

**NOTES:**

- Statement (1): FALSE, because the direct effect ( $i$ ) is statistically different from zero.
- Statement (2): TRUE, because the direct effect ( $j$ ) is statistically different from zero.



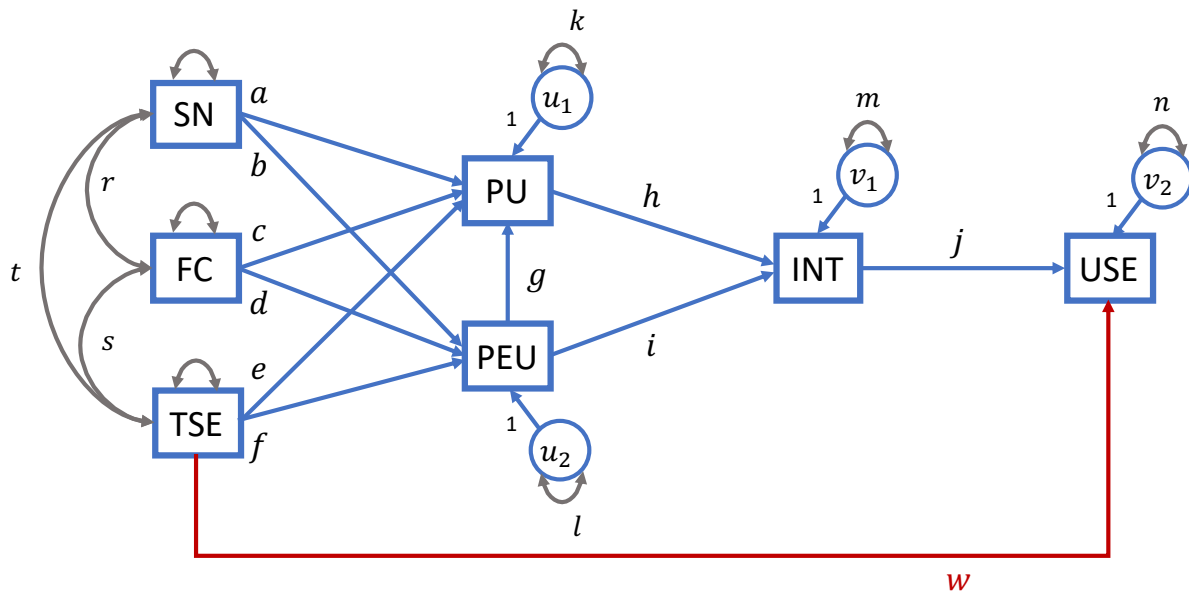
- Statement (3): FALSE, because the effects have different values and the researchers did not explicitly test for their equality.

SCORING:

- 1 credit per statement
- Total: 3 credits**

g) The researcher modified this model by including one additional path connecting technology self-efficacy (TSE) and technology use (USE). This new model is labelled **Model 2**, and the new path coefficient is labelled  $w$ .

Modified TAM (Model 2):



Fit index	Model 1	Model 2
$\chi^2$	123.014, $p < .001$	8.561, $p = .286$
CFI	0.912	0.999
RMSEA	0.158	0.020
SRMR	0.066	0.021
AIC	11551.255	11438.802
BIC	11638.377	11530.280
Model comparison	$\Delta\chi^2(1, N = 576) = 114.45, p < .001$	

Evaluating the model fit indices and the results of the model comparison (Model 1 vs. Model 2), decide which of the two models represents the data better. Provide a brief reasoning for your decision that considers at least two sources of information.

Your evaluation:

Model 2 represents the data better than Model 1. Model 2 is preferred over Model 1.

Possible reasoning:

- Model 2 fits the data perfectly, while Model 1 exhibits only a marginal fit to the data. The latter is indicated by the CFI = .912, the RMSEA = .066, and the significant

chi-square statistic (indicating a significant discrepancy between the observed and the model-implied covariance matrices).

