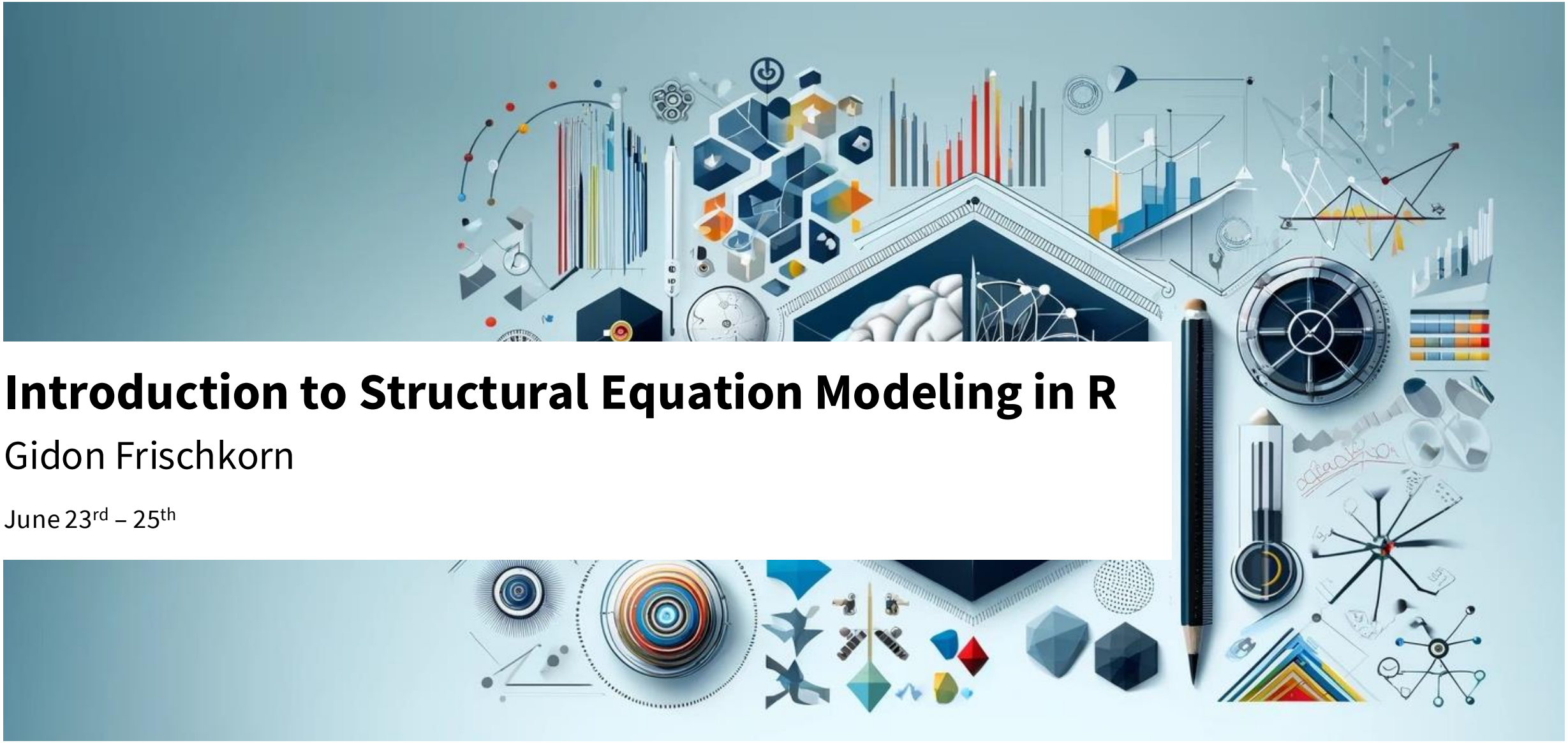
Universität Zürich[UZH]

Psychologisches Institut



Introduction to Structural Equation Modeling in R

Gidon Frischkorn

June 23rd – 25th

Who am I? Who are you?

Academic Career

B.Sc. & M.Sc. Psychology Heidelberg

PostDoc & SNF Ambizione Fellow Zurich

Ph.D. Cognitive Individual Differences Heidelberg

Research Focus

Individual Differences in cognitive processes

Advanced Bayesian Methods

Mathematical Modelling of Cognition

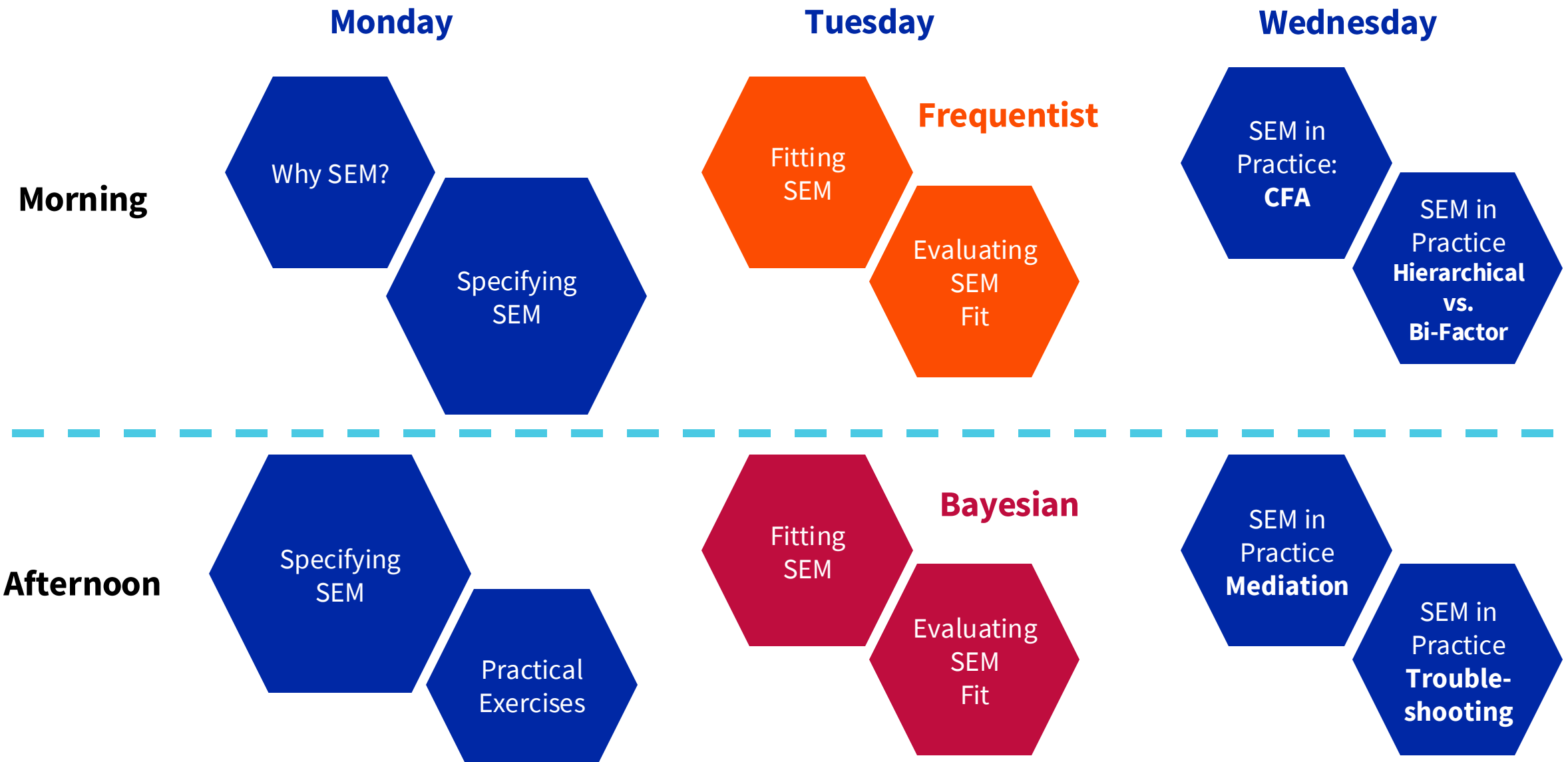
Who are you?

What do you hope to learn?

