Final Exam *(Evaluator's Version)*

MAE4101 Measurement Models (Spring Term 2024)

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

Exam date: 06 June 2024, 9:00-13:00 Oslo time, Silurveien 2, NO-0380 Oslo, Room 4D

Welcome to the MAE4101 Measurement Models exam!

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.
- For simplicity, χ^2 is written as **chi-square**.

Abbreviations you can use without explanation throughout the exam:
CFA—Confirmatory factor analysis
EFA—Exploratory factor analysis
MGCFA—Multi-group CFA
SEM—Structural equation model
df—Degrees of freedom

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

Best regards,
Ronny and Tony

Name:	SUGGESTED SOLUTIONS + GRADING

Results

Task	Credits	Max. credits
Concept, Terms, and Procedures		14
Path Model of Intentions to Donate Blood		16
Confirmatory Factor Analyses of Cognitive Skills Tests		10
Mastering Exploratory Factor Analysis		6
Quality of Instruction		12
Transformative TPACK Model		8
TOTAL:		**66**

Expected time on task

- Approximate max. time spent by teachers: 70 minutes
- Estimated max. time for students: 210 minutes

Grading

Grade	Credits range	Percentages correct
A	66-59	100-90%
B	58-53	89-80%
C	52-46	79-70%
D	45-40	69-60%
E	39-33	59-50%
F	32-0	49-0%

Concept, Terms, and Procedures (14 credits)

In the following tasks, you can demonstrate your understanding of the **key concepts, terms, and procedures** concerning structural equation models.

a) Define the following terms briefly.

Term	Definition
Manifest variable	An observed variable that is contained in the data set. Directly observed or measured variable.
Factor loading	Pearson correlation between the factor and the indicator.
Reflective measurement model	A measurement model hypothesizing that a latent variable causes the (co-)variation in the manifest (observed) variables. Examples are common factor models, such as CFA and EFA.
Model-implied covariance matrix	A covariance matrix that contains the variances and covariances that are implied by the model and its parameters. It is compared to the observed covariance matrix.

SCORING:

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

b) In structural equation modeling, a **discrepancy function** is used to estimate the model parameters. Which of the following statements about this discrepancy function are true or false?

Statement	True	False
The discrepancy function maximizes the similarity between the observed and the model-implied distribution of the data.	☐	X
The discrepancy function is used to minimize the probability of the observed data under the model.	☐	X
The discrepancy function minimizes the difference between the observed and model-implied covariance matrices.	X	☐
The discrepancy function is used to confirm the number of factors in a single EFA model.	☐	X

NOTES:

- **Statement 1:** False, because this function minimizes the discrepancy between the model-implied and observed covariance matrices. Moreover, the distribution of the data does not include the covariances, typically.
- **Statement 2:** False, because the probably of the data under the model is not necessarily the main element of this function. Only for maximum-likelihood estimation, a discrepancy function relates directly to the likelihood of the data, but this function is then maximized, not minimized.

- **Statement 3:** True. This is the definition of a discrepancy function in structural equation modeling.
- **Statement 4:** False, because this function only provides information about the fit of one model. It would need multiple models with varying number of factors to use it for determining how many factors may underlie the data.

SCORING:
- 1 credit per correct response
- **Sub-total: 4 credits**

c) **Name two indices or statistics** to evaluate the fit of a structural equation model. Provide the common cut-off value (as per Hu and Bentler, 1999) or the needed result for an acceptable model fit for these indices or statistics.

Note: There is no need to spell out the abbreviations of the fit indices.

Two fit indices or statistics
Two of the following indices or statistics: ▪ Chi-square statistic of the model: insignificant chi-square test suggests a good model fit, because the discrepancy between the observed and model-implied covariance matrices is insignificant. ▪ Comparative Fit Index (CFI): CFI above or equal to .95 ▪ Root Mean Square Error of Approximation (RMSEA): RMSEA below or equal to .06 ▪ Standardized Root Mean Square Residual (SRMR): SRMR below or equal to .08 ▪ Tucker Lewis Index (TLI): TLI above or equal to .95 ▪ …

SCORING:
- 1 credit for the correct name or abbreviation of fit index or statistic each
- 1 credit for the correct provision of the cut-off values or results for an acceptable model fit each
- **Sub-total: 2 credits**

d) Name and describe **two steps (levels) of measurement invariance testing**, except for the separate testing of factor models for the data of each group. When describing the steps, name which parameters are constrained across groups.

