Final Exam *(Evaluator Version)*

Measurement Models Course (Spring Term 2022)

MAE4101 Measurement Models
Centre for Educational Measurement at the University of Oslo (CEMO)

Exam date: 08 June 2022, 9:00-13:00, Oslo time, Silurveien 2, Sal 4D

Welcome to the MAE4101 Measurement Models on-site 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* versus *explain* something).
- You may **simplify subscripts** wherever appropriate (e.g., Y1 instead of Y_1).
- You may **simplify Greek letters** wherever appropriate (e.g., lambda1 instead of λ_1).
- Use **dots instead of commas** to indicate decimals.
- Keep your **explanations and descriptions brief**.
- **Partial credits** will be given.

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

Best regards,
Denise, Jarl, Kseniia, and Ronny

Abbreviations you may want to use throughout the exam:
CFA—Confirmatory factor analysis
CFI—Comparative fit index
EFA—Exploratory factor analysis
MGCFA—Multiple-groups CFA
RMSEA—Root mean square error of approximation
SEM—Structural equation model
SRMR—Standardized root mean squared residual

Name:	SUGGESTED SOLUTIONS + GRADING

Results

Task	Credits	Max. credits	Evaluator(s)
Path Models		16	RS
The Subjective Happiness Scale		10	DRC
Equivalent CFA Models		16	RS
Exploratory Factor Analysis		12	JKK
Complex SEM		16	RS
Multiple-Groups CFA		10	KM
TOTAL:		**80**	

Expected time on task

- Time spent on the exam by teachers: 50 minutes
- Estimated time for students: 50 minutes*2.5 = 125 minutes

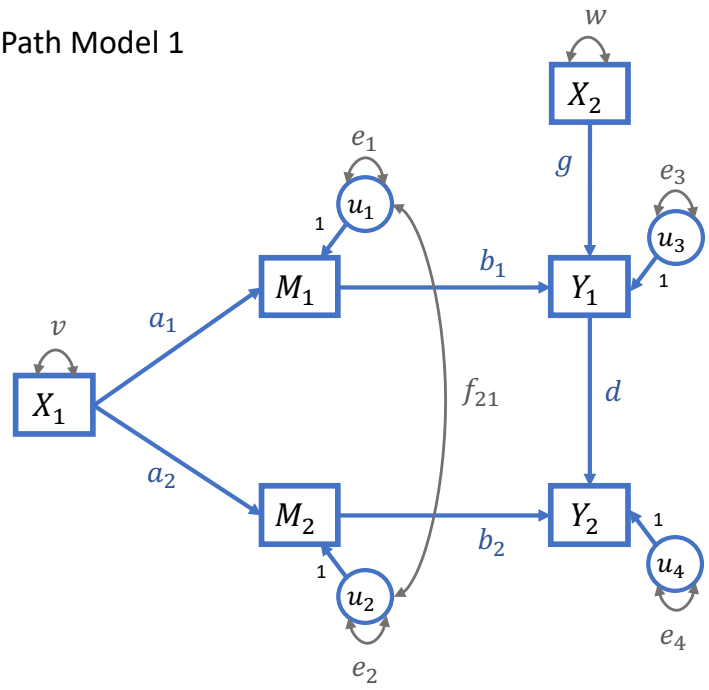
Grading

Grade	Credits range	Percentages correct
A	80-72	100-90 %
B	71-64	89-80 %
C	63-56	79-70 %
D	55-48	69-60 %
E	47-40	59-50 %
F	39-0	49-0 %

Path Models (16 credits)

Path models describe relationships between observed (manifest) variables. In this task, we consider the following path model (**Path Model 1**):

Path Model 1



Note: This model does not include a mean structure.

a) *Divide and conquer:* Identify the roles of the following variables in the model as **predictors, outcomes, and residuals**. Check the appropriate boxes below.

Note: Some variables may have multiple roles.

Variable	Predictor	Outcome	Residual
X_1	X	☐	☐
M_1	X	X	☐
M_2	X	X	☐
u_1	☐	☐	X
u_2	☐	☐	X
X_2	X	☐	☐
Y_1	X	X	☐
Y_2	☐	X	☐

u_3	☐	☐	X

NOTE:
This task is implemented slightly differently in Inspera due to the technical limitations of the system. The overall credits are still 3. Yet, the response options have changed into: Predictor, Outcome, Residual, Predictor and Outcome, Predictor and Residual, Residual and Outcome, and Predictor, Outcome, and Residual. Each correct answer is credited with 0.33 credits.