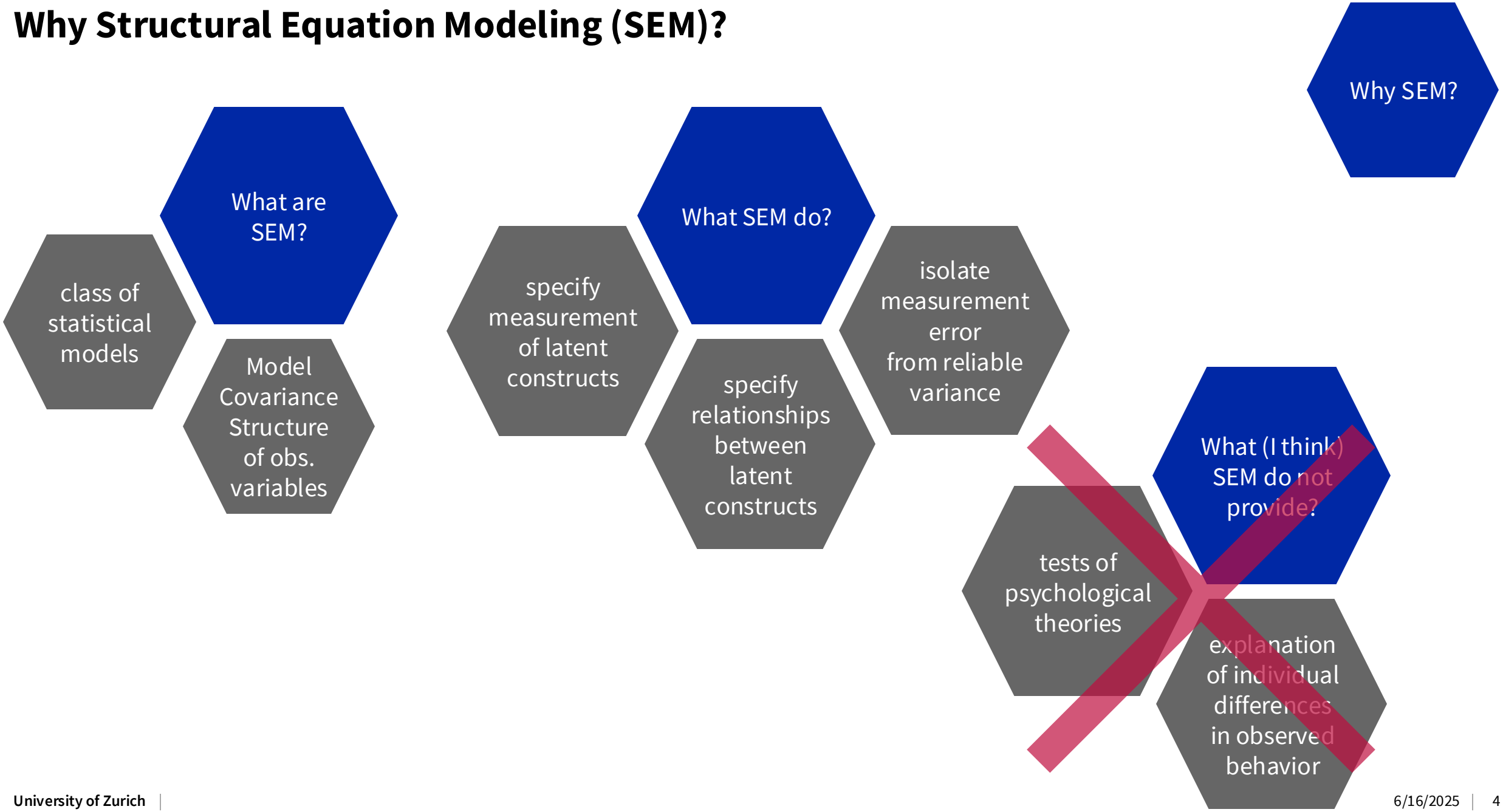
How would you rate your Statistics & R level?

What challenges do you anticipate?

SEM Workshop: Overview

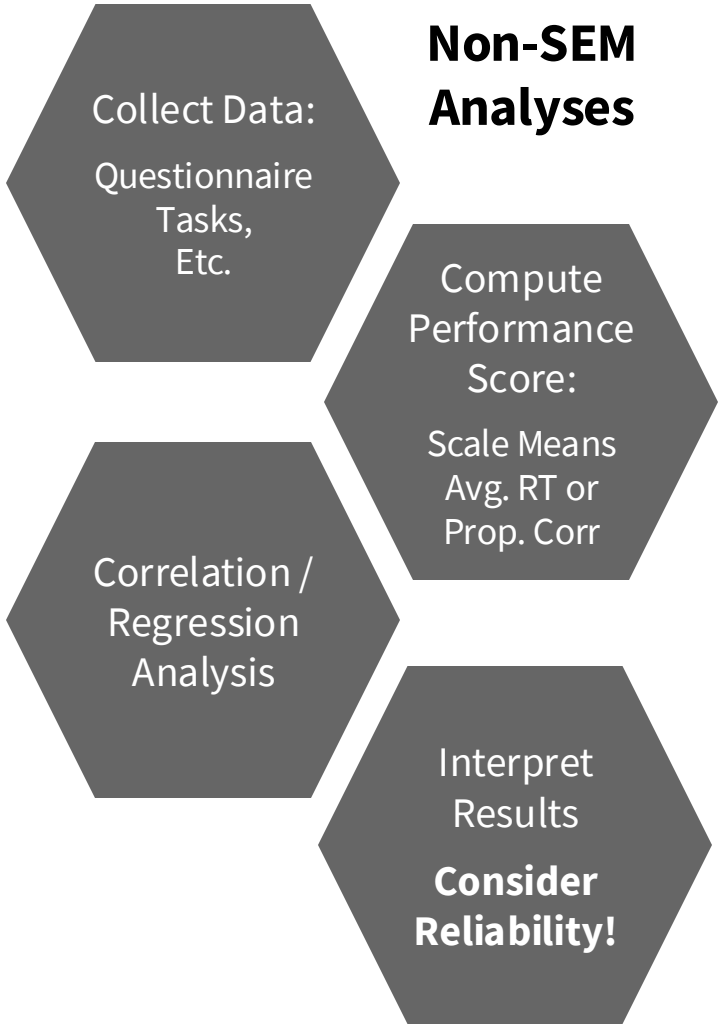


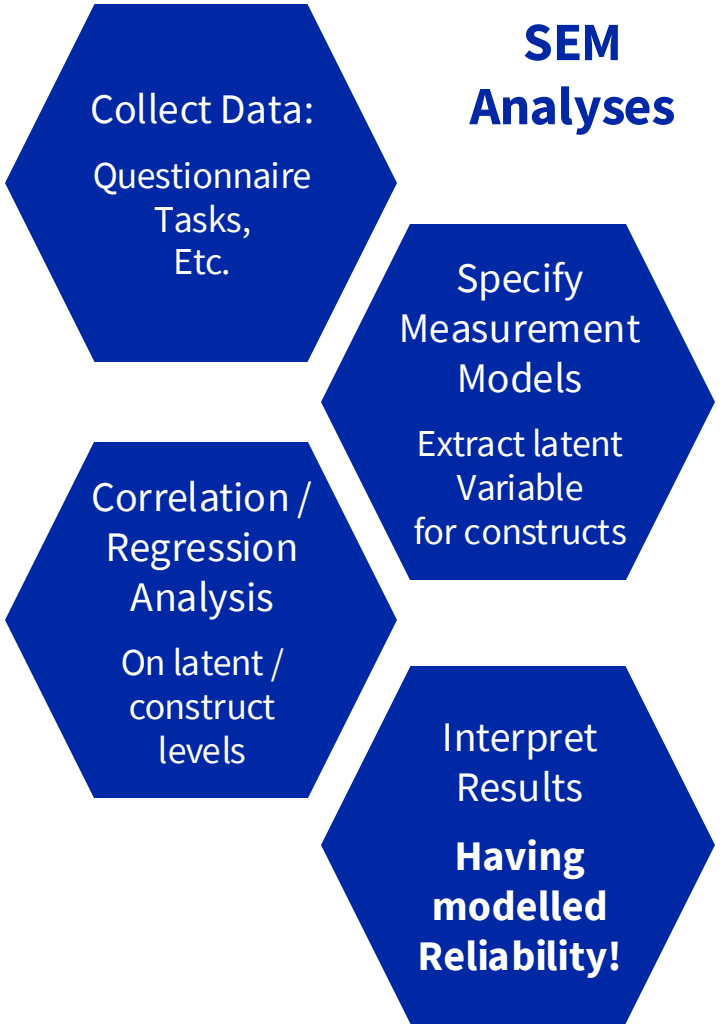
Why Structural Equation Modeling (SEM)?



Why Structural Equation Modeling (SEM)?

Why SEM?



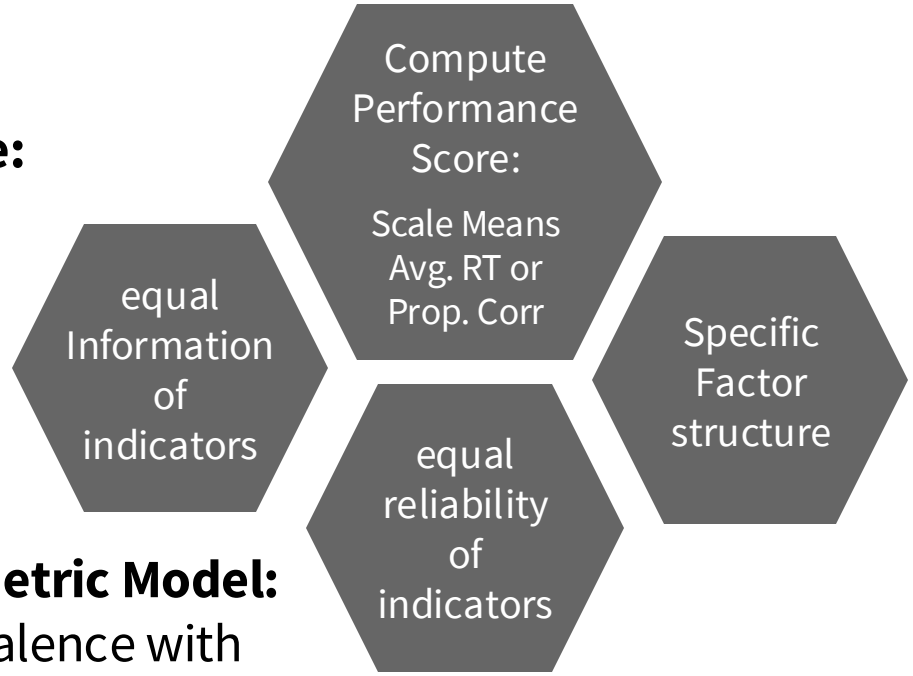


Why Structural Equation Modeling (SEM)?