Two steps (levels) of measurement invariance testing
Two of the following steps/levels: ▪ *Configural invariance*: same number of factors and loading pattern across groups, all other parameters are freely estimated ▪ *Metric invariance*: configural invariance plus the equality of factor loadings ▪ *Scalar invariance*: metric invariance plus the equality of indicator intercepts ▪ *Strict invariance*: scalar invariance plus the equality of indicator residual (co-) variances

SCORING:

- 1 credit for the correct name of the invariance step/level each
- 1 credit for the correct naming of the constraints each
- **Sub-total: 4 credits**

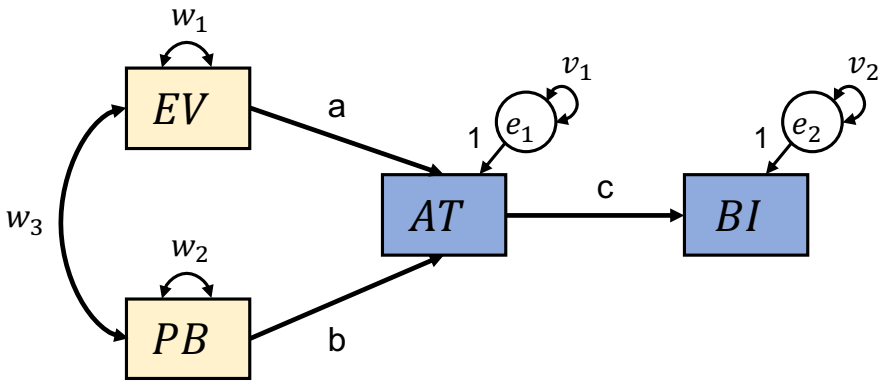
Path Model of Intentions to Donate Blood (16 credits)

In 1988, Bagozzi and Yi published a path model that represented their theory about a person's intentions to donate blood (i.e., behavioral intentions, **BI**). The authors hypothesized that these intentions are explained by the attitudes towards blood-giving (**AT**), which in turn are explained by the expectancy-value (**EV**) and the past behavior (**PB**).

The full path model is shown below (labelled **Model 1**).

Note: This model does not have a mean structure.

Model 1



a) Classify all variables in Model 1 as **endogenous manifest, exogenous manifest,** or **residual variables**. Check the boxes below.

Variable	Endogenous manifest	Exogenous manifest	Residual
EV	☐	X	☐
PB	☐	X	☐
AT	X	☐	☐
BI	X	☐	☐
e_1	☐	☐	X
e_2	☐	☐	X

SCORING:

- 0.5 credit per correct response
- **Sub-total: 3 credits**

b) Provide the **model-specifying equations** of Model 1.

Fill in the model equations here.

Outcome variable	Model equation
$AT =$	$aEV + bPB + e_1$
$BI =$	$cAT + e_2$

SCORING:

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

c) Derive the **model-implied covariance** between EV and AT: $Cov(EV, AT)$.

Note: You may apply either the variance-covariance rules or Wright's path tracing rules. Make sure to provide the steps you have taken to arrive at the end result.

Model-implied covariance	Result
$Cov(EV, AT)$	$= Cov(EV, aEV + bPB + e_1)$ $= Cov(EV, aEV) + Cov(EV, bPB) + Cov(EV, e_1)$ $= aCov(EV, EV) + bCov(EV, PB) + 0$ $= aVar(EV) + bCov(EV, PB)$ $= aw_1 + bw_3$ *Alternative:* Using the path tracing rules, there is one eligible pathway connecting EV and AT via a and the variance of EV. This gives aw_1. There is a second pathway through the covariance between EV and PB and then through b. This gives bw_3.

SCORING:

- 1 credit for the correct end results
- 1 credit for the correct explanation/derivation steps
- **Sub-total: 2 credits**

d) Show that **Model 1 is (over-)identified** by providing the number of available pieces of information (*p*), the number of freely estimated parameters (*q*), and the degrees of freedom of the model (*df*).

Note: Once again, ignore the mean structure.

Elements	Numbers

Number of observed pieces of information (p)	4 observed (manifest) variables with 4 variances and 6 covariances ⇒ p = 10
Number of parameters in the model (q)	To be estimated: 3 path coefficients (a, b, c); 2 residual variances (v_1 and v_2); 2 variances of the exogenous manifest variables (w_1 and w_2); 1 covariance between X_1 and X_2 (w_3) ⇒ q = 8
Degrees of freedom of the model (*df*)	$\mathrm{df}_M = p - q = 10 - 8 = 2 > 0$

SCORING:

- 1 credit per correct response element
- **Sub-total: 3 credits**

e) Model 1 contains two indirect effects. How can you **compute these two indirect effects** from the model parameters?