SCORING:

- 0.25 credit per correct response:
  - Predictor column: 1 credit
  - Outcome column: 1 credit
  - Residual column: 1 credit
- **Sub-total: 3 credits**

b) Specify Path Model 1 by providing the **model equations for the variables that have at least one predictor** (i.e., endogeneous variables).

Fill in the model equations here.

Outcome variable	Model equation
$M_1 =$	$a_1X_1 + u_1$
$M_2 =$	$a_2X_1 + u_2$
$Y_1 =$	$b_1M_1 + gX_2 + u_3$
$Y_2 =$	$b_2M_2 + dY_1 + u_4$

SCORING:

- 1 credit per correct equation
- **Sub-total: 4 credits**

c) Decide which of the following matrices represent the **variances and covariances of the residuals** in Path Model 1.

Covariance matrix of residuals	Yes, this matrix represents the variances and covariances of residuals.

$\begin{bmatrix} f_{21} & e_1 & e_2 & e_3 \\ e_1 & f_{21} & e_3 & e_4 \\ e_2 & e_3 & f_{21} & e_4 \\ e_3 & e_4 & e_4 & f_{21} \end{bmatrix}$	☐
$\begin{bmatrix} e_1 & 0 & 0 & 0 \\ 0 & e_2 & 0 & 0 \\ 0 & 0 & e_3 & 0 \\ 0 & 0 & 0 & e_4 \end{bmatrix}$	☐
$\begin{bmatrix} e_1 & f_{21} & 0 & 0 \\ f_{21} & e_2 & 0 & 0 \\ 0 & 0 & e_3 & 0 \\ 0 & 0 & 0 & e_4 \end{bmatrix}$	X
$\begin{bmatrix} e_1 & 0 & 0 & 0 \\ f_{21} & e_2 & 0 & 0 \\ 0 & 0 & e_3 & 0 \\ 0 & 0 & 0 & e_4 \end{bmatrix}$	☐

NOTES:

- Matrix 1: Incorrect matrix, because the covariances are placed in the diagonal, and the matrix is not symmetric.
- Matrix 2: Incorrect matrix, because it assumes that all residuals are independent, that is, have no covariances.
- Matrix 3: Correct matrix: the residual variances are stored in the diagonal, and the covariances are off-diagonal; plus, the matrix is symmetric.
- Matrix 4: Incorrect matrix, because it is not symmetric.

SCORING:

- 1 credit for the correct response
- **Sub-total: 1 credit**

d) Provide the following **two elements of the model-implied covariance matrix** from Path Model 1: $Var(M_1)$ and $Cov(X_1, M_2)$.

Fill in the equations here.

Element	Model equation
$Var(M_1)$	$= Var(a_1X_1 + u_1)$ $= Var(a_1X_1) + Var(u_1) + 2Cov(a_1X_1, u_1)$ $= a_1^2 Var(X_1) + e_1 + 2a_1 Cov(X_1, u_1)$ $= a_1^2 v + e_1 + 2a_1 \cdot 0$ $= a_1^2 v + e_1$
$Cov(X_1, M_2)$	$= Cov(X_1, a_2X_1 + u_2)$ $= Cov(X_1, a_2X_1) + Cov(X_1, u_2)$

	$= a_2 Cov(X_1, X_1) + 0$ $= a_2 Var(X_1)$ $= a_2 v$

NOTES:

- The end result counts; students do not need to provide the steps they have taken to arrive there.
- Students may apply the Wright tracing rules.

SCORING:

- 1 credit for the correct number of terms in each element
- 1 credit for the correct indices in each element
- **Sub-total: 2 credits**

e) Provide the code for specifying Path Model 1 in the R package lavaan.
This code should estimate all direct effects. Indirect effects are not needed. If the model contains **exogeneous manifest variables**, include their variances as well.

Note: The model is named "PathModel1" in this task. You do not need to name the path coefficients, variances, or covariances in the code.