Why SEM?

Non-SEM Analyses

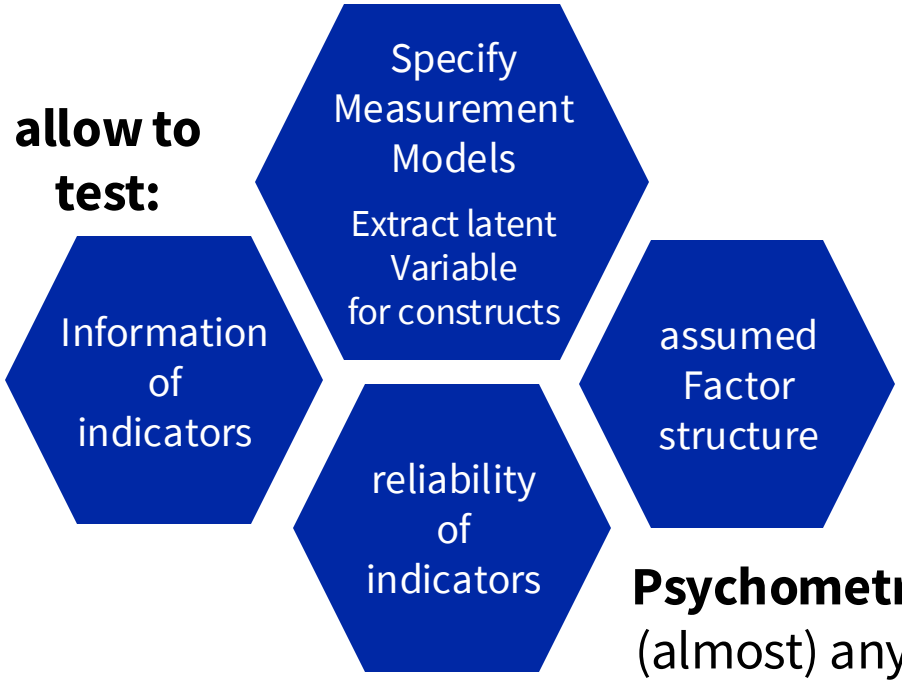
assume:



Psychometric Model:
τ-equivalence with equal error variances

SEM Analyses

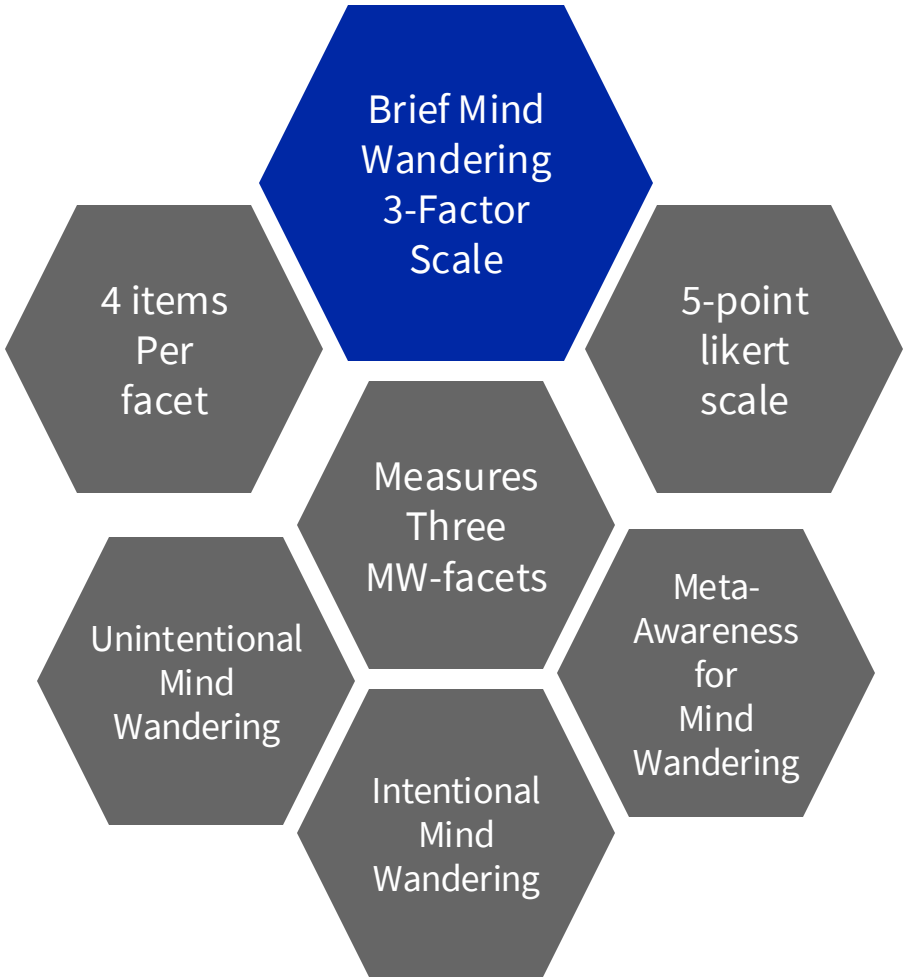
allow to test:



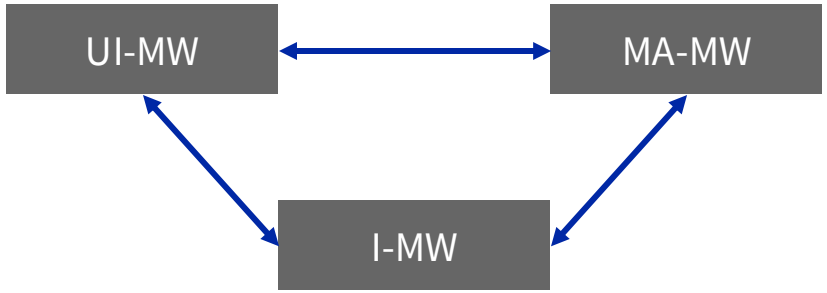
Psychometric Model:
(almost) anything can be specified & tested!

Why Structural Equation Modeling (SEM)?