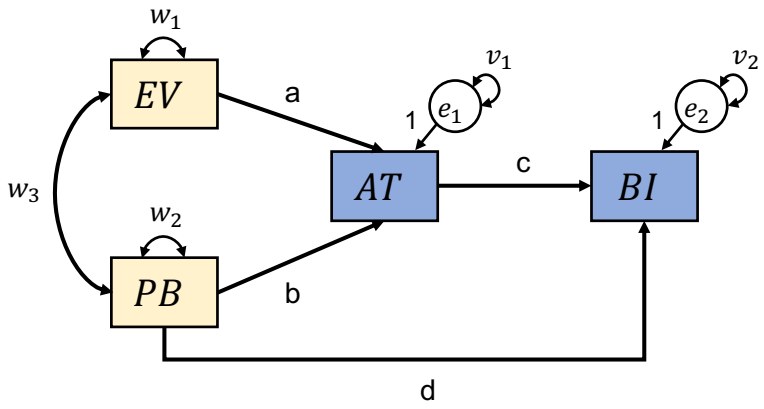
Computation of the two indirect effects
The two indirect effects can be computed as $a \cdot c$ and $b \cdot c$.

SCORING:

- 0.5 credit per correct response element
- **Sub-total: 1 credit**

Bagozzi and Yi (1988) modified their model by adding another path. This new and modified model is shown below (labelled **Model 2**).

Model 2



f) Provide the **lavaan code for specifying Model 2**.

Note: While you may skip the residual variances in the code, do not forget to specify the variances and covariances among the exogenous variables in the model. There is no need to label the path coefficients, and you do not have to include any indirect effect.

Lavaan code for the model specification:

```
Model2 <-  '
           BI ~ AT + PB
           AT ~ EV + PB

           EV ~~ PB
           EV ~~ EV
           PB ~~ PB

           '
```

SCORING:

- 1 credit for each line specifying the endogenous variables BI and AT
- 1 credit for the covariance between EV and PB
- 0.5 credit for each variance of the exogenous variables EV and PB
- **Sub-total: 4 credits**

g) Suppose you have labelled the path coefficients in the lavaan code (with letters as in the path diagram of Model 2). Which of the following code lines do you have to add to **estimate the indirect effect** (ind) of *EV* on *BI* via *AT*? Check the box with the correct code line.

	Code line
☐	`ind :~ a*c`
☐	`ind == a*c`
X	`ind := a*c`
☐	`ind =~ a*c`

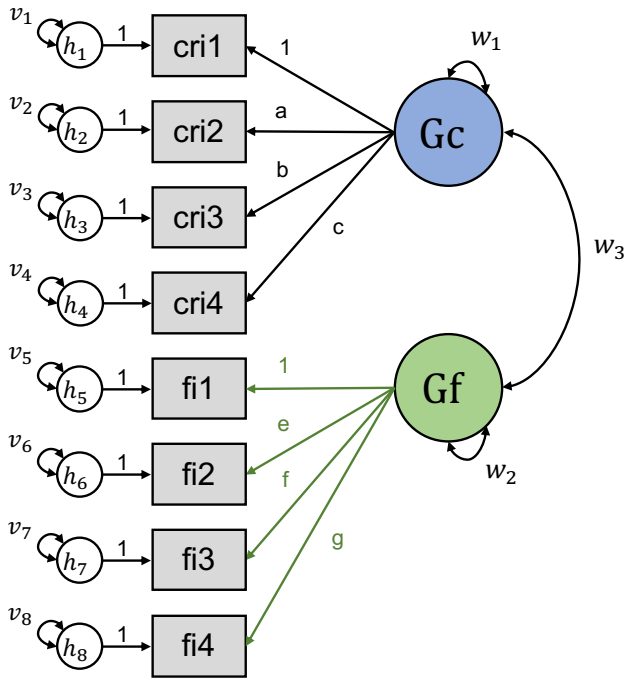
SCORING:

- 1 credit per correct response
- **Sub-total: 1 credit**

Confirmatory Factor Analyses of Cognitive Skills Tests (10 credits)

The figure below shows a CFA model that describes the structure underlying eight tests of cognitive skills. This structure contains two factors Gc and Gf. The model is named "Model CFA1" and does *not* have a mean structure.

Model CFA1



a) Decide whether Model CFA1 represents a **reflective, causal-formative, or composite** measurement model and explain why.

Your decision and explanation
The model is **reflective**. All indicators are considered to be caused by the two latent variables. The latent variables (exogenous variables) are the "predictors" of the observed indicators.

SCORING:

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

b) Verify that Model CFA1 **is over-identified**.

Note: Ignore the mean structure.