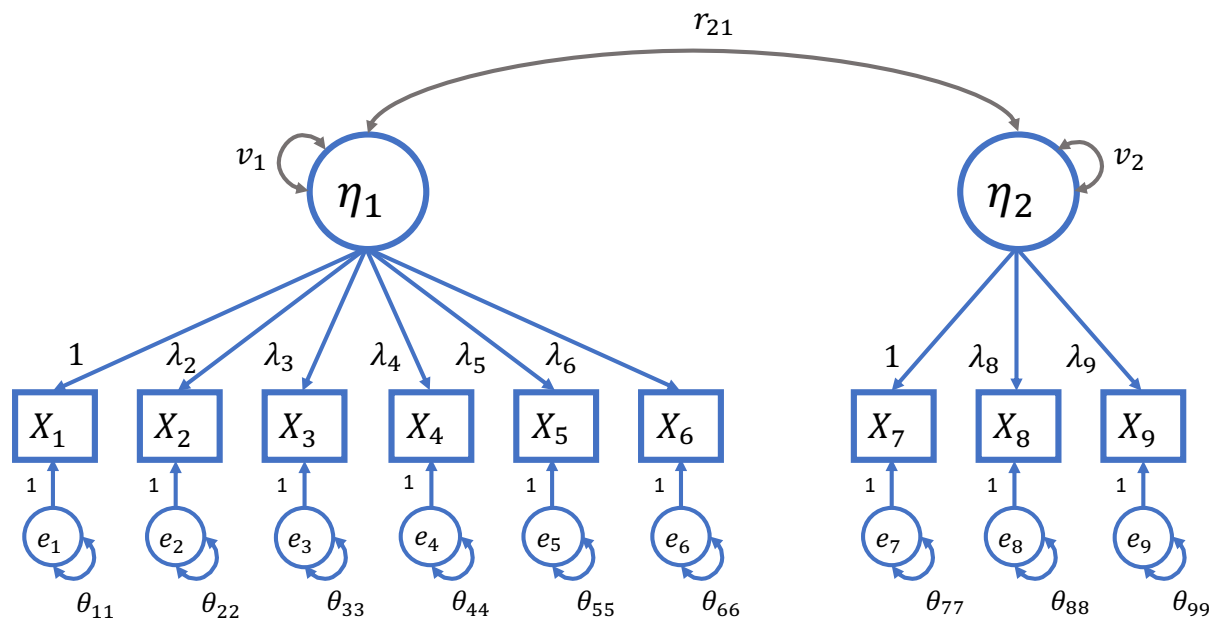
- The information criteria of Model 1 are consistently larger than those for Model 2, pointing to the preference of Model 2.
- The model comparison shows that the chi-square values of the two models differ significantly in favor of Model 2.

**SCORING:**

- 1 credit for the correct decision for Model 2
- 1 credit per correct argument (two are needed in total)
- **Total: 3 credits**

## Confirmatory factor analysis (15 credits)

The figure below shows a confirmatory factor analysis (CFA) model involving two factors  $\eta_1$  and  $\eta_2$ . Note that this model does *not* have a mean structure.



- a) Determine whether Model 1 represents a **reflective** or a **formative** measurement model and explain why.

Provide your answer here.

### Explanation

The model is **reflective**. All indicators are considered to be caused by latent variable. The latent variable (exogenous variable) is a “predictor” of indicators. Indicators have measurement error since not all variance is explained by latent variable. When we remove latent variable, we expect no covariation between indicators. Therefore, the only correlation between indicators should be (ideally) due to latent variable.

### SCORING:

- 1 credit for correct type of measurement model
- 1 credit for the correct explanation
- **Total: 2 credits**

- b) Verify that the model is **identified with 26 degrees of freedom**.

### Model identification

Degrees of freedom of the model are determined by the following elements:

- Pieces of information from the data ( $p$ ): 9 variances and  $8+7+6+5+4+3+2+1$  covariances of the 9 manifest variables  $X_1, \dots, X_9 \rightarrow p = 45$
- Number of parameters in the model (without a mean structure): 2 factor variances, 1 factor covariance, 5+2 factor loadings, 9 residual variances  $\rightarrow q = 19$

- Degrees of freedom of the model:  $df_M = 45 - 19 = 26 > 0 \rightarrow$  The model is (over-) identified.

**SCORING:**

- 1 credit for the correct number p
- 1 credit for the correct number q
- 1 credit for the difference p-q and the conclusion (>0)
- **Total: 3 credits**

c) Write out the model equations for the indicator variables  $X_1$ ,  $X_3$ , and  $X_4$ .

Fill in the equations here.

Indicator variables	Model equation
$X_1 =$	$\eta_1 + e_1$
$X_3 =$	$\lambda_3 \eta_1 + e_3$
$X_4 =$	$\lambda_4 \eta_1 + e_4$

**SCORING:**

- 1 credit per correct equation
- **Total: 3 credits**

d) Using the variance rules and the model specification, verify that the model-implied variance of  $X_3$  is  $Var(X_3) = \lambda_3^2 \cdot v_1 + \theta_{33}$ .

Fill in the equations here.