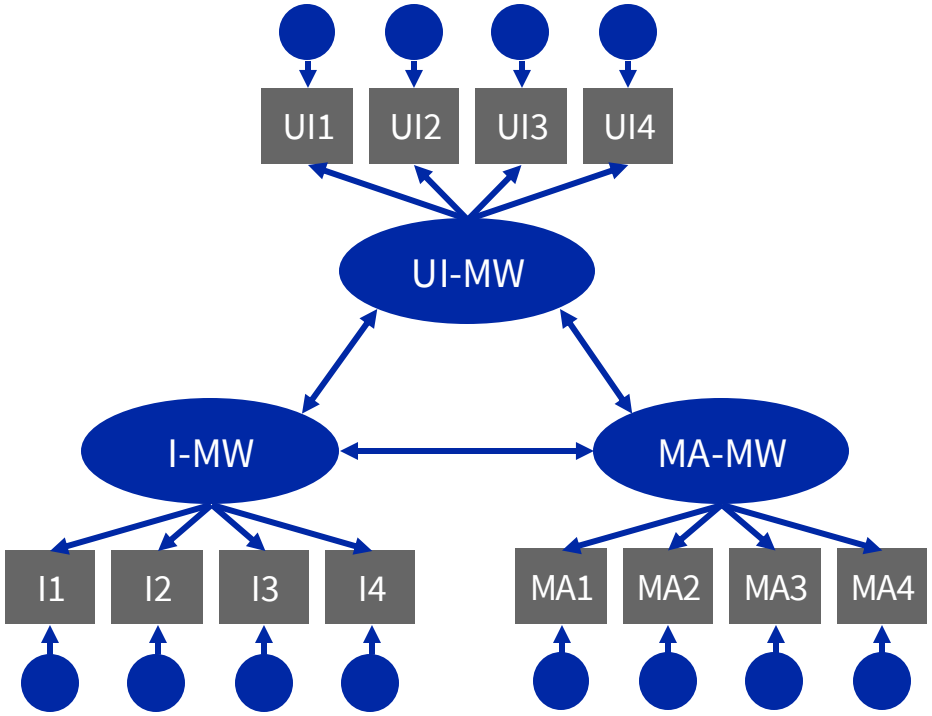
Practical Example



Non-SEM Analyses



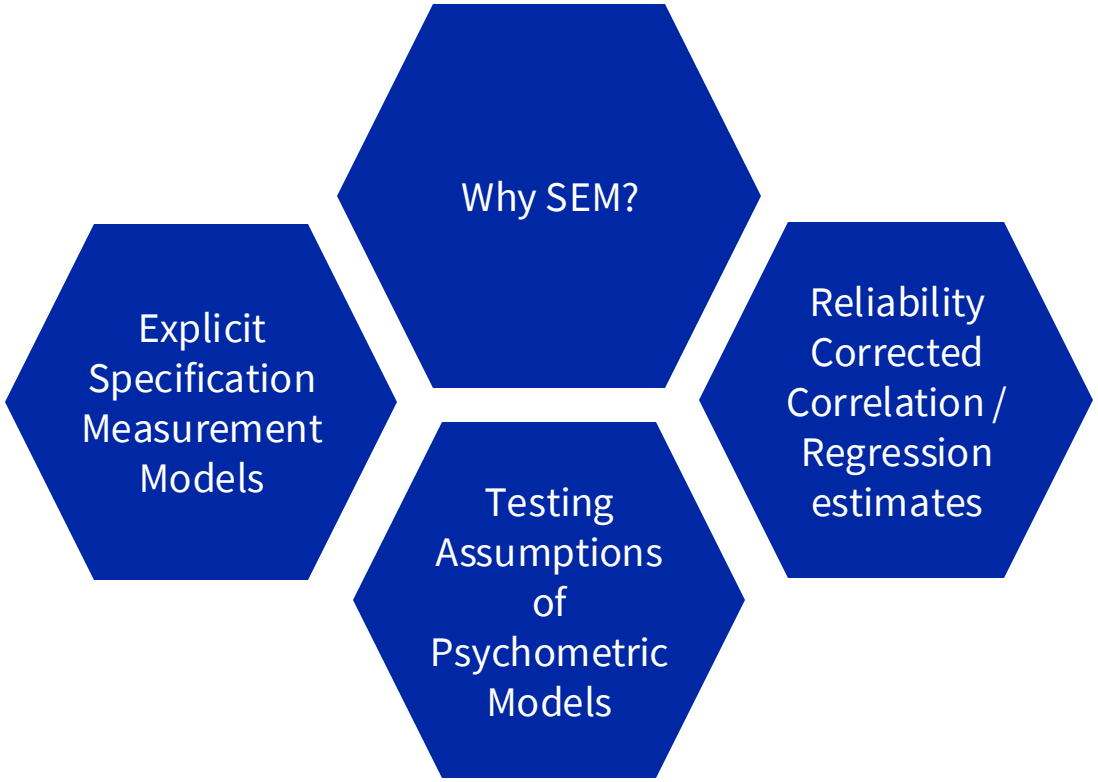
SEM Analyses



Why SEM?

Why Structural Equation Modeling (SEM)?

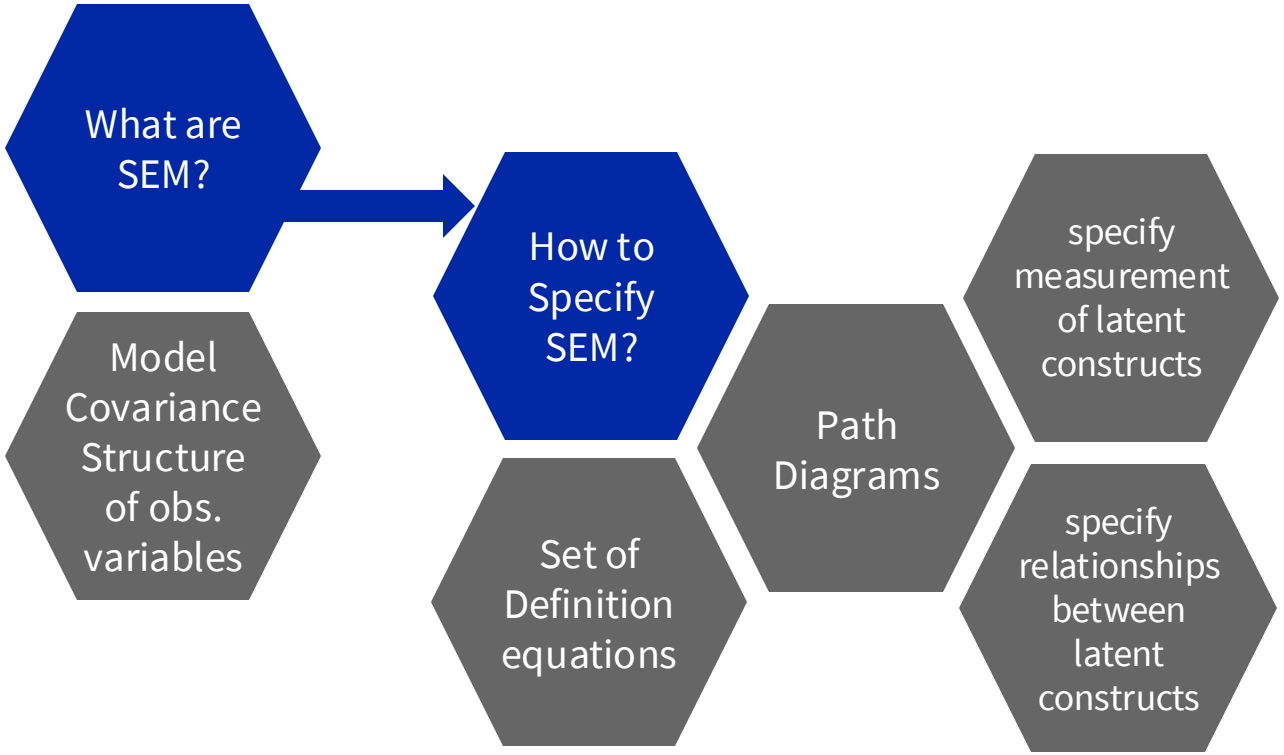
Summary

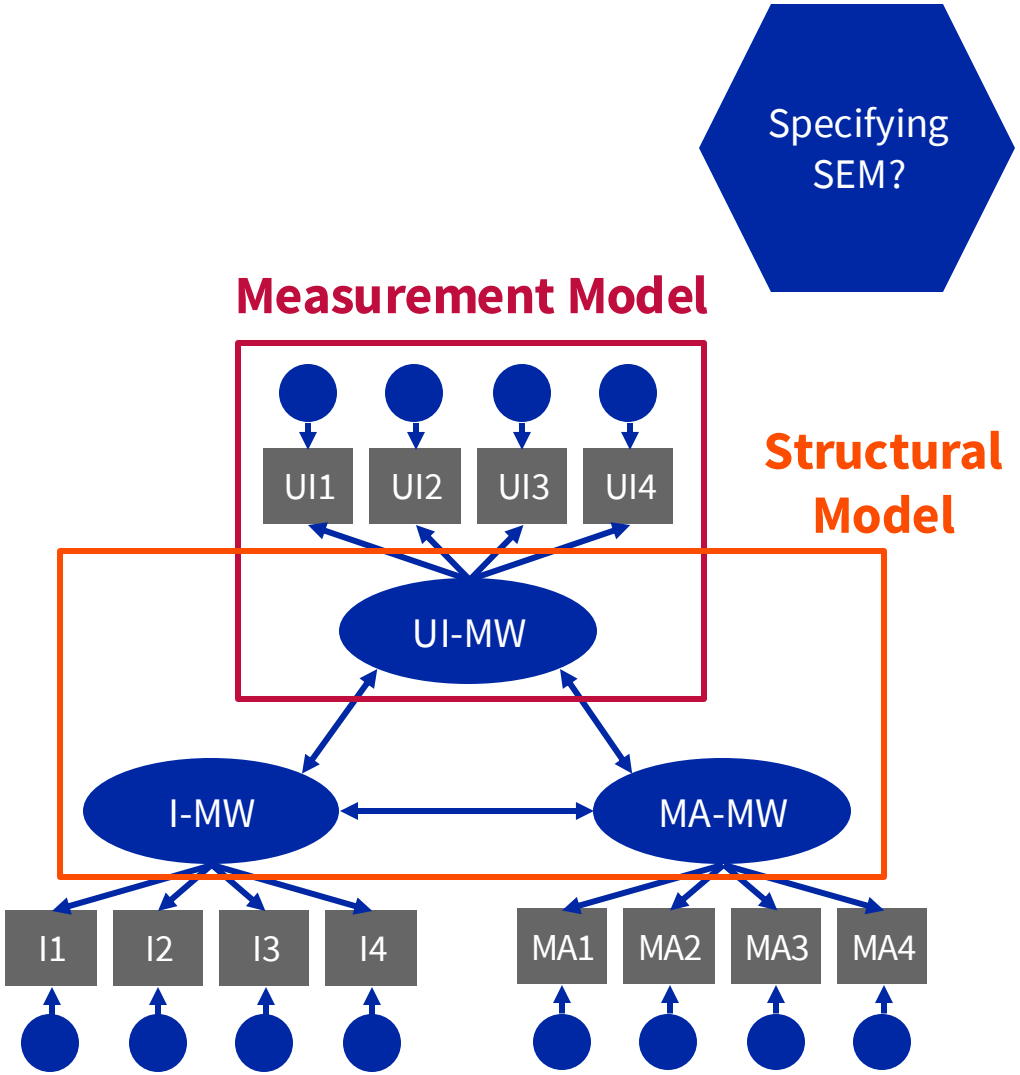


What are your questions so far?