Lavaan code for model specification:

```
PathModel1 <- '
              # Structural model
              M1 ~ X1
              M2 ~ X1
              Y1 ~ M1 + X2
              Y2 ~ M2 + Y1

              # Additional (co-)variances
              M1 ~~ M2     # Residual covariance
              X1 ~~ X1     # Exogeneous variable
              X2 ~~ X2     # Exogeneous variable

              '
```

SCORING:

- 1 credit for each correctly specified outcome variable (4 credits)
- 1 credit for including the residual covariance
- 1 credit for the inclusion of the two variances of the exogeneous variables X1 and X2
- **Sub-total: 6 credits**

The Subjective Happiness Scale (10 credits)

Lyubomirsky and Lepper (1999) developed a 4-item self-report measure of happiness, the so-called "Subjective Happiness Scale (SHS)". The response format is a 7-point Likert-type scale with higher scores reflecting greater happiness. The item wordings are provided below.

Item	1	2	3	4	5	6	7
Q1) In general, I consider myself…	Not a very happy person						A very happy person
Q2) Compared to most of my peers, I consider myself…	Less happy						More happy
Q3) Some people are generally very happy. They enjoy life regardless of what is going on, getting the most out of everything. *To what extent does this characterization describe you?*	Not at all						A great deal
Q4) Some people are generally not very happy. Although they are not depressed, they never seem as happy as they might be. *To what extent does this characterization describe you?*	Not at all						A great deal

Source: Lyubomirsky & Lepper (1999, DOI: 10.1023/A:1006824100041).

a) A researcher wants to represent the construct underlying the SHS by a reflective measurement model with Q1-Q4 as indicators. In the path diagram of this measurement model, how are the following elements represented? Check the respective geometric shape.

Element	Geometric shape
Construct "Subjective Happiness"	☐ Triangle ☐ Square/rectangle X Circle/ellipse ☐ Hexagon
Construct indicators Q1-Q4	☐ Triangle X Square/rectangle ☐ Circle/ellipse ☐ Hexagon
Intercepts of the construct indicators Q1-Q4	X Triangle ☐ Square/rectangle ☐ Circle/ellipse ☐ Hexagon
Residuals of the construct indicators Q1-Q4	☐ Triangle ☐ Square/rectangle X Circle/ellipse

	☐ Hexagon

SCORING:

- 1 credit per correct answer
- **Sub-total: 4**

b) After collecting data, the researcher ran a CFA and reported the following conclusions in a scientific paper (see below). As a reviewer, evaluate whether each statement is true or false *and* explain your decision briefly.

Statement	Decision	Explanation
A *p*-value of the chi-square test statistic equal to 0.76 indicates that the model fits the data well.	X True ☐ False	An insignificant *p*-value suggests that the data do not support the alternative hypothesis S≠∑. Hence, good model fit is indicated.
A modification index of 3.14 for "Q3~~Q4" reflects an approximation of how much the overall χ2 will decrease if we fix the covariance between Q3 and Q4 to 3.14.	☐ True X False	A modification index reflects an approximation of how much the overall χ2 will decrease if a fixed or constrained parameter is freely estimated.
An RMSEA of 0.038 and an SRMR of 0.020 indicate an acceptable model fit.	X True ☐ False	Following Hu's and Bentler's (1999) fixed cut-offs, both values are below their suggestions, RMSEA ≤ .06 and SRMR ≤ .08.

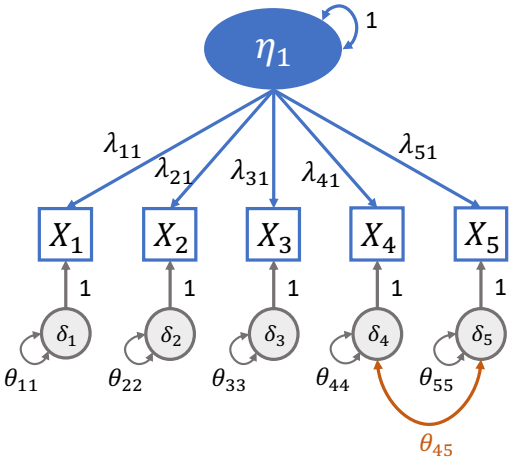
SCORING:

- 1 credit per correct answer
- 1 credit for a correct explanation
- **Sub-total: 6**