Elements	Numbers

Number of observed pieces of information (p)	8 observed (manifest variables, cri1-cri4 and fi1-fi4) with 8 variances and 7+6+5+4+3+2+1 = 28 covariances; alternative computation with 8 manifest variables: 8*(8+1)/2 ⇒ p = 36
Number of parameters in the model (q)	To be estimated: 6 factor loadings (a, b, c, e, f, g) + 8 residual variances (v1-v8) + 2 factor variances + 1 factor covariance ⇒ q = 17
Degrees of freedom of the model (df_M)	$\mathrm{df}_M = p - q = 36 - 17 = 19 > 0$

SCORING:

- 1 credit for the correct number p
- 1 credit for the correct number q
- 1 credit for the correct degrees of freedom of the model df_M
- **Sub-total: 3 credits**

c) Provide the **model equations of the indicator variables cri1 and cri4** in Model CFA1.

Fill in the equations here.

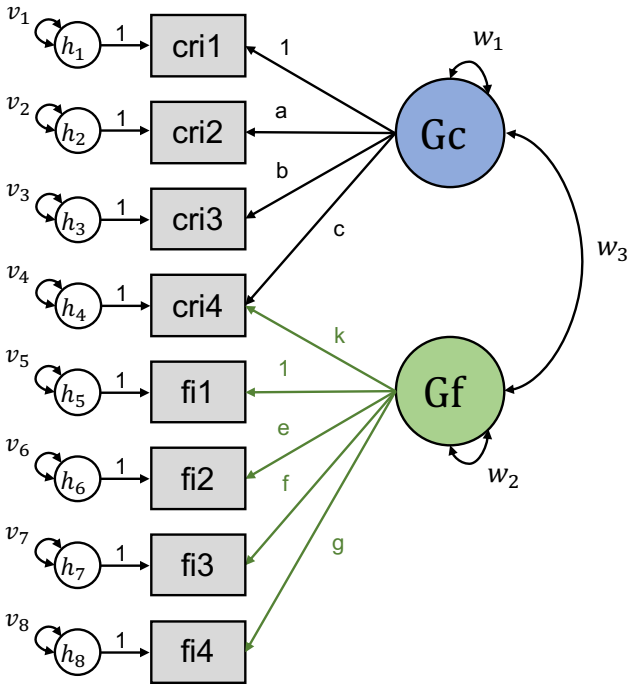
Indicator variables	Model equation
cri1 =	$Gc + h_1$
cri4 =	$c \cdot Gc + h_4$

SCORING:

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

Some researchers have inspected the fit of Model CFA1 and decided to modify this model. The modified model is named “Model CFA2” and is shown below.

Model CFA2



d) Provide the **model equation of the indicator variable cri4** in Model CFA2.

Fill in the equation here.

Indicator variables	Model equation
cri4 =	$c \cdot Gc + k \cdot Gf + h_4$

SCORING:

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

e) Describe **one method** to decide which of the two models (Models CFA1 and CFA2) may fit the data better. What is the criterion you would apply to decide for one model?

Provide your answer here.

Method of model comparison
One of the following methods: ▪ Model comparison with chi-square difference testing: If the outcome is statistically significant, then the model with the cross-loading (loading k) may fit the data better. ▪ Model comparison with differences in fit indices: The model with better indices is preferred.

- Estimate Model CFA2 and perform a parameter test of the factor loading *k*: Test whether *k* is statistically significantly different from zero.
- Model comparison with differences in information criteria: The model with lower values of the AIC and/or BIC is preferred.