Component	Formula
$Var(X_3) =$	$  \begin{aligned}  Var(\lambda_3 \eta_1 + e_3) &= \lambda_3^2 \cdot Var(\eta_1) + Var(e_3) + 2Cov(\lambda_3 \eta_1, e_3) \\  &= \lambda_3^2 \cdot v_1 + \theta_{33} + 2\lambda_3 Cov(\eta_1, e_3) \\  &= \lambda_3^2 \cdot v_1 + \theta_{33} + 2\lambda_3 \cdot 0 = \lambda_3^2 \cdot v_1 + \theta_{33}  \end{aligned}  $

**SCORING:**

- 1 credit for inserting the model specification equation for  $X_3$
- 1 credit for applying correctly the variance rule to the sum  $\lambda_3 \eta_1 + e_3$
- 1 credit for handling correctly the coefficients in brackets ( $\lambda_3$ )
- 1 credit for recognizing the zero covariance  $Cov(\eta_1, e_3)$
- **Total: 4 credits**

e) Derive the following element of the model-implied covariance matrix in Model 1:  
 $Cov(X_3, X_4)$ .

Fill in the equation here.

Component	Formula
$Cov(X_3, X_4) =$	$\lambda_3\lambda_4v_1$

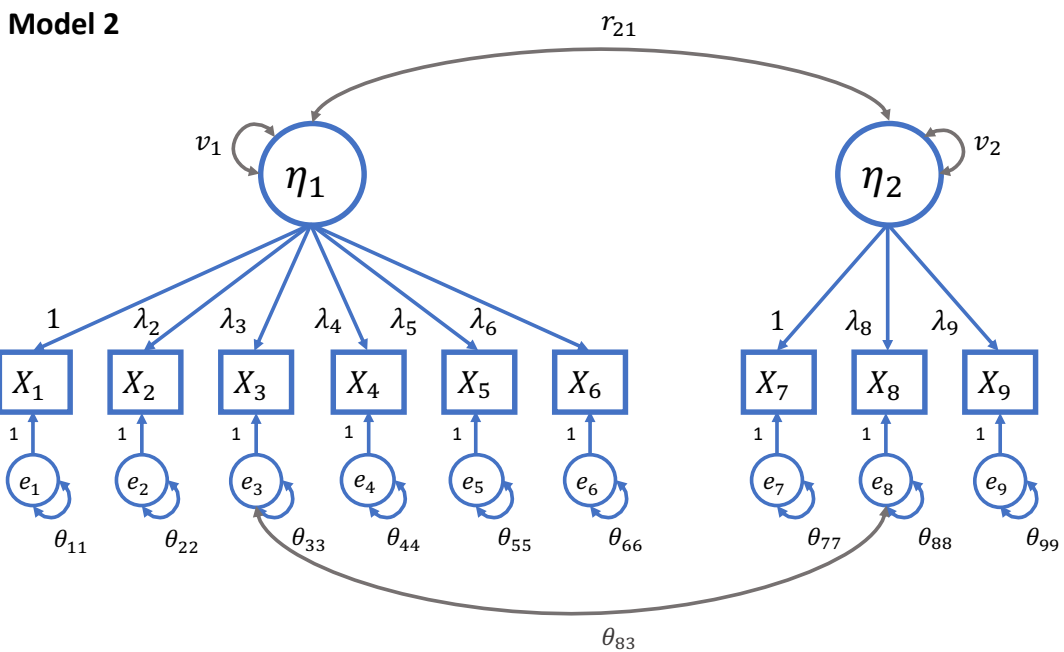
**NOTES:**

- $Cov(X_3, X_4) = Cov(\lambda_3\eta_1 + e_3, \lambda_4\eta_1 + e_4) = \lambda_3\lambda_4Cov(\eta_1, \eta_1) + \lambda_3Cov(\eta_1, e_4) + \lambda_4Cov(e_3, \eta_1) + Cov(e_3, e_4) = \lambda_3\lambda_4Var(\eta_1) + 0 + 0 + 0 = \lambda_3\lambda_4v_1$

**SCORING:**

- 1 credit for the correct formula
- Students may derive these elements using the variance and covariance rules or Wright's tracing rules.
- **Total: 1 credit**

After some inspection of the model fit and parameters, Model 1 has been slightly modified, as shown below (Model 2).



f) What is the meaning of the new parameter  $\theta_{83}$  in Model 2? How can researchers test whether adding this new parameter to the original model (Model 1) is justified?

Provide your answer here.