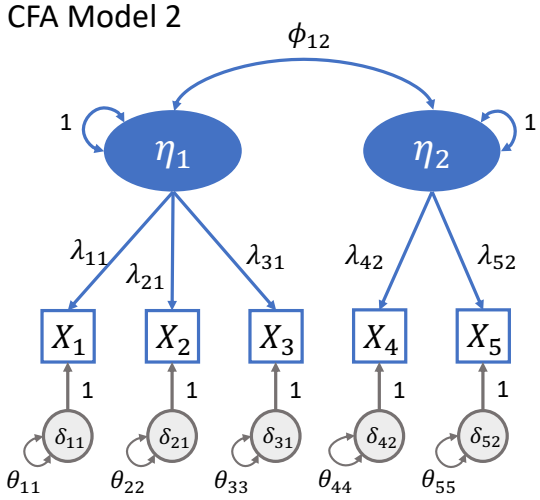
Equivalent CFA Models (16 credits)

In the following, you will find the path diagrams of two CFA models, “CFA Model 1” and “CFA Model 2”. These two models do not contain a mean structure.

CFA Model 1



CFA Model 2



a) How many **pieces of information are available** to estimate the two models?

	CFA Model 1	CFA Model 2
Available pieces of information (p)	15	15

NOTES:

- 5 variances (X1-X5)
- 4+3+2+1=10 covariances (among X1-X5)
- Alternative: p = 5*(5+1)/2

SCORING:

- 0.5 credit for each correct number
- **Sub-total: 1 credit**

b) Count the **number of freely estimated model parameters** in the two models and complete the table below.

Model parameters	CFA Model 1	CFA Model 2
Factor loadings	5	5
Residual variances	5	5
Residual covariances	1	0
Factor variances	0	0
Factor covariances	0	1

SCORING:

- 0.4 credits for each correct number
- **Sub-total: 4 credits**

c) How many **degrees of freedom** do the two models have?

	CFA Model 1	CFA Model 2
Degrees of freedom of the model (df_M)	4	4

NOTES:

- Degrees of freedom of the models: $df_M = p - q = 15 - 11 = 4$

SCORING:

- 0.5 credit for each correct number
- **Sub-total: 1 credit**

d) Which of the following statements are true or false?

Statement	True	False
The two models, CFA Model 1 and 2, are exactly identified (i.e., just-identified).	☐	X
The degrees of freedom of the two models CFA Model 1 and CFA Model 2 are the same.	X	☐

NOTES:

- Statement (1): False, because the degrees of freedom of these models are positive, yet not zero.
- Statement (2): True. See question (c).

SCORING:

- 1 credit for each correct response
- **Sub-total: 2 credits**

e) Provide the **model-specifying equations of the indicators X_3 and X_4** for CFA Models 1 and 2.

Fill in the model equations here.

Outcome variables	CFA Model 1	CFA Model 2
$X_3 =$	$\lambda_{31}\eta_1 + \delta_3$	$\lambda_{31}\eta_1 + \delta_{31}$
$X_4 =$	$\lambda_{41}\eta_1 + \delta_4$	$\lambda_{42}\eta_2 + \delta_{42}$

SCORING:

- 1 credit for each correct equation
- **Sub-total: 4 credits**

f) **Show** that the **model-implied covariance between X_3 and X_4 in CFA Model 1** is $Cov(X_3, X_4) = \lambda_{31}\lambda_{41}$.

Note: Provide the steps you have taken to show this result.

Fill in the equations and any comments here.

Model-implied element	CFA Model 1
$Cov(X_3, X_4)$	$= Cov(\lambda_{31}\eta_1 + \delta_3, \lambda_{41}\eta_1 + \delta_4)$ $= Cov(\lambda_{31}\eta_1, \lambda_{41}\eta_1) + Cov(\lambda_{31}\eta_1, \delta_4) + Cov(\delta_3, \lambda_{41}\eta_1) + Cov(\delta_3, \delta_4)$ $= \lambda_{31}\lambda_{41}Cov(\eta_1, \eta_1) + \lambda_{31}Cov(\eta_1, \delta_4) + \lambda_{41}Cov(\delta_3, \eta_1) + Cov(\delta_3, \delta_4)$ $= \lambda_{31}\lambda_{41}Var(\eta_1) + \lambda_{31} \cdot 0 + \lambda_{41} \cdot 0 + 0$ $= \lambda_{31}\lambda_{41} \cdot 1$ $= \lambda_{31}\lambda_{41}$

SCORING:

- 1 credit for inserting the model-specifying expressions of X3 and X4
- 1 credit for applying correctly the covariance rules to the sum of random variables
- 1 credit for correctly applying the covariance rules to the product of an integer and a random variable
- 1 credit for stating the independence of some random variables (zero covariances)
- **Sub-total: 4 credits**

Exploratory Factor Analysis (12 credits)

Tibi and Kirby (2017) investigated the dimensionality of a test battery assessing children's morphological knowledge with a sample of *N* = 102 third graders. They administered ten tests of morphological knowledge and explored the dimensionality using EFA.

