
Final Exam (*Student Version*)

Measurement Models Course (Spring 2021)

MAE4101 Measurement Models

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

Release date: 11 June 2021, 1:00 PM, Oslo time

Submission deadline: 18 June 2021, 3:00 PM, Oslo time

Welcome to the MAE4101 Measurement Models home 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. *explain/derive* something).
- You may **simplify subscripts** or Greek symbols wherever appropriate (e.g., Y_1 instead of Y_1 , λ_1 instead of λ_1).
- Keep your explanations and descriptions **to the point**.
- **Partial credits** will be given.
- Starting with this home exam, you declare that you will work on the tasks **without any help of others**.
- You will be presented with **three tasks**, each of which is accompanied by a dataset. Download the data sets here: <https://cemouio.shinyapps.io/mae4101/>.
- **Scheduling:** We recommend spending at least one day on each of the tasks.
- Submit the **responses to the tasks in the Inspera system**.
- Submit the **R code for each of the three main tasks**. Create one R code file for each of the three tasks and label them "PathModels.R", "FactorAnalysis.R", and "SEM.R".
- **Help!** If needed, please contact us via email or the exam forum on the course Canvas page, and we shall get back quickly.

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

Best regards,

Denise, Jelena, Kseniia, Jarl, and Ronny

Results

Task	Credits	Max. credits
Path Models		40
Factor Analysis: Conspiracy Theory Scale (CTS)		40
Structural Equation Models		40
TOTAL:		120

Grading

Grade	Credits range	% Correct threshold
A	108–120	90 %
B	96–107	80 %
C	84–95	70 %
D	72–83	60 %
E	60–71	50 %
F	0–60	< 50 %

Data Sets

Using your personal access code, please download the three datasets for this exam from the following page: <https://cemouio.shinyapps.io/mae4101/>

Declaration

Before starting the exam, confirm the following:

Declaration	Yes, I confirm.
I hereby confirm that I will work on this home exam without the help of others.	<input type="checkbox"/>
I will not exchange any responses, material, or ideas with the other examinees during the exam.	<input type="checkbox"/>
I confirm that I am the sole author of the responses provided in this exam.	<input type="checkbox"/>
I am aware of the possible consequences any form of cheating during this exam may have (https://www.uio.no/english/studies/examinations/cheating/).	<input type="checkbox"/>

Data Check

Check the **PathModel data set** you have downloaded by indicating its sample size and the mean of the variable X_1 below.

(a) Which **sample size** does your PathModel data set have?

500 625 750

(b) What is the **mean of the variable X_1** in the PathModel data set?

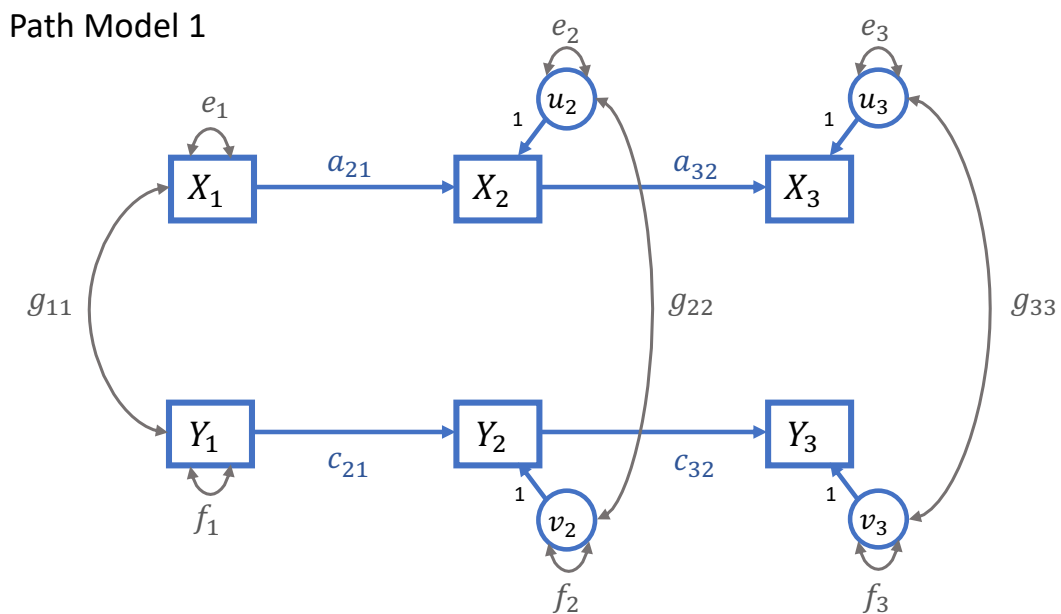
-0.045 0.033 0.038

Path Models

A group of researchers followed university students from their first year until their last year of a Bachelor program. At three time points, these students reported their interest in the subject they were studying and their life satisfaction. The variables indicating students' **interest** are labelled X_1 , X_2 , and X_3 , and the variables indicating students' **life satisfaction** are labelled Y_1 , Y_2 , and Y_3 .