Model parameter  $\theta_{83}$

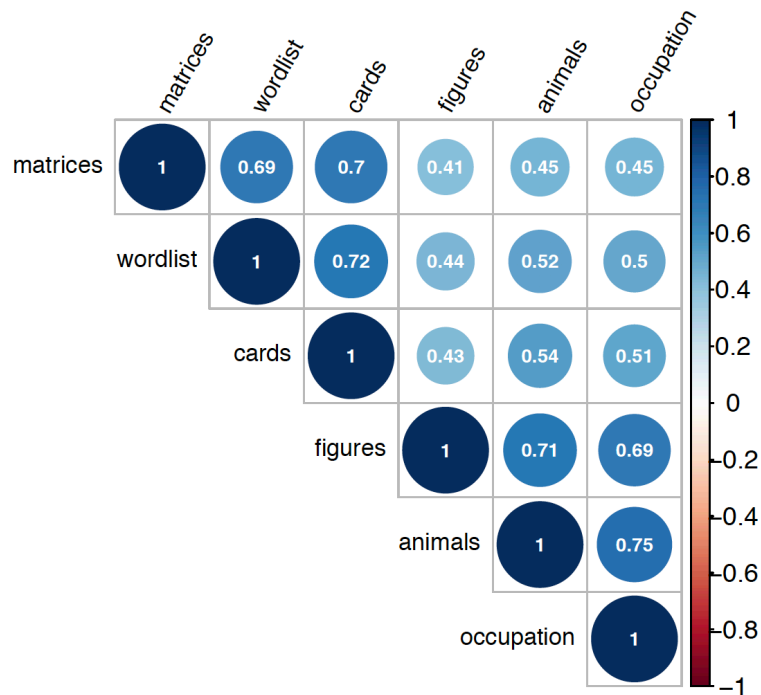
The new model parameter  $\theta_{83}$  represents the **covariance between the residuals  $e_3$  and  $e_8$** . To test whether adding this parameter is justified, researchers can **compare the fit of the two CFA models, that is, the original model and the modified model** via chi-square difference testing, differences in information criteria or other fit indices. Inspecting its confidence intervals may provide evidence for its significant deviation from zero.

SCORING:

- 1 credit for the explanation of the model parameter
- 1 credit for the description of a testing procedure
- **Total: 2 credits**

## Family IQ Study (5 credits)

The famous “Family IQ Study” obtained eight test scores of cognitive skills from  $N = 399$  children. The following correlogram is based on these scores.



a) What does this correlogram show? Interpret the findings.

### Interpretation

- The correlogram shows the correlation matrix, that is, shows the correlations among the eight scores.
- Inspecting this matrix, it is possible to conclude that the indicators are positively correlated with each other. This suggests that there is a general tendency for children who score high in one indicator will also score high in the others.

### SCORING:

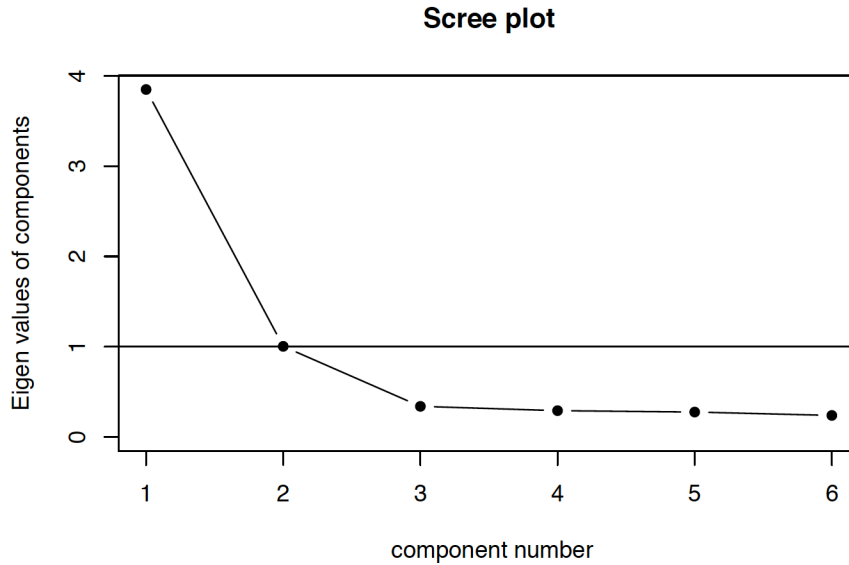
- 1 credit for indicating that correlations are shown.
- 1 credit for explaining the positive relations among the variables.
- **Total: 2 credits**

b) Extracting the eigenvalues from the correlation matrix provided the following output:

```
# Extract the eigenvalues
evals <- eigen(dat.cor)$values
evals

## [1] 3.8494373 1.0041103 0.3394470 0.2918007 0.2763080 0.2388966
# Ratios of eigenvalues
(evals/sum(evals)*100)[1:4]

## [1] 64.157289 16.735172 5.657451 4.863346
```



How many factors can be extracted from the data? Explain your choice and provide the proportion of variance explained by each of the selected factor(s).