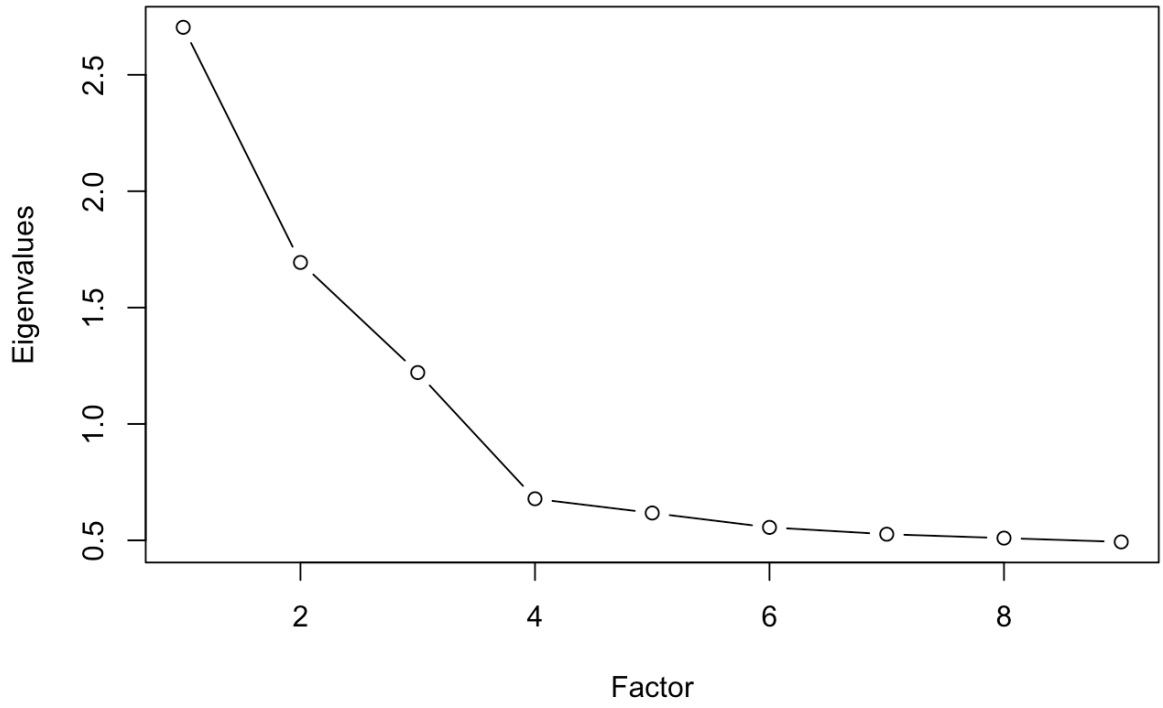
SCORING:

- 1 credit for correct naming/description of the method
- 1 credit for the correct criterion
- **Sub-total: 2 credits**

Mastering Exploratory Factor Analysis (6 credits)

A team of students in the Master of Assessment, Measurement, and Evaluation program developed a questionnaire scale that contained 9 items (named Y1-Y9) and collected data from 750 young adults in Norway. Without any prior knowledge about the factor structure of that scale, the students conducted an exploratory factor analysis.

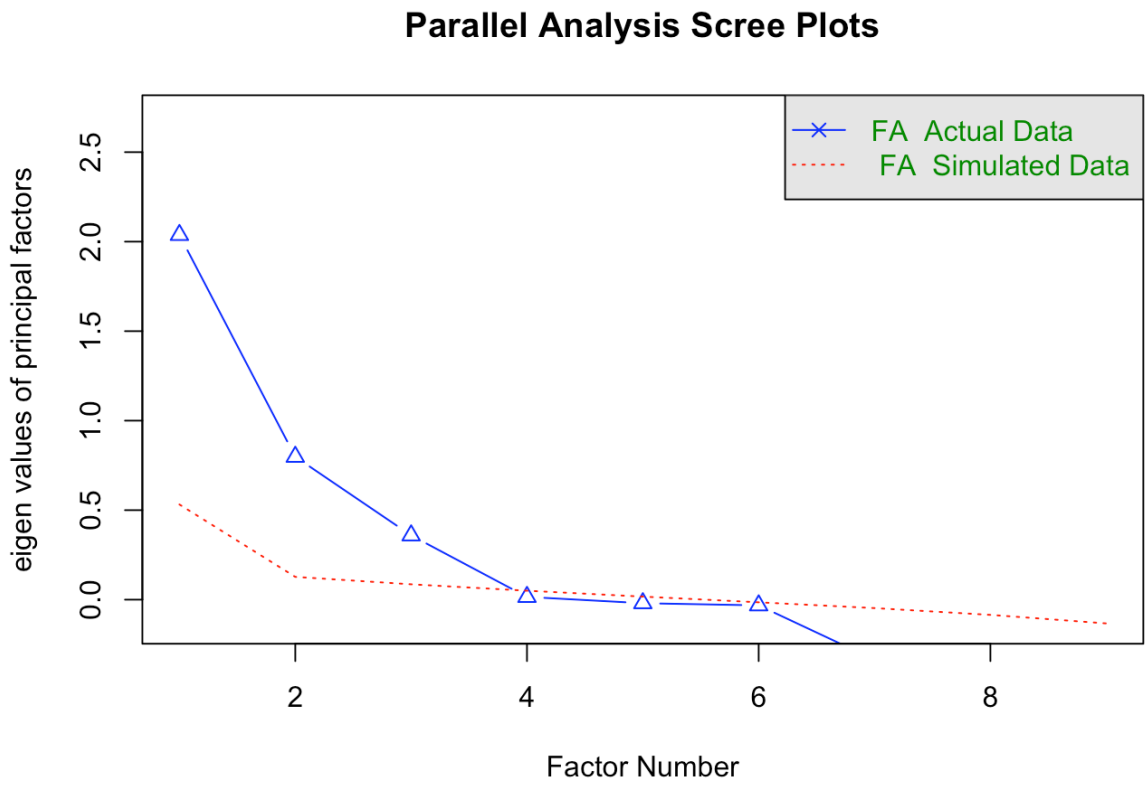
a) The students concluded that **3 factors underlie the 9 items (Y1-Y9)**. Below, you will find the outputs of the tests the students have performed to arrive at this conclusion. Explain how these tests indicate that 3 factors exist.