The tests are named as follows:

- Word selection
- Sentence completion
- Morphological decomposition
- Morphological composition
- Morphological relation judgment
- Standard word analogy
- Local word analogy
- Sentence analogy
- Word analysis
- Picture choice

a) Name and describe the **four fundamental steps of EFA**.

Step	Explanation
1.	**Factor extraction**: Define the parameter estimation procedure (e.g., ML, PF, PCA)
2.	**Factor selection**: Define the strategy for factor selection (e.g., Kaiser-Guttman rule, scree test, parallel analysis)
3.	**Factor rotation**: Define the rotation technique (e.g., Varimax, Promax)
4.	**Factor score**: Define the method for the calculation of individual scores (e.g., Thurstone's regression)

SCORING:

- 0.5 credit per correct naming of each step
- 0.5 credit per correct description
- **Sub-total: 4 credits**

b) Explain what **"common variance" and "unique variance"** mean.

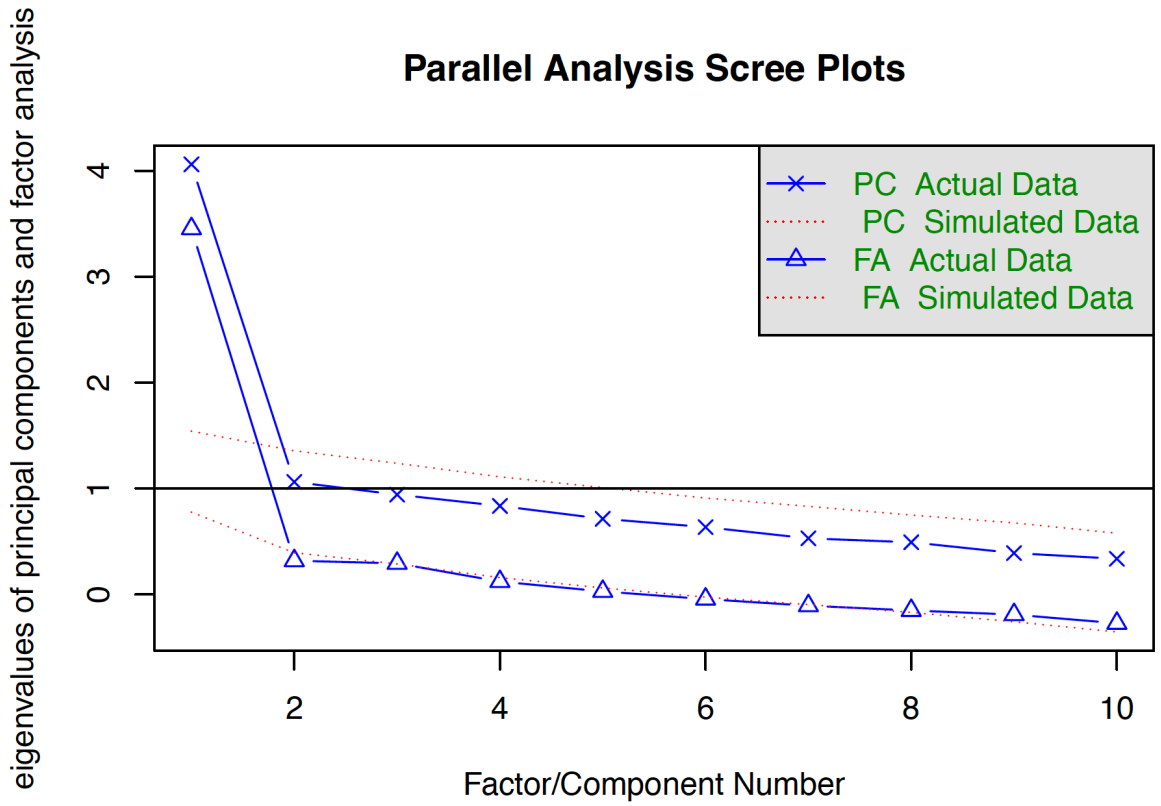
Term	Explanation
Common variance	Common variance is the variance accounted for by the factor, which is estimated on the basis of variance shared with other indicators in the analysis.
Unique variance	Unique variance is a combination of reliable variance that is specific to the indicator (i.e., systematic factors that influence only one indicator) and random error variance (i.e., measurement error or unreliability in the indicator).