Number of factors	Explanation
2	<p><b>Examples:</b></p> <ul style="list-style-type: none"> <li>▪ Based on the screeplot, the last substantial decline in the magnitude of the eigenvalues indicates the number of factors to be extracted.</li> <li>▪ Based on the Kaiser rule, two eigenvalues above 1 exist.</li> <li>▪ The elbow criterion suggests one strong elbow from 1→2 and one minor elbow from 2→3 factors.</li> </ul>
Proportion of explained variance by factor(s)	
<ul style="list-style-type: none"> <li>▪ Proportion of variance explained by the first factor: 64.16 %</li> <li>▪ Proportion of variance explained by the second factor: 16.74 %</li> </ul>	

**SCORING:**

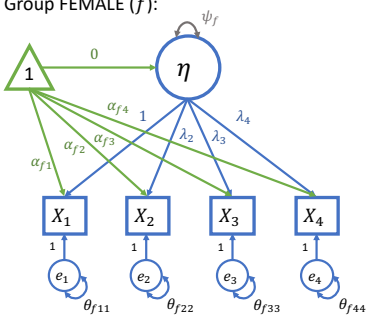
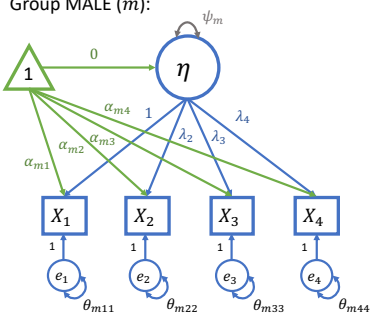
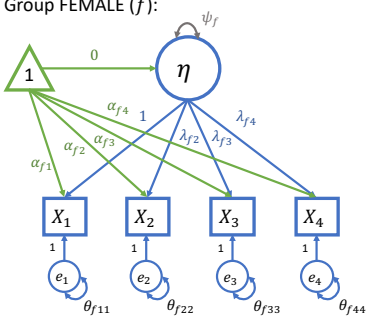
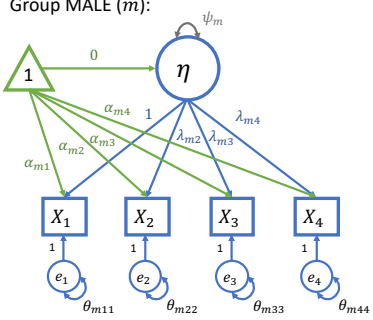
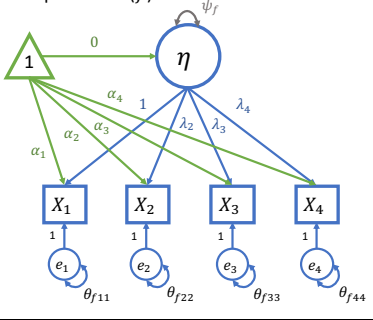
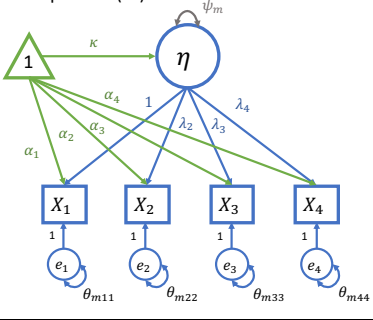
- 1 credit the correct number of factors
- 1 credit for the correct explanation (one reason/criterion is sufficient)
- 1 credit for the correct variance explanations
- **Total: 3 credits**



## Measurement invariance testing (4 credits)

The following figures represent three models of multi-group confirmatory factor analysis for the latent variable  $\eta$  across gender (**Models 1-3**). These figures depict several invariance constraints.

a) For each of the figures, indicate which level(s) of measurement invariance are shown in the path diagram. Check the appropriate box(es).