The researchers hypothesized that the interest and life satisfaction variables are related as shown in **Path Model 1**.

Note: This model does not include a mean structure.



(P1) Specify **Path Model 1**, providing the model equations for the dependent variables and the covariance matrix of the residuals.

Fill in the model equations here.

Dependent variables	Model equation

Covariance matrix of the residuals:

--

- (P2) Using the information about the model specification, **derive** the following two elements of the model-implied covariance matrix underlying **Path Model 1**: $Var(X_2)$ and $Cov(X_1, Y_2)$.

Fill in the equations here.

Independent variables	Model equation

- (P3) Specify and estimate **Path Model 1 in lavaan**, using the **PathModel data set** and maximum-likelihood estimation. Provide the R code for the model specification and the estimation.

*Note: You may read the data file using the `read.csv2()` function in R. Decimals are indicated by commas, and the data file contains the variable names in the first row.
Example: `read.csv2(file = "PathModel.csv", header = TRUE, dec = ",")`*

Model specification in lavaan:

```
PathModel1 <- ‘
```

Model estimation in lavaan:

--

- (P4) Request a **model summary** which contains the standardized model parameters, the variance explanations of all dependent variables (R^2), and the model fit indices. Provide the R code for calling this model summary.

Model summary in lavaan:

--

(P5) The chi-square test statistic of Path Model 1 is accompanied by **8 degrees of freedom**. Explain why there are 8 degrees of freedom in this model.

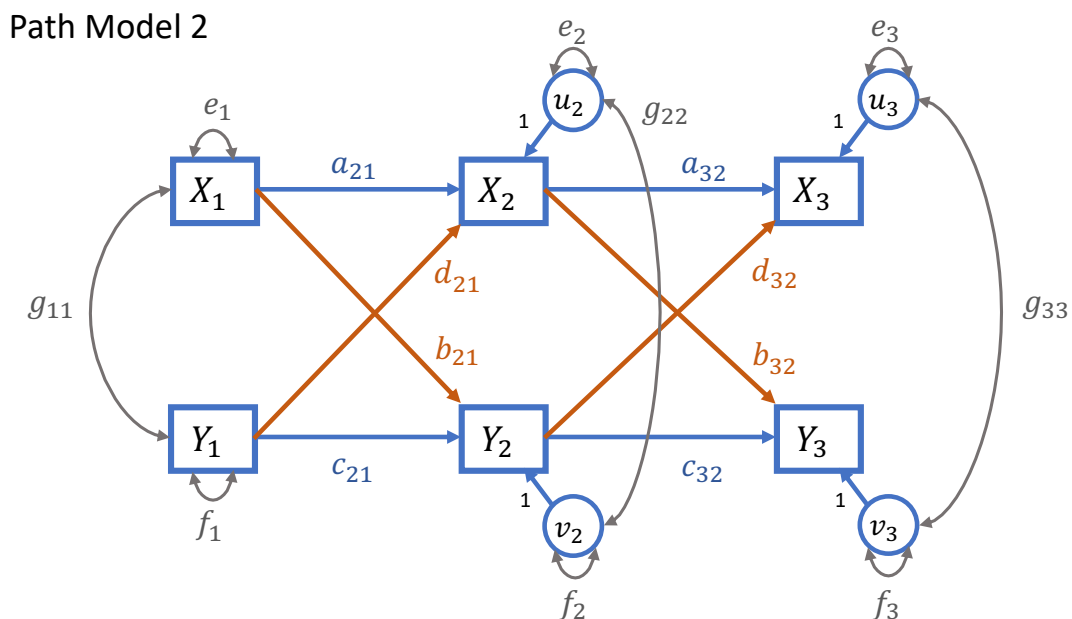
Model identification:

(P6) Determine whether **Path Model 1** exhibits a **reasonably good fit** (Hu & Bentler, 1999), focusing on the chi-square statistic, CFI, RMSEA, and SRMR. Provide the values of these indices with three digits after the comma.

Model fit evaluation:

Fit index	Values
Chi-square statistic	
CFI	
RMSEA	
SRMR	

Having evaluated the fit of Path Model 1, the researchers decided to revise their model. The revised model is labelled **Path Model 2** and is shown below.