Model Specification

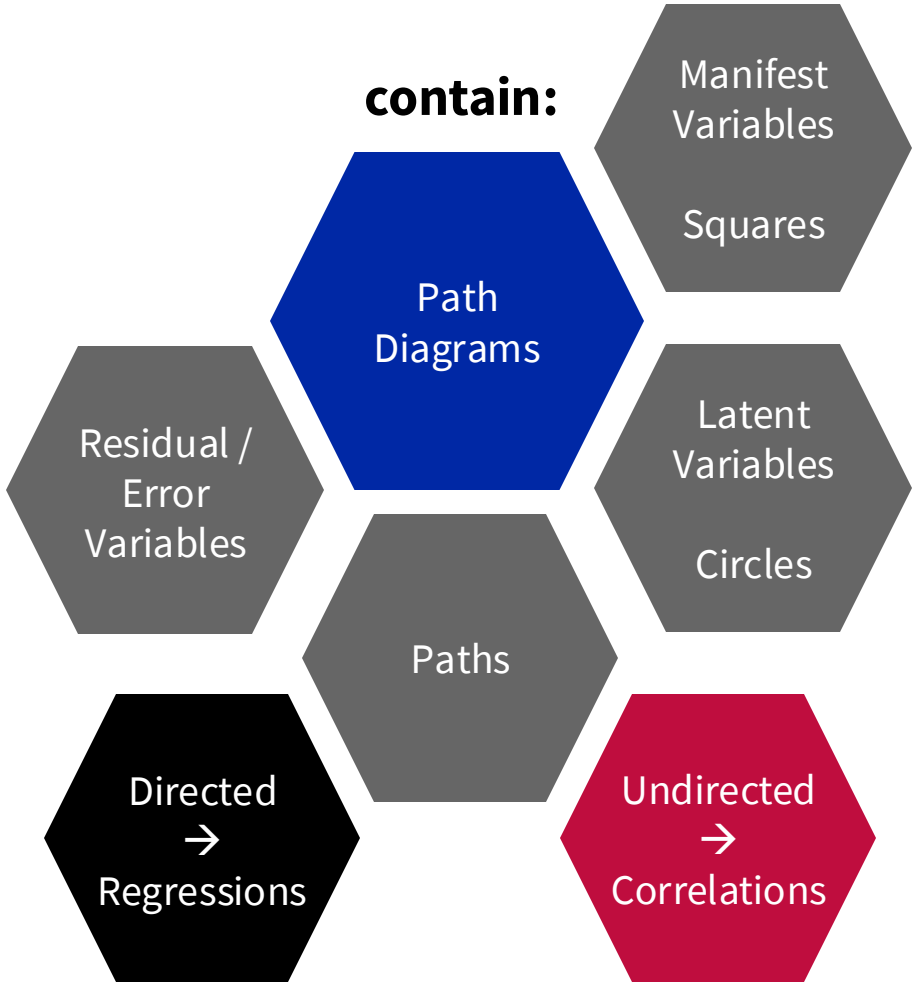
Path diagrams & terms



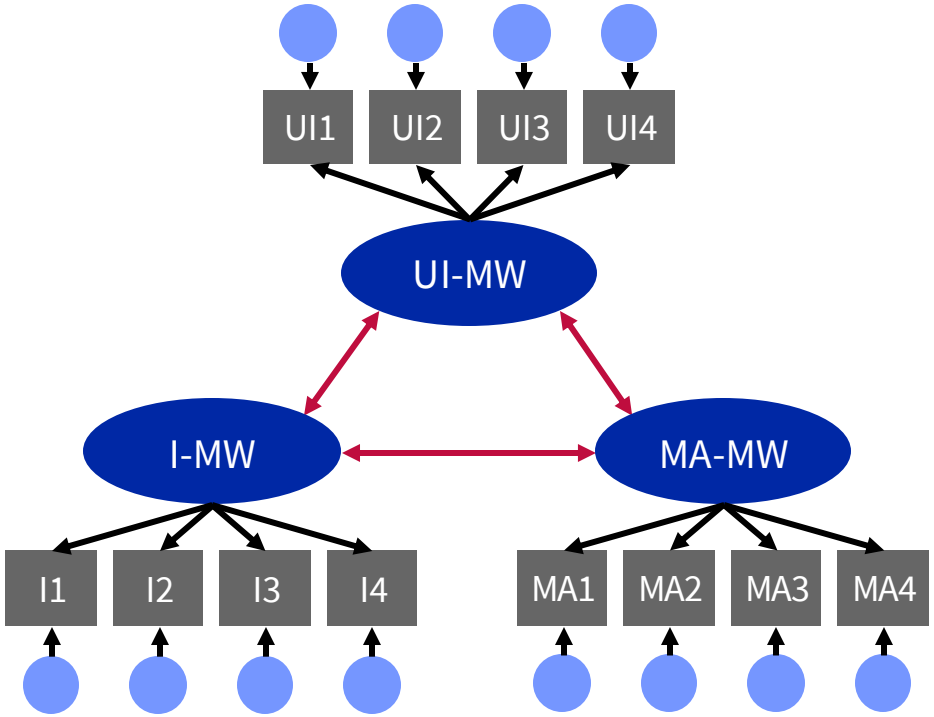


Model Specification

Path diagrams & terms



Specifying SEM?



Model Specification
Path diagrams & terms

Specifying SEM?

Parameters:

- *Fixed* = set to specific values
- *Free* = estimated by data
- *Constrained* = required to satisfy numerical conditions

Path Diagrams

Variances

Exogeneous Variables

Co-Variances

Endogenous Variables

Paths

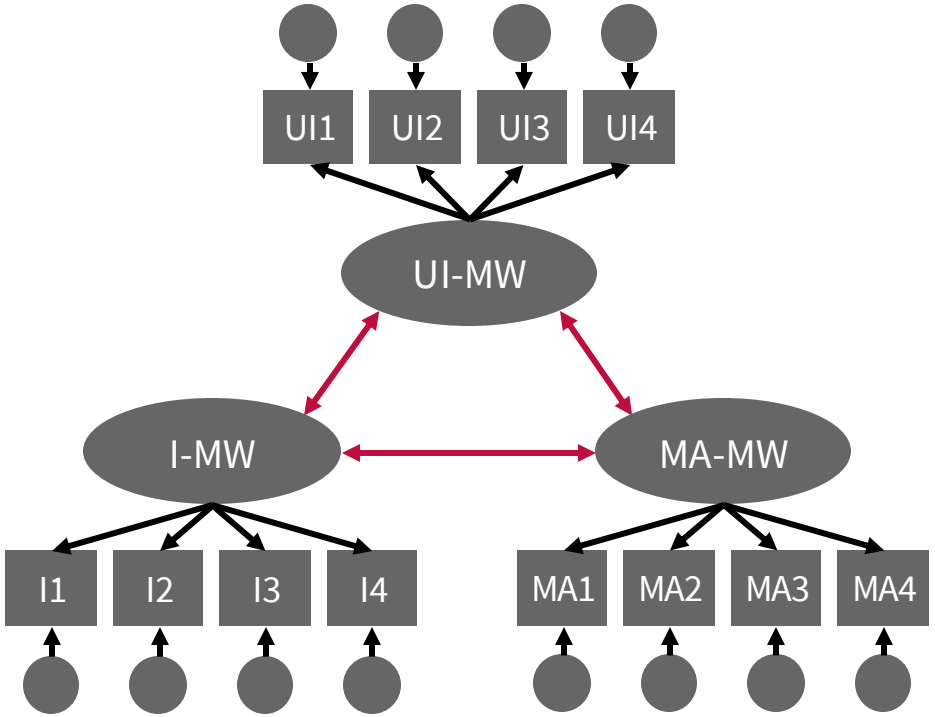
Directed → Regressions

Undirected → Correlations

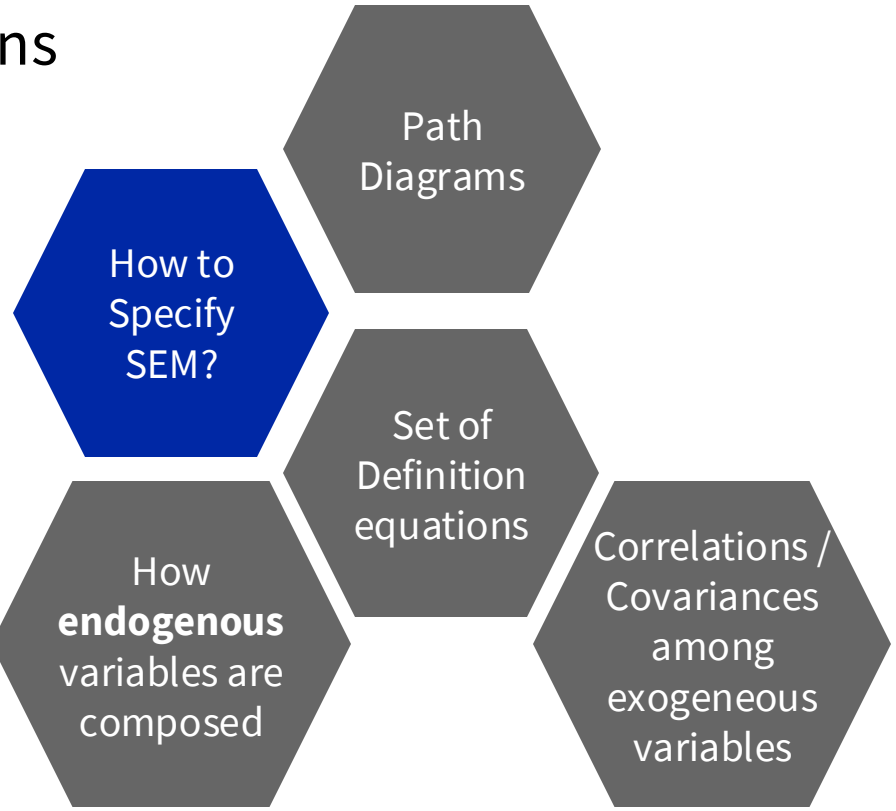
No Value = fixed to one

Path Coefficients

No Arrow = fixed to zero



Model Specification
Definition equations



Technically:

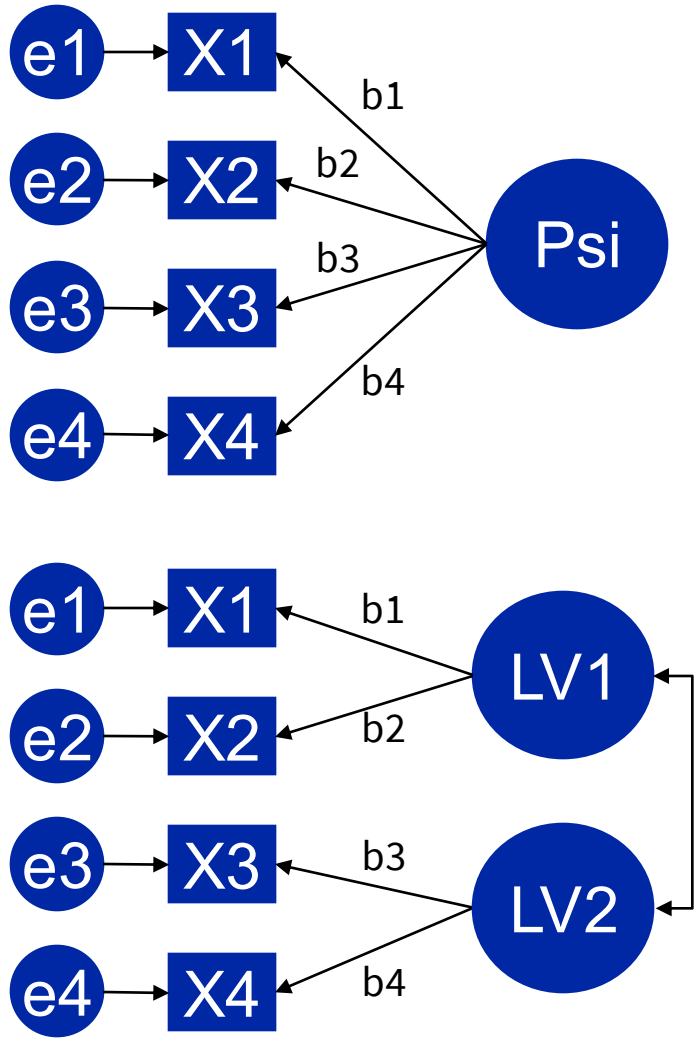
Specify all correlations that are zero.

Top Model:

X1 = b1*Psi + e1
X2 = b2*Psi + e2
X3 = b3*Psi + e3
X4 = b4*Psi + e4
Cov(e,e) = Cov(Psi, e) = 0

Bottom Model:

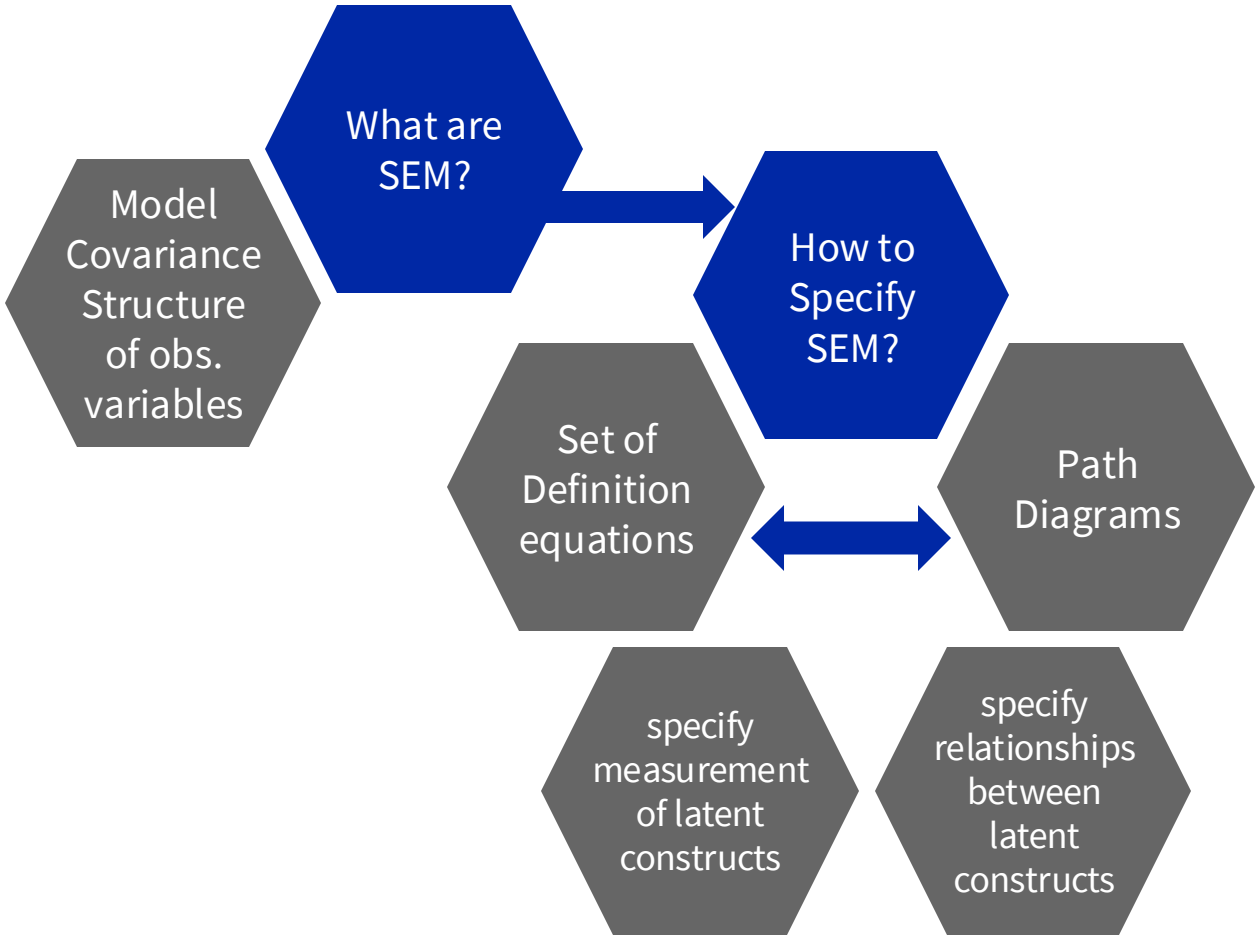
X1 = b1*LV1 + e1
X2 = b2*LV1 + e2
X3 = b3*LV2 + e3
X4 = b4*LV2 + e4
Cov(LV1,LV2)
Cov(e,e) = Cov(LV1, e) = Cov(LV2,e) = 0



Model Specification

Summary: Path diagrams & Definition Equations

Specifying SEM?



Model Specification

Exercise: Path diagrams & Definition Equations

Specifying SEM?

Specify the definition equations.

Definition Equations:

Y11 =

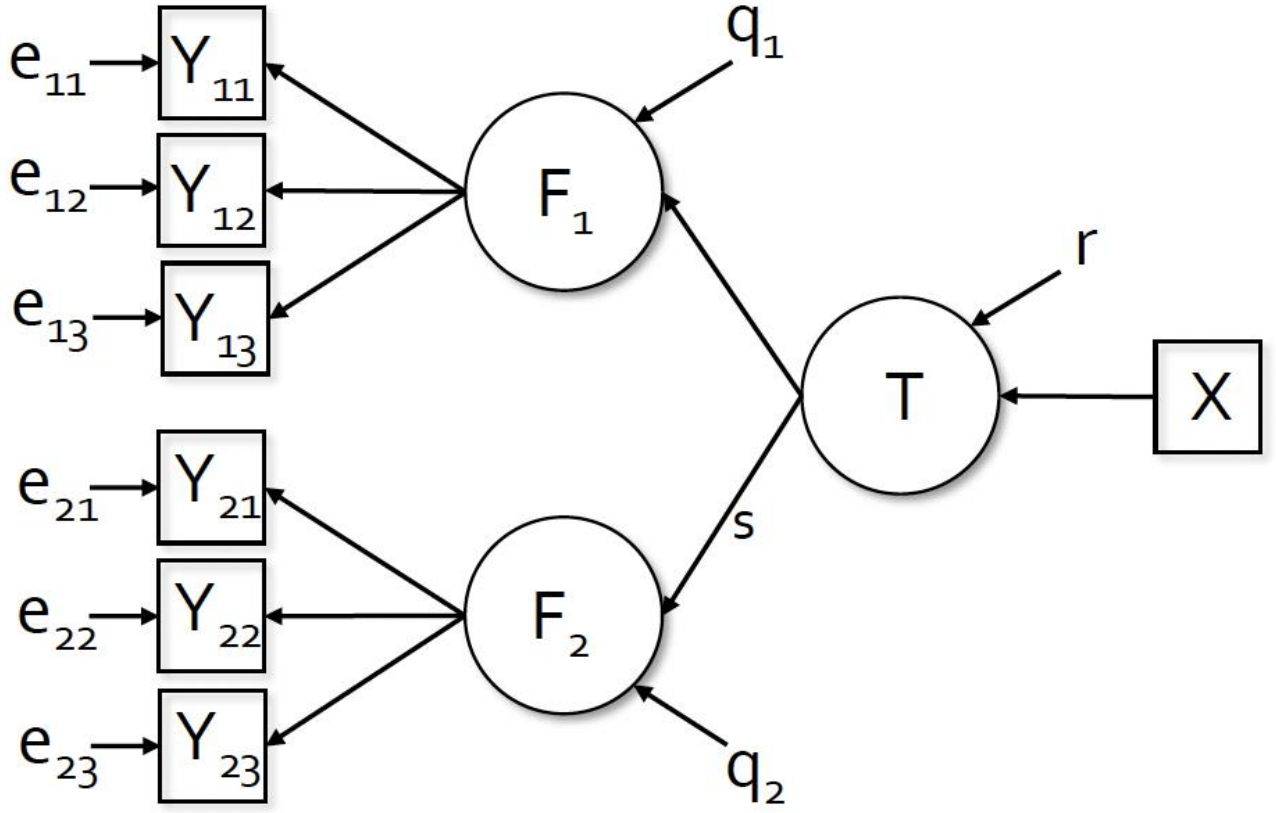
Y12 =

Y13 =

Y11 =

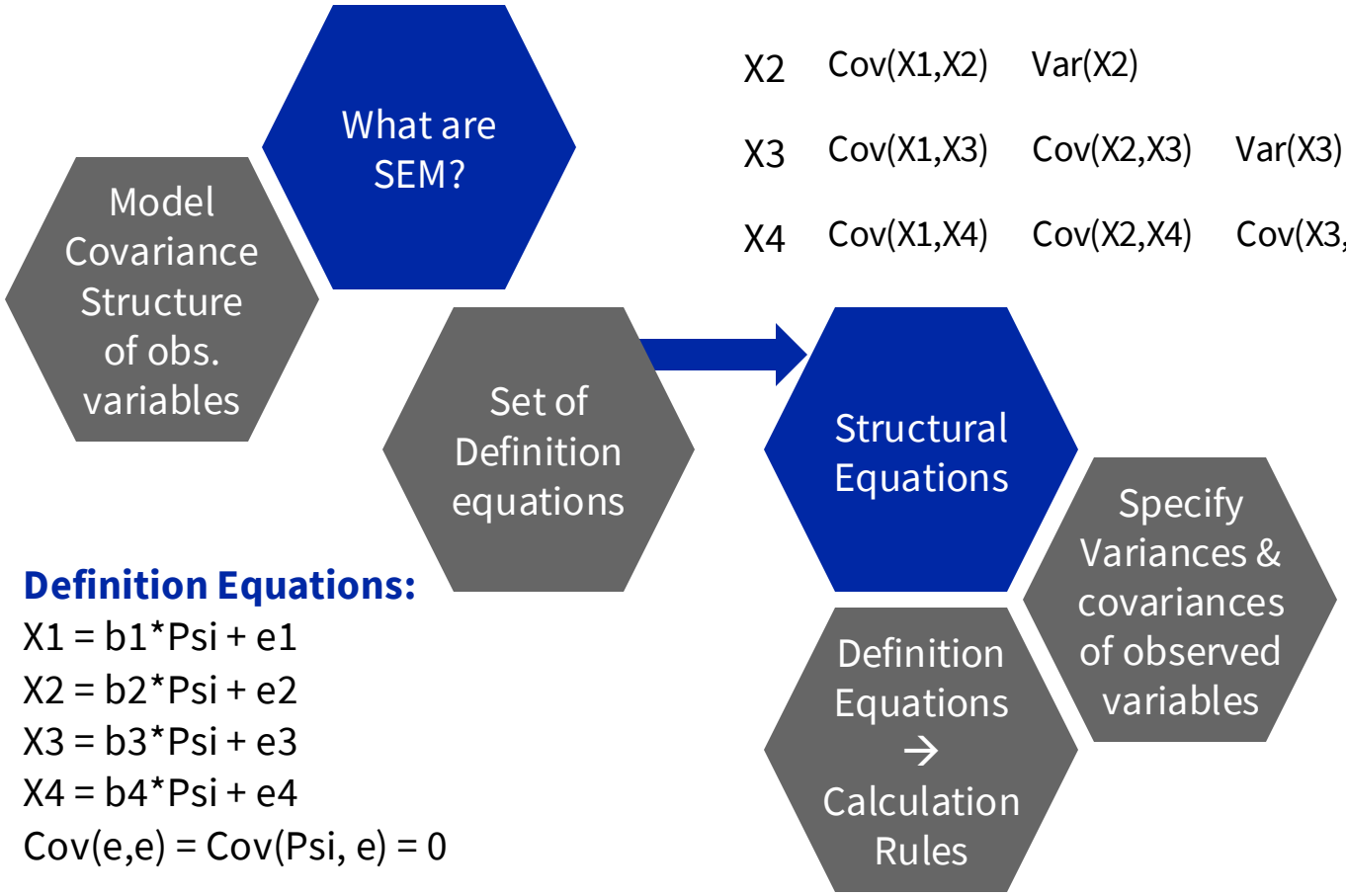
Y12 =

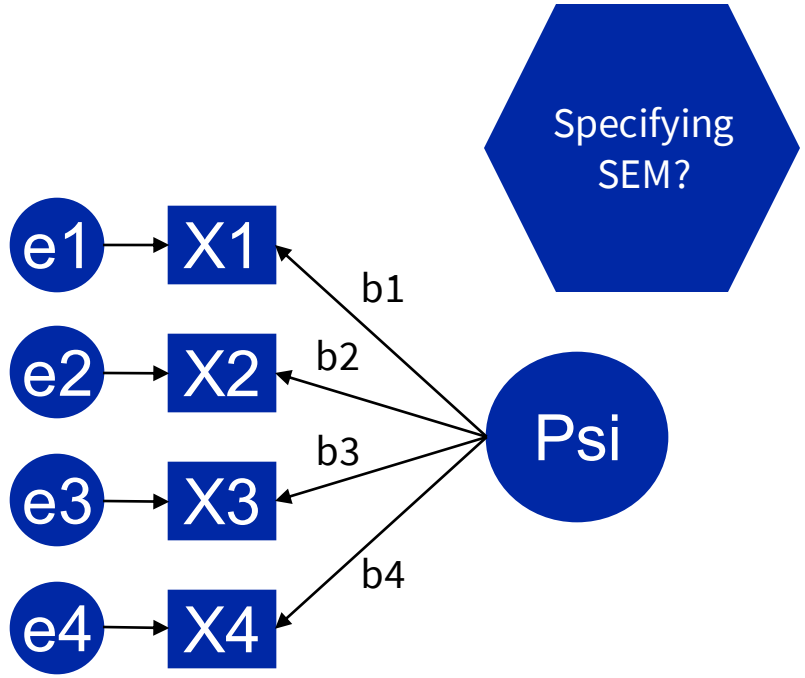
Y13 =



Model Specification
Structural Equations

	X1	X2	X3	X4
X1	Var(X1)			
X2	Cov(X1,X2)	Var(X2)		
X3	Cov(X1,X3)	Cov(X2,X3)	Var(X3)	
X4	Cov(X1,X4)	Cov(X2,X4)	Cov(X3,X4)	Var(X4)





Definition Equations:

X1 = b1*Psi + e1
X2 = b2*Psi + e2
X3 = b3*Psi + e3
X4 = b4*Psi + e4
Cov(e,e) = Cov(Psi, e) = 0

Var(X1) = Var(b1*Psi + e1)
= Var(b1*Psi) + Var(e1) + 2*b1*Cov(Psi,e1)
= $b1^2$*Var(Psi) + Var(e1)

Cov(X1,X2) = Cov(b1*Psi + e1,b2*Psi + e2)
= b1*b2*Cov(Psi,Psi) + b1*Cov(Psi,e2) + b2*Cov(e1,Psi) + Cov(e1,e2)
= b1*b2*Var(Psi)

Calculation Rules for Variances & Covariances:

- Cov(X,X) = Var(X)
- Cov(aX + bY, cV + dW) = ac * Cov(X,V) + ad * Cov(X,W) + bc * Cov(Y,V) + bd * Cov(Y,W)
- Var(aX + bY) = a^2 * Var(X) + b^2 * Var(Y) + 2ab * Cov(X,Y)

Model Specification

Exercise: Path diagrams & Definition Equations

Specifying SEM?

Specify the structural equations.

Definition Equations:

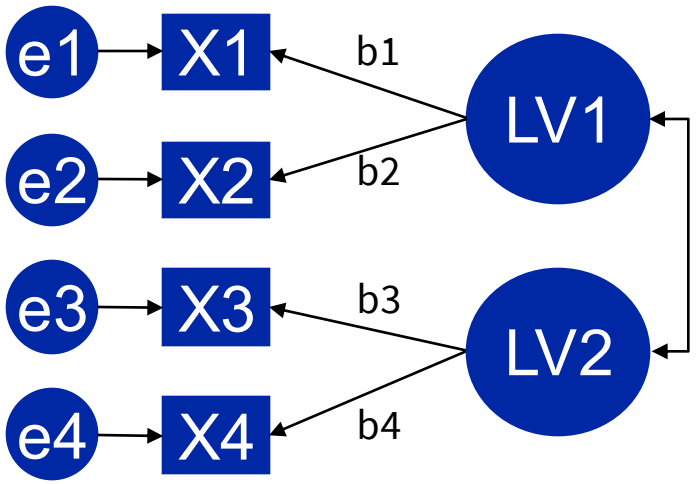
X1 = b1*LV1 + e1
X2 = b2*LV1 + e2
X3 = b3*LV2 + e3
X4 = b4*LV2 + e4

Cov(LV1,LV2)
Cov(e,e) = Cov(LV1, e) = Cov(LV2,e) = 0
Cov(q,X) = Cov(r,X) = 0

Structural Equations:

Var(X1) =
Var(X2) =
Var(X3) =
Var(X4) =

Cov(X1,X2) =
Cov(X1,X3) =
Cov(X1,X4) =
Cov(X2,X3) =
Cov(X2,X4) =
Cov(X3,X4) =



Calculation Rules for Variances & Covariances:

- Cov(X,X) = Var(X)
- Cov(aX + bY, cV + dW) = ac * Cov(X,V) + ad * Cov(X,W) + bc * Cov(Y,V) + bd * Cov(Y,W)
- $Var(aX + bY) = a^2 * Var(X) + b^2 * Var(Y) + 2ab * Cov(X,Y)$

What are your questions so far?