<p><b>Model 1</b></p> <p>Group FEMALE (<math>f</math>):</p>  <p>Group MALE (<math>m</math>):</p> 	Configural invariance	<input checked="" type="checkbox"/>
	Metric invariance	<input checked="" type="checkbox"/>
	Scalar invariance	<input type="checkbox"/>
	Strict invariance	<input type="checkbox"/>
<p><b>Model 2</b></p> <p>Group FEMALE (<math>f</math>):</p>  <p>Group MALE (<math>m</math>):</p> 	Configural invariance	<input checked="" type="checkbox"/>
	Metric invariance	<input type="checkbox"/>
	Scalar invariance	<input type="checkbox"/>
	Strict invariance	<input type="checkbox"/>
<p><b>Model 3</b></p> <p>Group FEMALE (<math>f</math>):</p>  <p>Group MALE (<math>m</math>):</p> 	Configural invariance	<input checked="" type="checkbox"/>
	Metric invariance	<input checked="" type="checkbox"/>
	Scalar invariance	<input checked="" type="checkbox"/>
	Strict invariance	<input type="checkbox"/>

### SCORING:

- 1 credit for each figure with the correct solution
- **Total: 3 credits**

b) Which of the three models must fit the data in order for researchers to compare the means of the latent variable  $\eta$  across gender?

Model	Selection
Model 1	<input type="checkbox"/>
Model 2	<input type="checkbox"/>
Model 3	X

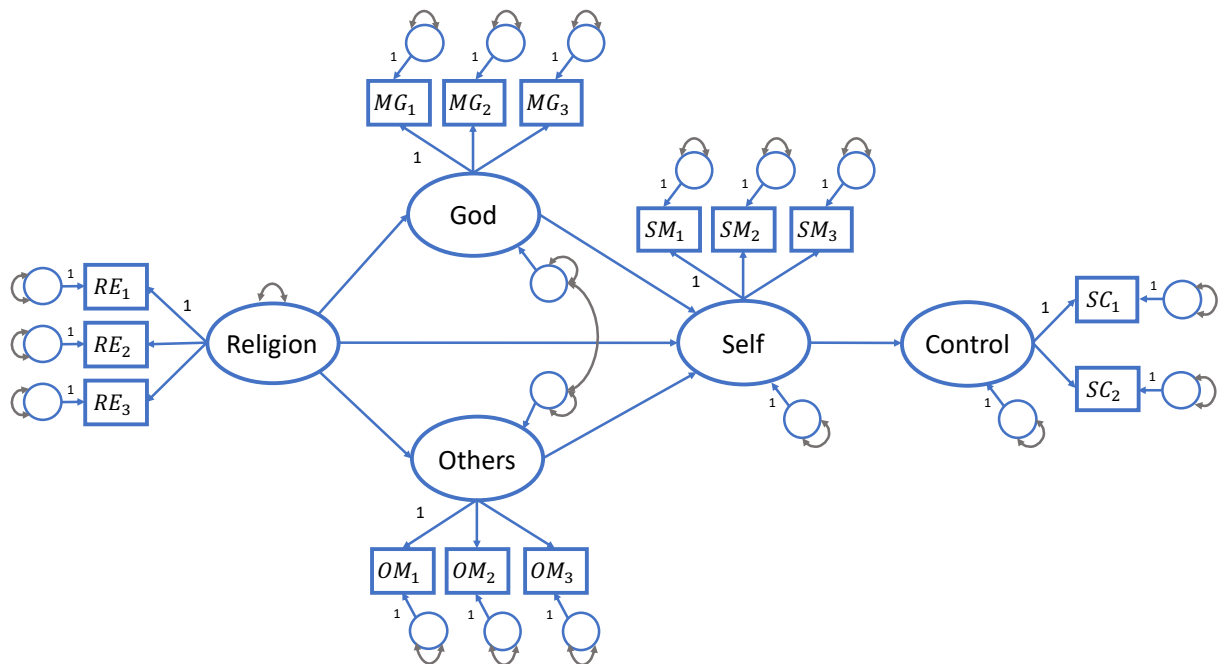
SCORING:

- 1 credit for the correct solution
- **Total: 1 credit**

## Monitoring and Self-Control (14 credits)

Carter et al. (2012, DOI:10.1177/1948550612438925) proposed a model that describes the relations between religiosity (**Religion**), perceived monitoring by oneself (**Self**), by others (**Others**), by God (**God**), and people's self-control (**Control**).

The corresponding structural equation model is shown below.



- a) Using the syntax of the R package lavaan, specify this model by providing the command lines below. The model is labelled "model1".