Path Model 2 contains so-called “cross-lagged paths” (i.e., b - and d -paths). These paths connect the interest and life satisfaction variables at different time points.

(P7) Specify and estimate **Path Model 2 in lavaan**, using the same data and estimation procedure as for Path Model 1. Provide the R code for the **model specification**. Compute the **indirect effects** for $X_1 \rightarrow X_2 \rightarrow X_3$ and $Y_1 \rightarrow Y_2 \rightarrow Y_3$.

Model specification in lavaan:

PathModel2 <- ‘

,

(P8) Provide and describe the evidence supporting the **reasonably good fit of Path Model 2**, focusing on the chi-square statistic, CFI, RMSEA, and SRMR. Provide the values of these indices with three digits after the comma.

Model fit evaluation:

Fit index	Values
Chi-square statistic	
CFI	
RMSEA	
SRMR	

(P9) Provide the unstandardized estimates and their standard errors of the **two indirect effects** with three digits after the comma and interpret them.

Indirect effects:

Parameters	Estimate	SE
Indirect effect (indX) $X_1 \rightarrow X_2 \rightarrow X_3$		
Indirect effect (indY) $Y_1 \rightarrow Y_2 \rightarrow Y_3$		

File upload

Please upload the **R code** file for the path model task and label it **“PathModels.R”**.

File upload
(max. 2 GB):

Conspiracy Theory Scale (CTS)

Exploratory Factor Analysis (EFA)

A group of researchers wants to evaluate the dimensionality of the 15-item “Conspiracy Theory Scale” (CTS) for a sample of $N = 2400$ students. Half of the students were sampled from Statistics and half from Psychology courses.

The following table describes the 15 items (variables) in the dataset (**Q1-Q15**). Participants rated these items on a 5-point Likert-type scale: *1 = Definitely not true, 2 = Probably not true, 3 = Not sure/cannot decide, 4 = Probably true, 5 = Definitely true.*

Variable	Description
Q1	The government is involved in the murder of innocent citizens and/or well-known public figures, and keeps this a secret.
Q2	The power held by heads of state is second to that of small unknown groups who really control world politics.
Q3	Secret organizations communicate with extraterrestrials, but keep this fact from the public.
Q4	The spread of certain viruses and/or diseases is the result of the deliberate, concealed efforts of some organization.
Q5	Groups of scientists manipulate, fabricate, or suppress evidence in order to deceive the public.
Q6	The government permits or perpetrates acts of terrorism on its own soil, disguising its involvement.
Q7	A small, secret group of people is responsible for making all major decisions, such as going to war.
Q8	Evidence of alien contact is being concealed from the public.
Q9	Technology with mind-control capacities is used on people without their knowledge.
Q10	New and advanced technology which would harm current industry is being suppressed.
Q11	The government uses people as patsies to hide its involvement in criminal activity.
Q12	Certain significant events have been the result of the activity of a small group who secretly manipulate world events.
Q13	Some UFO sightings and rumors are planned or staged in order to distract the public from real alien contact.
Q14	Experiments involving new drugs or technologies are routinely carried out on the public without their knowledge or consent.
Q15	A lot of important information is deliberately concealed from the public out of self-interest.

The **FactorAnalysis data set** contains the raw data of these variables. In the following, you will be asked to perform analyses with these variables. Please make sure to perform these analyses and submit the respective code.

(F1) **Correlation:** Obtain the Pearson correlation between variables Q3 and Q8 with two digits after the comma. Interpret this value and indicate the assumptions behind this correlation (i.e., level of measurement, type of relationship).

Measure	Value	Interpretation
Correlation between Q3 and Q8		
Assumptions		

(F2) **Factor selection:** Using the **principal axis factor method**, how many factors can be extracted from the variables Q1-Q15? Apply the Kaiser-Guttman rule and parallel analysis, provide the respective eigenvalues for both methods, and explain your decision for the number of factors.

Method	Number of factors	Eigenvalues	Explanation
Kaiser-Guttman rule			
Parallel analysis			

(F3) **Factor rotation:** A researcher wants to compare the results of the **maximum-likelihood factor extraction** for three factors of the CTS without rotation to the extraction with the Promax rotation. Using the Promax and no rotation in two separate analyses, provide the following estimates and their interpretations. Explain briefly why one would choose the results for the rotated solution.