SCORING:

- 1 credit per correct explanation
- **Sub-total: 2 credits**

c) The table below shows eigenvalues of the data from Tibi's and Kirby's (2017) study, and the figure shows the corresponding scree plots.

Eigenvalues	1	2	3	4	5	6	7	8	9	10
Eigenvalues of factors	3.45	0.32	0.29	0.12	0.03	-0.04	-0.11	-0.15	-0.19	-0.27
Eigenvalues of simulated factors	0.78	0.39	0.29	0.16	0.06	-0.03	-0.10	-0.17	-0.26	-0.35
Eigenvalues of components	4.06	1.06	0.94	0.84	0.71	0.64	0.53	0.49	0.39	0.34
Eigenvalues of simulated components	1.54	1.36	1.24	1.11	1.01	0.91	0.83	0.75	0.68	0.58



How many factors are suggested using the scree plot, parallel analysis, and the Kaiser-Guttman criteria for principal components (PC) and factor analysis (FA)?

Method	Number of factors (PC)	Number of factors (FA)
Scree plot	1	1
Parallel analysis	1	1
Kaiser-Guttman criterion	2	1

SCORING:

- 0.5 credits per correct number
- **Sub-total: 3 credits**

d) Having conducted a principal axis factor analysis, Tibi and Kirby (2017) presented the factor loadings of a one-factor and a two-factor solution in the following table.

Note: Tibi and Kirby labelled the two factors "Oral MA" and "Written MA" (MA = Morphological awareness).

Source: Tibi & Kirby (2017), DOI:10.1017/S0142716417000029, p. 1035.

	Factor Loading		
		Two Factor Solution	
Test	One Factor Solution	Oral MA	Written MA
Word selection	**.586**	**.341**	**.308**
Sentence completion	**.701**	.018	**.769**
Morphological decomposition	**.621**	.107	**.578**
Morphological composition	**.631**	−.005	**.708**
Morphological relation judgment	**.361**	**.584**	−.153
Standard word analogy	**.733**	**.522**	.298
Local word analogy	**.664**	**.453**	.288
Sentence analogy	**.610**	**.485**	.201
Word analysis	**.367**	.286	.126
Picture choice	**.473**	**.407**	.127

Note: Direct oblimin rotation used in two-factor solution. Loadings > .300 are in bold.

Which of the following statements are true or false?

Statement	True	False
In the two-factor solution, each test loads on only one factor.	☐	X
In the two-factor solution, all tests have loadings above .30 on at least one factor.	☐	X
Factor loadings of the one-factor solution are larger than the factor loadings for the two-factor solution on all tests.	☐	X

NOTES:

- Statement (1): False, because some tests have non-zero loadings on both factors (e.g., Word selection).
- Statement (2): False, because the test Word analysis does not have loadings above .30.

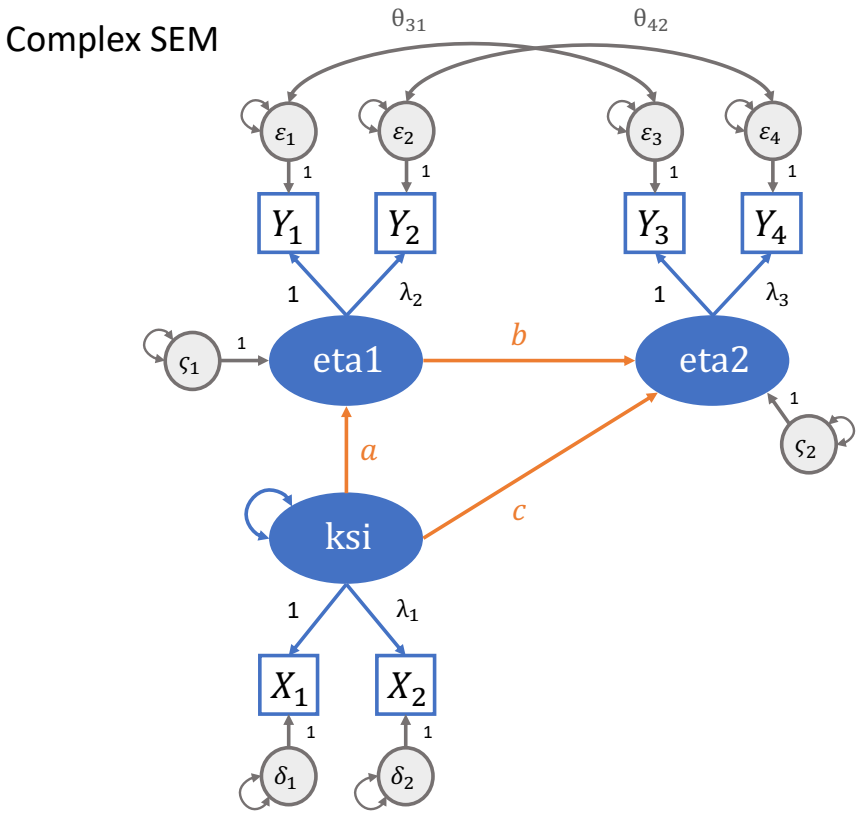
- Statement (3): False, because some loadings are smaller (e.g., Sentence completion).