Lavaan code for model specification:

```
model1 <- '
  # Measurement models
  RELIGION =~ RE1+RE2+RE3
  GOD =~ MG1+MG2+MG3
  SELF =~ SM1+SM2+SM3
  OTHERS =~ OM1+OM2+OM3
  CONTROL =~ SC1+SC2

  # Structural relations
  CONTROL ~ SELF
  SELF ~ OTHERS+GOD+RELIGION
  GOD ~ RELIGION
  OTHERS ~ RELIGION

  # Covariances
  GOD ~~ OTHERS
'
```

### SCORING:

- 1 credit per correct line for the measurement models (5 in total)

- 1 credit per correct line for the structural relations (4 in total)
- 1 credit for specifying correctly the covariance between God and Others
- **Total: 10 credits**

b) Carter et al. (2012) evaluated this model for a sample of  $N = 583$  participants and obtained the following fit indices:

```
## Model Test User Model:
##
## Test statistic 60.201
## Degrees of freedom 70
## P-value (Chi-square) 0.792
##
## Model Test Baseline Model:
##
## Test statistic 2129.580
## Degrees of freedom 91
## P-value 0.000
##
## User Model versus Baseline Model:
##
## Comparative Fit Index (CFI) 1.000
## Tucker-Lewis Index (TLI) 1.006
##
## Loglikelihood and Information Criteria:
##
## Loglikelihood user model (H0) -13213.024
## Loglikelihood unrestricted model (H1) -13182.924
##
## Akaike (AIC) 26496.048
## Bayesian (BIC) 26648.935
## Sample-size adjusted Bayesian (BIC) 26537.823
##
## Root Mean Square Error of Approximation:
##
## RMSEA 0.000
## 90 Percent confidence interval - lower 0.000
## 90 Percent confidence interval - upper 0.017
## P-value RMSEA <= 0.05 1.000
##
## Standardized Root Mean Square Residual:
##
## SRMR 0.023
```

Examine which of the following fit indices are acceptable according to the criteria proposed by Hu and Bentler (1999). Check the appropriate boxes.

Fit index	Not acceptable	Acceptable
Chi-square test statistic	<input type="checkbox"/>	X
CFI	<input type="checkbox"/>	X
RMSEA	<input type="checkbox"/>	X
SRMR	<input type="checkbox"/>	X

SCORING:

- 0.5 credit per correct decision
- **Total: 2 credits**

c) Carter et al. (2012) hypothesized that the **direct effects** of the variables **Religion, God,** and **Others** on the variable **Self** are the same. To test this hypothesis, they compared two models:

- The model with freely estimated path coefficients (**model1**).
- The model that constrains the path coefficients to equality (**model2**).

Carter et al. (2012) have obtained the following output in lavaan:

```
# Model comparison
anova(model1.fit, model2.fit)

## Chi-Squared Difference Test
##
##           Df   AIC   BIC  Chisq Chisq diff Df diff Pr(>Chisq)
## model1.fit 67 26508 26674 55.119
## model2.fit 69 26506 26663 56.512    1.3928     2    0.4984
```

Decide whether or not these results support their hypothesis and provide a brief reasoning.

Your decision:

These numbers support the hypothesis.

Reasoning:

There is no significant difference in the model-data fit (chi-square statistics), suggesting that the model constraints of equality do not deteriorate the model fit significantly. The information criteria AIC and BIC are lower for the model with equal path coefficients.

SCORING:

- 1 credit for the correct decision
- 1 credit for the correct reasoning (information and/or chi-square difference testing)
- **Total: 2 credits**

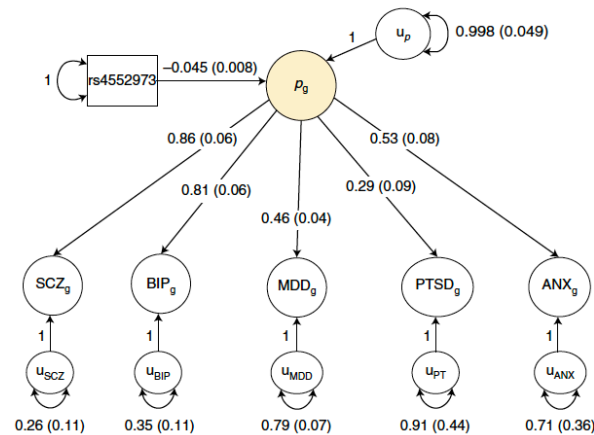
## Genomics (7 credits)

Confirmatory factor analysis can be extended by adding some predictor variables. A recently published paper on genomics and personality presented two examples of these extended models (Grotzinger et al., 2019, DOI: 10.1038/s41562-019-0566-x, p. 516):