Element	Without rotation	With Promax rotation	Interpretation
Factor loading of Q9 for the first factor			

Communality of Q12			
Residual variance of Q10			
Explanation for the use of the rotated solution			

Confirmatory Factor Analysis (CFA)

(F4) **Model specification and estimation:** Using the CTS data, run a CFA model with maximum likelihood estimation with two factors. This model is named “**CFA1**” and contains two factors: “GOV” and “UFO”. GOV is measured by the variables Q1, Q6, and Q11. UFO is measured by the variables Q3, Q8, and Q13. This model does not include a mean structure.

Provide the R code for the specification and the estimation of model CFA1 in lavaan.

Model specification in lavaan:

--

Model estimation in lavaan:

--

(F5) **Model identification:** A researcher claims that model **CFA1 is over-identified**. Test his or her claim, indicating all the elements that support your conclusion (i.e., number of available pieces of information, model parameters, degrees of freedom).
Note: Assume that the factor variances are freely estimated.

Model identification:

--

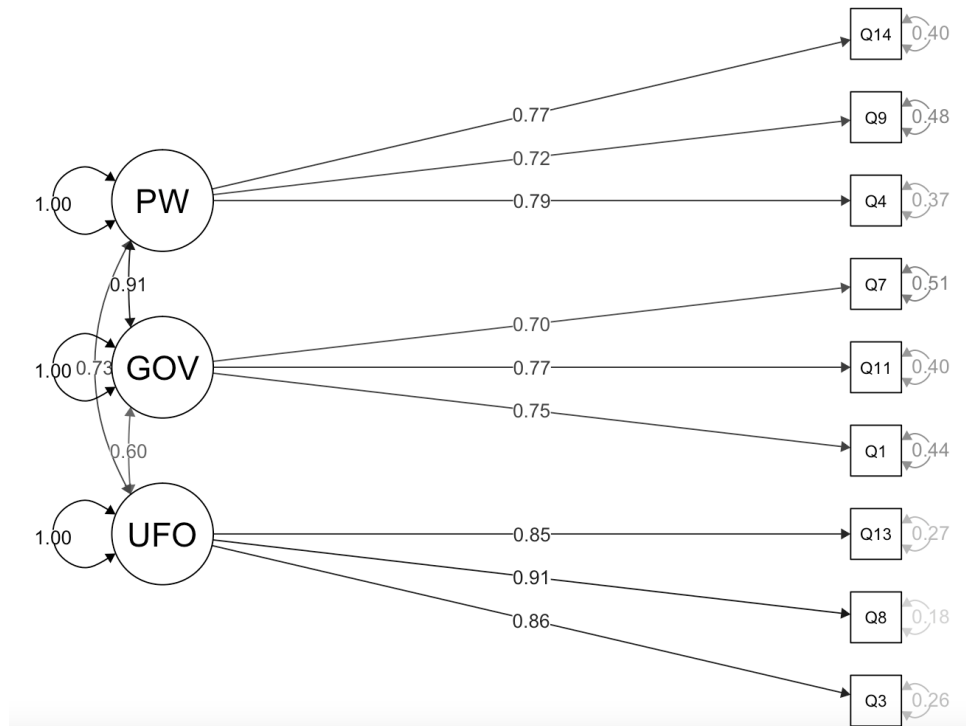
(F6) **Goodness-of-fit indices:** Provide the RMSEA and CFI of model **CFA1**. Evaluate these two indices and decide whether model CFA1 fits the data.

Element	Value	Interpretation
RMSEA		
CFI		

(F7) **Areas of local misfit:** Provide and interpret the following indices.

Element	Value	Interpretation
Modification indices Q3 \sim Q13		
Modification indices UFO \sim Q6		
Standardized residual for Q1 and Q11		

(F8) **Path diagram and model specification:** Another researchers has included more CTS variables and specified a CFA model with three factors (model **CFA2**). Model **CFA2** is shown in the figure below.



Using the information from the path diagram, extract the following **standardized parameter estimates of model CFA2** from the path diagram and interpret them.

Element	Estimate	Explanation
PW= \sim Q9		
GOV \sim UFO		
Residual variance of Q11		

File upload

Please upload the R code file for the factor analysis task and label it " FactorAnalysis.R ".	File upload (max. 2 GB):	
---	-----------------------------	--