SCORING:

- 1 credit for each correct response
- **Sub-total: 3 credits**

Complex SEM (16 credits)

Researchers in the Social Sciences are often interested in the direct and indirect effects of a predictor on an outcome variable. The following structural equation model (without a mean structure) quantifies such effects.



a) Which of the following statements about the model "Complex SEM" are true or false?

Statement	True	False
The structural part of this model describes the relationships between the latent variables ksi, $eta1$, and $eta2$.	X	☐
The latent variable ksi is exogeneous in this model.	X	☐
This model contains 6 parameters that need to be estimated.	☐	X
The indirect effect of ksi on $eta2$ via $eta1$ is $ab + c$.	☐	X

NOTES:

- Statement (1): True, because the model defines the structural relationships among these three variables.
- Statement (2): True, because this latent variable does not have a cause in the model.
- Statement (3): False, because the model contains 3 factor loadings, 5+2 residual variances, 2 residual covariances, 3 structural path coefficients, and 1 additional variance of the exogeneous latent variable ksi that need to be estimated.
- Statement (4): False, because the indirect effect only contains the term ab.

SCORING:

- 1 credit for each correct response
- **Sub-total: 4 credits**

f) Describe **one method to examine if the two residual covariances θ_{31} and θ_{42} are needed** in this model.

Description of one method:
Possible methods: ▪ Estimate the model with the residual covariances and check if these covariances are significantly different from zero ▪ Estimate two models, one with and one without the two residual covariances, and compare them via chi-square difference testing or differences in model fit indices ▪ Estimate the model without the residual covariances and check if the modification indices suggest an improvement in model fit when including the two residual covariances ▪ Estimate the model without the residual covariances and check if the model residuals flag the two residual covariances (i.e., have large values)

SCORING:

- 1 credit for describing an appropriate method
- 1 credit for the correct use of terminology
- **Sub-total: 2 credits**

g) Provide the code for specifying the model in the R package lavaan.
This code should allow you to estimate the **indirect effect** of ksi on $eta2$ via $eta1$. If the model contains **exogeneous latent variables**, include their variances as well.

Note: The model is named "ComplexSEM" in this task.

Lavaan code for model specification:

```
ComplexSEM <-  '
               # Measurement models
               ksi =~ X1 + X2
               eta1 =~ Y1 + Y2
               eta2 =~ Y3 + Y4

               # Structural model
               eta2 ~ b*eta1 + c*ksi
               eta1 ~ a*ksi

               # Additional (co-)variances
               Y1 ~~ Y3     # Residual covariance
               Y2 ~~ Y4     # Residual covariance
               ksi ~~ ksi   # Exogeneous latent variable

               # Additional parameters
               ind:=a*b   # Indirect effect (ind)
```

'

NOTES:

- The line "`eta2~b*eta1+c*ksi`" can also be written as two separate lines.