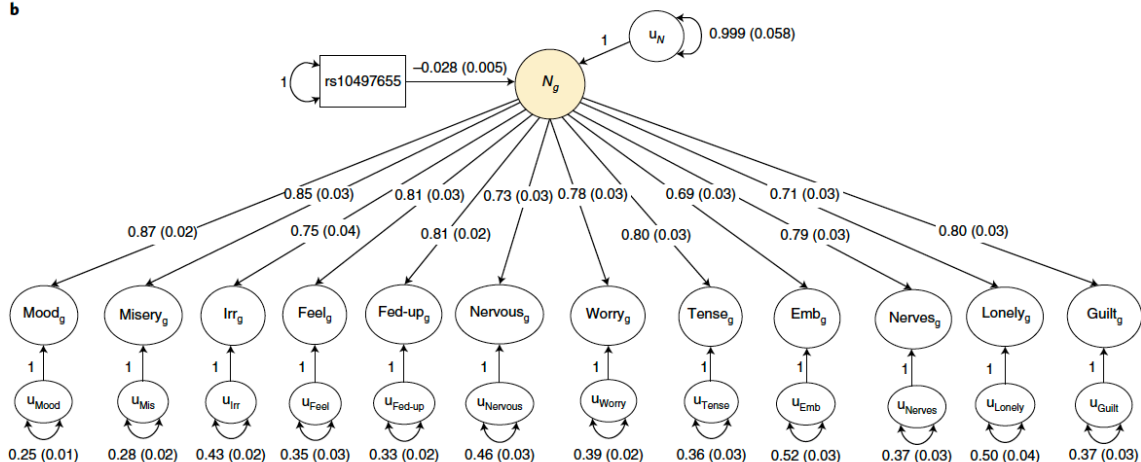
### ARTICLES

### NATURE HUMAN BEHAVIOUR

a



b



**Model A** presents the relation between a general factor of psychopathology ( $p_g$ ) and the predictor variable “individual single-nucleotide polymorphism rs455279”.

*Note. SCZ-ANX = Indicators of psychopathology (e.g., bipolar disorder, anxiety).*

**Model B** presents the relation between a general factor of neuroticism ( $N_g$ ) and the predictor variable “individual single-nucleotide polymorphism rs1049765”.

*Note. Mood-Guilt = Indicators of neuroticism (e.g., guilt, irritability).*

a) Interpret the standardized regression coefficient  $b = -0.028$  in **Model B**.

Your interpretation:

Examples:

- A one SD unit increase in the genomics variable rs1049765 results in a decrease of 0.028 SD units in neuroticism.

- The larger the score of the genomics variable rs1049765, the lower the neuroticism score.

SCORING:

- **Total: 1 credit**

b) What do the variable  $u_N$  and the corresponding parameter value of 0.999 represent in **Model B**? Explain briefly.

Your explanation:

- The latent variable  $u_N$  represents the residual of the latent variable  $N_g$ , that is, the difference between the value of  $N_g$  and the predicted value of  $N_g$  after for the regression model with rs1049765 as the predictor.
- The corresponding parameter value of 0.999 represents the corresponding residual variance (i.e., variance not explained by the predictor rs1049765).

SCORING:

- 1 credit for the correct explanation of what  $u_N$  represents.
- 1 credit for the correct explanation of the parameter value.
- **Total: 2 credits**

c) Confirmatory factor analysis allows researchers to estimate the **reliability of a scale**. For a CFA model without residual covariances, the scale reliability can be estimated as McDonald's Omega from the standardized model parameters of the items  $j = 1, \dots, J$  as

$$\omega = \frac{(\sum_{j=1}^J \lambda_j)^2}{(\sum_{j=1}^J \lambda_j)^2 + \sum_{j=1}^J \theta_{jj}}$$

How can the reliability of the **scale measuring  $p_g$  in Model A** be calculated from the standardized model parameters as they are shown in the figure?

*Note: You do not need to provide the final result of this calculation.*

Scale reliability	Calculation
$\omega =$	$\frac{(0.86 + 0.81 + 0.46 + 0.29 + 0.53)^2}{(0.86 + 0.81 + 0.46 + 0.29 + 0.53)^2 + (0.26 + 0.35 + 0.79 + 0.91 + 0.71)}$

SCORING:

- 1 credit for the correct numerator
- 1 credit for the correct denominator
- **Total: 2 credits**

- d) The scale reliability of  $p_g$  is  $\omega = 0.74$  in Model A. The scale reliability of  $N_g$  is  $\omega = 0.95$  in Model B. Explain these differences in the two reliability coefficients.

#### Discussion

The scale reliability in Model A is smaller as the one in Model B for several reasons:

- Differences in factor loadings and residual variances between the two models
- Model A shows heterogeneous factor loadings (some are smaller than others), while Model B shows consistently high factor loadings.
- Model A has less indicator variables of the latent variable than Model B.

#### SCORING:

- 1 credit for recognizing that the reliability is a function of factor loadings, residual variances, and the number of indicators.
- 1 credit for a correct explanation of the differences.
- **Total: 2 credits**