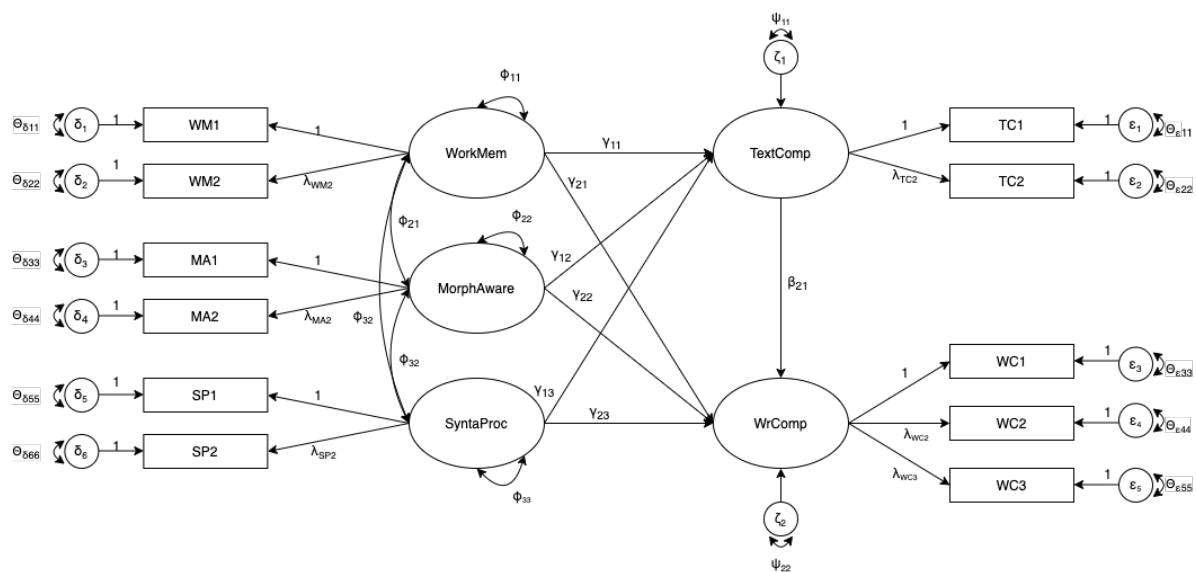
Structural Equation Models

A team of researchers wants to investigate the predictors of individual differences in written composition (**WrComp**) for grade 4 and grade 5 students. The **predictors** are:

- Morphological awareness (**MorphAware**)
- Syntactic processing (**SyntaProc**)
- Working memory (**WorkMem**)

The researchers further hypothesize that the observed predicted relations are **mediated by** text comprehension (**TextComp**). They posited a structural equation model, named “**SEM1**”, which is shown below. Notice that this model does not contain a mean structure.

Model SEM1:



(S1) Specify the **measurement models** for the latent variables **WorkMem**, **SyntaProc**, and **WrComp** by providing their model equations. Fill in the equations here.

Manifest indicators	Model equation

(S5) Request a **SEM1 model summary** which contains the standardized model parameters, the variance explanations of all dependent variables (R^2), and the model fit indices. Provide the R code for calling this model summary.

Model summary in lavaan:

(S6) Determine whether **SEM1** exhibits **acceptable fit**. Inspect the chi-square statistic, CFI, RMSEA, and SRMR. Provide the values of these indices with three digits after the comma. *Note: As an intermediate check, make certain your chi-square statistic has 34 degrees of freedom.*

Model fit evaluation:

Fit index	Value
Chi-square statistic	
CFI	
RMSEA	
SRMR	

(S7) In their study, the researchers tested the **mediation effect of text comprehension (TextComp)**. Describe **two procedures** which could be employed to test for the presence of mediation.

Mediation Testing

(S8) Name two procedures to deal with possible **multicollinearity** in model **SEM1**.

Multicollinearity procedures

The study includes students from two grades: grade 4 and grade 5. This brings the questions of whether **the measures and their structural relations are the same across groups**. Hence, the researchers would like to test for invariance.

To complete the following tasks, please use the grouping variable provided in the 12th column of the raw data. This variable “group” is a factor with two levels.

(S9) Using the specification of model **SEM1** from *task S3*, extend this model to a multi-group structural equation model and test the following **levels of measurement invariance**:

- a. Configural invariance
- b. Metric invariance
- c. Regression (structural relations) invariance

Provide the R code for each model estimation.

Model estimation in lavaan:

(S10) Using the estimated models from *task S9*, compare the invariance models via chi-square difference tests. Specifically, compare the configural model to the metric model, and the metric model to the structural invariance model. Provide the results of these comparisons and the respective chi-square values below.

Note: You may use anova()-function to perform the tests.

Measurement invariance level	Df	Chi2	Chi2 diff	Df diff	p
Configural			-	-	-
Metric					
Regression					

Explain briefly which level(s) of measurement invariance is (are) supported.

Explanation of invariance level

File upload

Please upload the **R code** file for the structural equation modeling task and label it "**SEM.R**".

File upload
(max. 2 GB):

Notes