Your explanation
At least one of the two explanations: ▪ The screeplot shows a knee, until the eigenvalue curve flattens out from 4 factors onwards. ▪ There are three eigenvalues above 1 (Kaiser criterion).

SCORING:

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



Your explanation
Three eigenvalues of the data (blue line, generated from a factor analysis) are larger than the eigenvalues of the simulated data (red line).

SCORING:

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

Model fit statistic/index	EFA with one factor	EFA with two factors	EFA with three factors
Chi-square value	501.6	160.7	11.9
df	27	19	12
p-value	< .001	< .001	.46
BIC (sample-adjusted)	408.6	95.2	-29.5

Explanation
The model with three factors fits the data best (e.g., insignificant chi-square test and lowest BIC).

SCORING:

- 1 credit for the correct response

- **Sub-total: 1 credit**

b) The students were unsure which factor rotation, oblique or orthogonal, is suitable to describe the factor structure. Below, you will find the output from an EFA with 3 factors and an oblique (Oblimin) and, respectively, an orthogonal (Varimax) factor rotation.

Oblique factor rotation (Oblimin)

```
## Standardized loadings (pattern matrix) based upon correlation matrix
##      ML2   ML3   ML1   h2    u2 com
## y1 -0.03  0.71  0.03 0.50 0.501 1.0
## y2  0.02  0.62 -0.01 0.38 0.617 1.0
## y3  0.02  0.61 -0.04 0.38 0.623 1.0
## y4  0.70  0.02 -0.01 0.49 0.512 1.0
## y5  0.70  0.00  0.01 0.49 0.509 1.0
## y6  0.57 -0.02  0.05 0.34 0.663 1.0
## y7  0.67 -0.01 -0.02 0.44 0.563 1.0
## y8  0.13  0.08  0.43 0.25 0.746 1.3
## y9 -0.01 -0.01  1.00 1.00 0.005 1.0

##  With factor correlations of
##      ML2  ML3  ML1
## ML2 1.00 0.20 0.27
## ML3 0.20 1.00 0.16
## ML1 0.27 0.16 1.00


##      RMSEA      lower      upper confidence
## 0.00000000 0.00000000 0.04106741 0.95000000
```

Orthogonal factor rotation (Varimax)

```
## Standardized loadings (pattern matrix) based upon correlation matrix
##     ML2  ML3  ML1   h2    u2 com
## y1 0.04 0.70 0.10 0.50 0.501 1.0
## y2 0.07 0.61 0.05 0.38 0.617 1.0
## y3 0.07 0.61 0.03 0.38 0.623 1.0
## y4 0.69 0.09 0.10 0.49 0.512 1.1
## y5 0.69 0.07 0.11 0.49 0.509 1.1
## y6 0.56 0.04 0.14 0.34 0.663 1.1
## y7 0.65 0.05 0.08 0.44 0.563 1.0
## y8 0.19 0.11 0.45 0.25 0.746 1.5
## y9 0.11 0.05 0.99 1.00 0.005 1.0


##      RMSEA      lower      upper confidence
## 0.00000000 0.00000000 0.04106741 0.95000000
```

Decide which factor rotation is more suitable to describe the data.

	Factor rotation method
X	Oblique rotation
☐	Orthogonal rotation

SCORING:

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

Explain your decision for your choice of a factor rotation method.

Explanation
The factor correlations are not zero.

SCORING:

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

c) Using the output of the rotated factor solution you have chosen, indicate which items (Y1-Y9) belong to which factor (Factor 1, Factor 2, or Factor 3). In other words, indicate which items are mainly explained by which of the three factors. Please check the respective boxes.