SCORING:

- 1 credit for each correct measurement model (3 credits)
- 1 credit for each correct element in the structural model (3 credits for a, b, and c)
- 1 credit for each correct residual covariance (2 credits)
- 1 credit for the variance of ksi
- 1 credit for the correct definition of the indirect effect (*Note:* 0 credit is given if the indirect effect is correctly defined, but the labelling of the structural path coefficients a and b is missing.)
- **Sub-total: 10 credits**

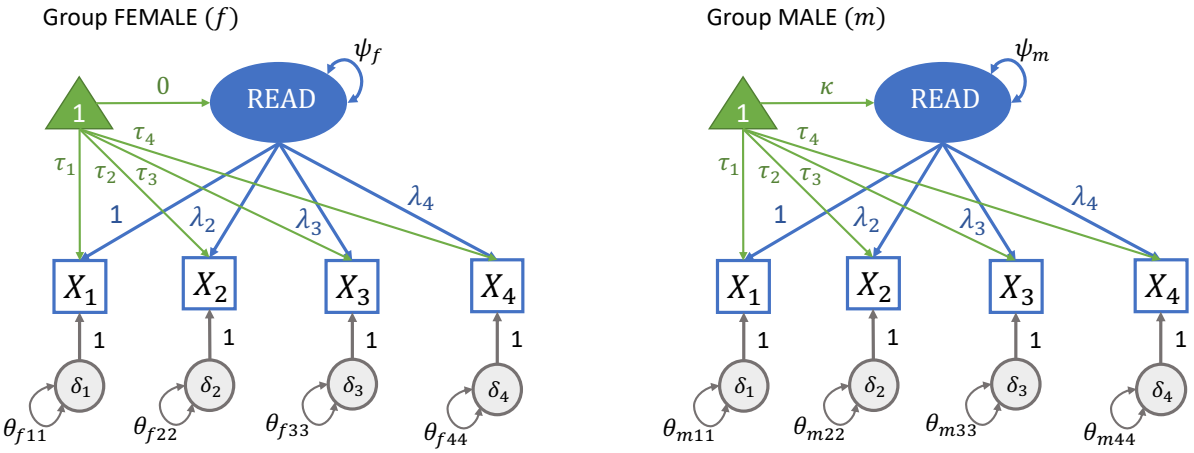
Multiple-Groups CFA (10 credits)

A team of researchers attempted to establish the invariance of a reading comprehension measure (READ) across gender (FEMALE [*f*] and MALE [*m*]).

The researchers have performed multiple-groups confirmatory factor analysis (MGCFA) to evaluate measurement invariance. In this procedure, they sequentially constrained parameters in the measurement model of READ to be equal across gender and evaluated the fit of multiple MGCFAs.

The path diagram below shows the **MGCFA1** model which had appropriate fit to the data and was the researchers' final model:

MGCFA1



a) In the models assuming configural, metric, scalar, and strict invariance, which **model parameters are constrained** to be equal across gender?

Level	Equality constraints
Configural	- Equality of form - Same number of factors across genders, same numbers, and patterns of salient loadings
Metric	- Equality of form (*same as the configural model*) - Equality of the factor loadings
Scalar	- Equality of form (*same as the configural model*) - Equality of the factor loadings (*same as the metric model*) - Equality of the indicator intercepts
Strict	- Equality of form (*same as the configural model*) - Equality of the factor loadings (*same as the metric model*) - Equality of the indicator intercepts (*same as the scalar model*) - Equality of the indicator (residual) error variances

SCORING:

- 1 credit per correct constraints for each level
- **Sub-total: 4 credits**

b) Which **levels of measurement invariance** are met in model MGCFA1?

Level	Yes, this level is met.	No, this level is not met.
Configural	X	☐
Metric	X	☐
Scalar	X	☐
Strict	☐	X

SCORING:

- 0.5 credit per correct answer
- **Sub-total: 2 credits**

c) Complete the code for specifying and estimating the model MGCFA1 in the R package lavaan.

Note: Be attentive to the parameter constraints specified in the path diagram. Do not forget to add a command that indicates a grouping by gender.

Lavaan code for model specification:

```
MGCFA1 <-  '
           READ =~ X1 + X2 + X3 + X4
           '
```

Lavaan code for model estimation:

```
MGCFA1.fit <-  cfa(MGCFA1,
               data = data,
               estimator = ML,
               meanstructure = TRUE,
               group = "GENDER",
               group.equal = c("loadings", "intercepts"))
```

SCORING:

- 1 credit for correctly specifying the model
- 1 credit for adding the grouping option
- 1 credit for specifying "loadings"
- 1 credit for specifying "intercepts"
- **Sub-total: 4 credits**