Item	Factor 1	Factor 2	Factor 3
Y1	X	☐	☐
Y2	X	☐	☐
Y3	X	☐	☐
Y4	☐	X	☐
Y5	☐	X	☐
Y6	☐	X	☐
Y7	☐	X	☐
Y8	☐	☐	X
Y9	☐	☐	X

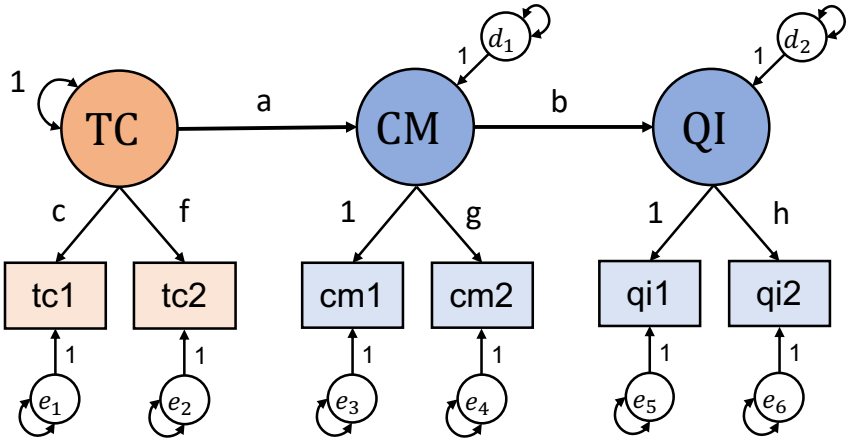
SCORING:

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

Quality of Instruction (12 credits)

Teachers play a vital role in students' learning. A group of researchers tested a model that describes the relations between **teacher competence (TC)**, **classroom management (CM)**, and their **quality of instruction (QI)**. The three constructs are measured by two observed variables each. This model of instructional quality (labelled as MIQ) is shown below.

Model of Instructional Quality (MIQ)



Note: This model does not contain a mean structure.

a) This model contains a measurement part and a structural part.

Provide the **equations that specify the measurement model part**.

Measurement part of the model:
Fill in the model equations here.

Outcome variable	Model equation
$tc1 =$	$c \cdot TC + e_1$
$tc2 =$	$f \cdot TC + e_2$
$cm1 =$	$CM + e_3$
$cm2 =$	$g \cdot CM + e_4$
$qi1 =$	$QI + e_5$
$qi2 =$	$h \cdot QI + e_6$

SCORING:

- 0.5 credit per correct equation
- **Sub-total: 3 credits**

Provide the **equations that specify the measurement model part**.
Structural part of the model:

Fill in the model equations here.

Outcome variable	Model equation
$CM =$	$a \cdot TC + d_1$
$QI =$	$b \cdot CM + d_2$

SCORING:

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

b) Provide the lavaan code that specifies this model. The model is labelled "MIQ" (Model of Instructional Quality).

Note: There is no need to label the path coefficients, and you do not have to estimate any indirect effect.

Lavaan code for the model specification:

```
MIQ <-  '
        # Measurement part
        TC =~ tc1+tc2
        CM =~ cm1+cm2
        QI =~ qi1+qi2

        # Structural part
        CM ~ TC
        QI ~ CM

        '
```

SCORING:

- 1 credit for each line specifying the measurement models
- 1 credit for each line specifying the structural model
- **Sub-total: 5 credits**

c) Formulate **one research question** that could be answered by estimating this model.

Note: Avoid formulating questions that can be answered with a simple yes or no.

Research question
▪ To what extent are teacher competence and the quality of instruction related via classroom management? ▪ What is the indirect effect of teacher competence on quality of instruction via classroom management? ▪ To what extent does a full mediation of the relation between TC and QI via CM hold?

SCORING:
- 1 credit for a correct research question
- **Sub-total: 1 credit**

d) What is the model-implied covariance between the variables *tc2* and *cm2*?

Note: You can provide the end result with deriving it.

Model-implied covariance:
$$Cov(tc2, cm2) = f \cdot a \cdot g$$ How come? ▪ *Wright's path tracing rules*: There is only one pathway connecting tc2 and cm2 via the direct effect of TC on CM (labelled a). All parameters along the way are multiplied. ▪ *Variance-covariance rules*: $Cov(tc2, cm2) = Cov(f \cdot TC + e_2, g \cdot CM + e_4)$ $= Cov(f \cdot TC, g \cdot CM) + Cov(f \cdot TC, e_4) + Cov(e_2, g \cdot CM) + Cov(e_2, e_4)$ $= fg \cdot Cov(TC, CM) + f \cdot Cov(TC, e_4) + g \cdot Cov(e_2, CM) + Cov(e_2, e_4) = fg \cdot Cov(TC, a \cdot TC + d_1) + 0 + 0 + 0$ $= fag \cdot Cov(TC, TC) + fg \cdot Cov(TC, d_1) = fag \cdot 1 + 0 = fag$

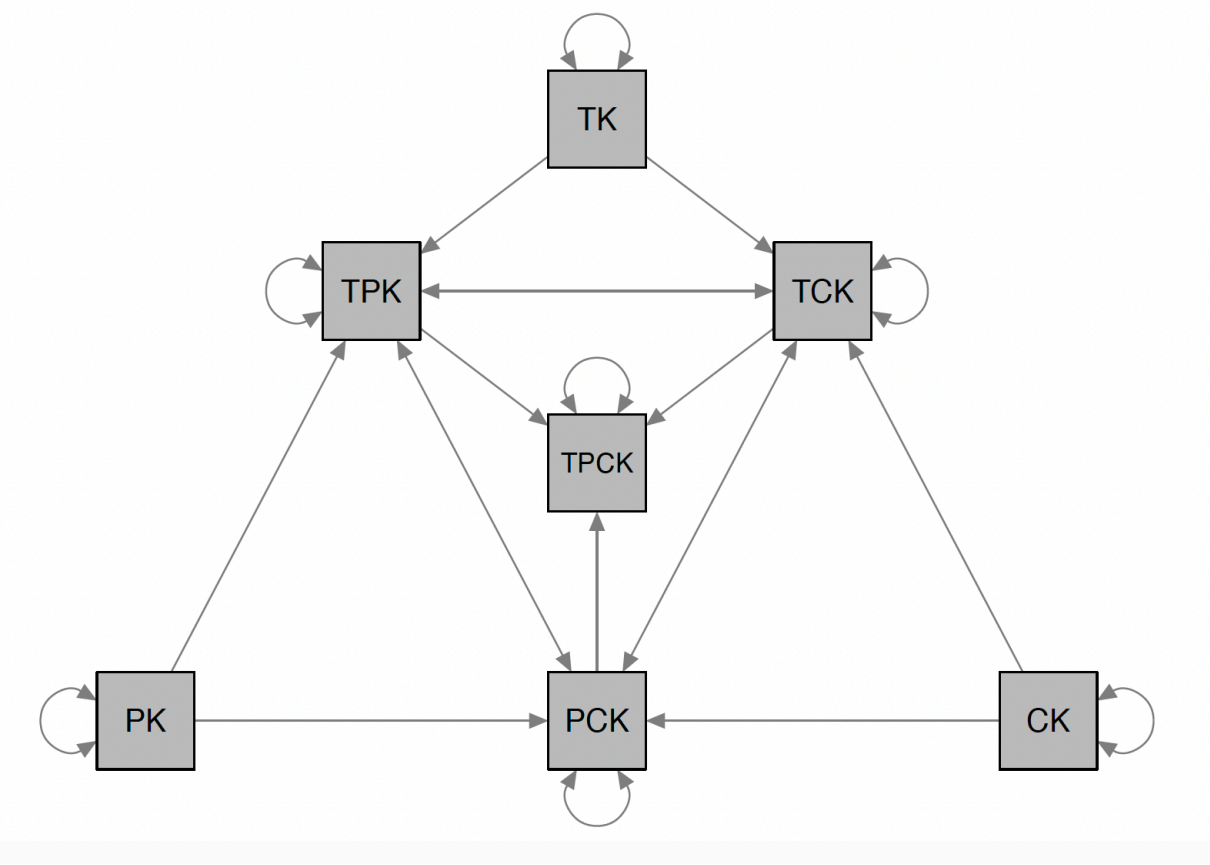
SCORING:
- 1 credit for the correct result
- **Sub-total: 1 credit**

Transformative TPACK Model (8 credits)

Schmid et al. (2021) proposed a path model that describes the relations between seven components of teachers' Technological and Pedagogical Content Knowledge (TPACK). This so-called "transformative model" contains the following teacher variables:

- CK: Content knowledge
- PK: Pedagogical knowledge
- PCK: Pedagogical and content knowledge
- TK: Technological knowledge
- TCK: Technological content knowledge
- TPK: Technological pedagogical knowledge
- TPCK: Technological pedagogical content knowledge

The model is shown in the figure below.



Note: This path diagram depicts the variances of exogenous variables and residuals in the same way.

Provide the lavaan code that specifies this model. The model is labelled "TPACK".

Note: You do not need to include in the code the variances of residuals, but their covariances. Moreover, include the variances of the exogenous variables.

Lavaan code for the model specification:

```
TPACK <-  '
          TPCK ~ TCK + TPK + PCK
          TPK ~ TK + PK
          TCK ~ TK + CK
          PCK ~ CK + PK

          PCK ~~ TPK
          PCK ~~ TCK
          TPK ~~ TCK

          CK ~~ CK
          PK ~~ PK
          TK ~~ TK

          '
```

SCORING:

- 1 credit for each line specifying the endogenous variables TPCK, TPK, TCK, and PCK
- 1 credit for each of the residual covariances (i.e., PCK-TPK, PCK-TCK, and TPK-TCK)
- 1 credit for all variances of the exogenous variables CK, PK, and TK
- **Sub-total: 8 credits**