Model Specification

Model identification

Are these models Identified?

Identifying SEM

Variances of observed variables

Covariances among observed variables

Number of observed variables

Variances of exogeneous variables

Covariances among exogeneous variables

Number of estimated parameters

Loadings / Regression weights

> Over identified

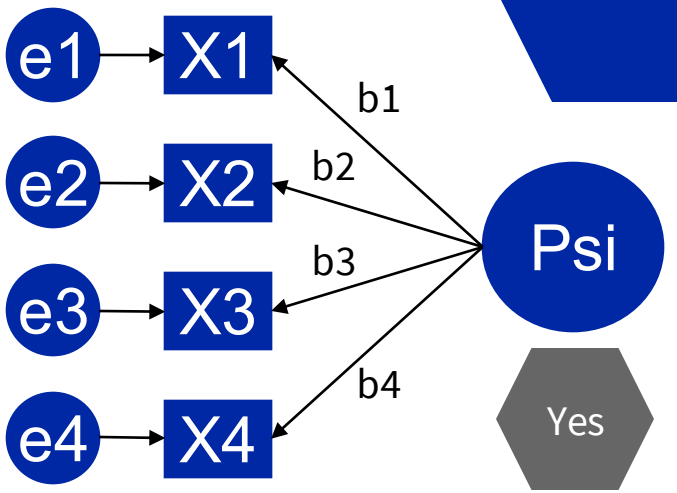
= exactly identified

< under identified

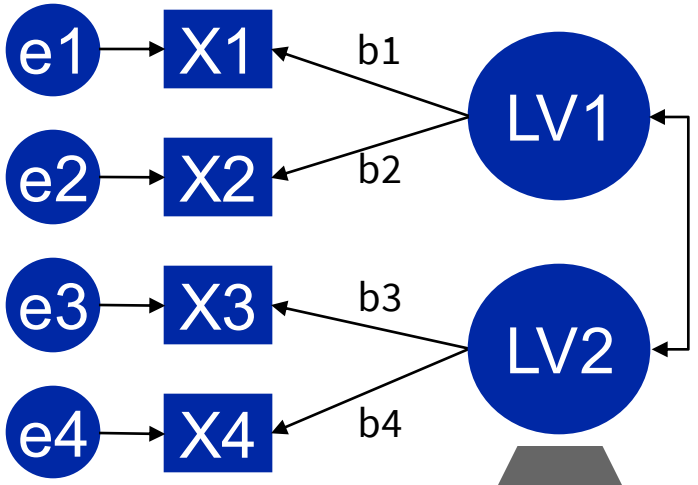
$df = N_{observed} - N_{parameters}$

→ Should be larger than 0

Specifying SEM?



Yes



No

Model Specification

Model identification

$df = N_{observed} - N_{parameters}$

→ Should be larger than 0

Number of observed variables

Number of estimated parameters

Identifying SEM

Normalization

Fix one loading (typically to 1)

All loadings from LV are free

Infinite Solutions if:

&

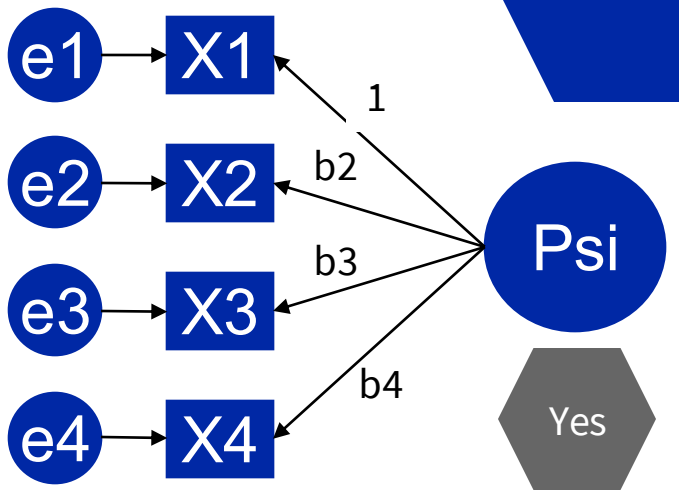
Variance of LV is free

Standardization

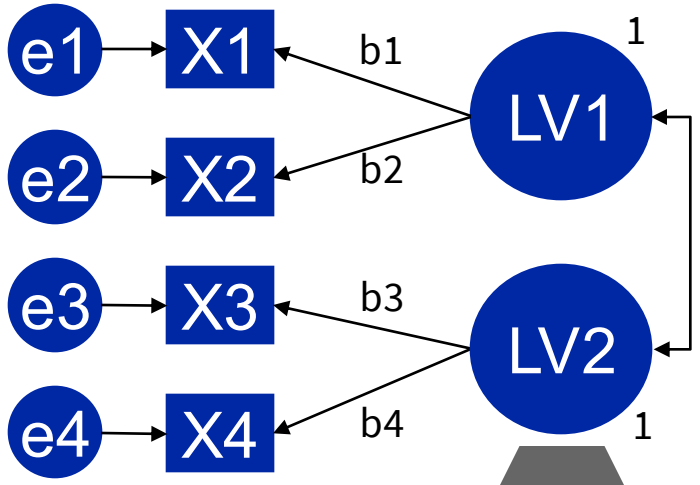
Fix LV variance (typically to 1)

Identifying these models

Specifying SEM?



Yes



Yes

What are your questions so far?

Model Specification

Introduction: lavaan syntax

Specifying SEM?

How to Specify SEM in lavaan?

specify measurement of latent constructs

=~
"is measured by"

UI-MW =~ UI1 + UI2 + UI3 + UI4

I-MW =~ I1 + I2 + I3 + I4

MA-MW =~ MA1 + MA2 + MA3 + MA4

Default Identification: Normalization

specify relationships between latent constructs

~~
"is correlated with"

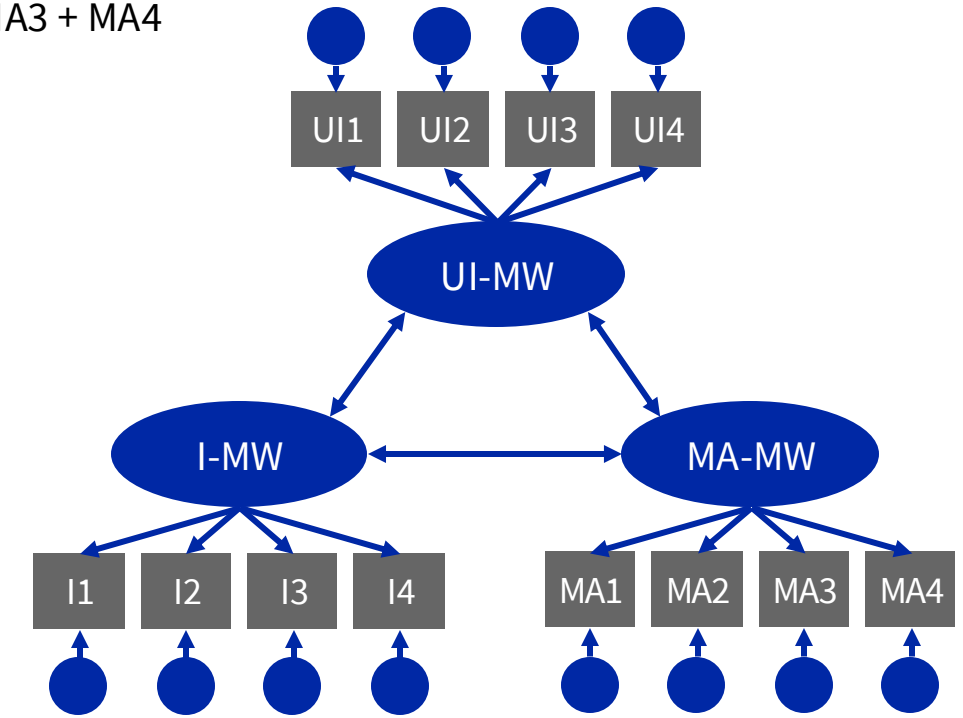
NA → should be freely estimated

UI-MW ~~ NA*I-MW + NA*MA-MW

I-MW ~~ NA*MA-MW

~
"is predicted by"

Default:
All exogeneous LV (except residuals) are correlated



Model Specification

Introduction: lavaan syntax

Specifying SEM?

How to Specify SEM in lavaan?

specify measurement of latent constructs

=~ "is measured by"

specify relationships between latent constructs

Residuals / Error Variances

~~ "is correlated with"

~ "is predicted by"

UI-MW =~ UI1 + UI2 + UI3 + UI4

I-MW =~ I1 + I2 + I3 + I4

MA-MW =~ MA1 + MA2 + MA3 + MA4

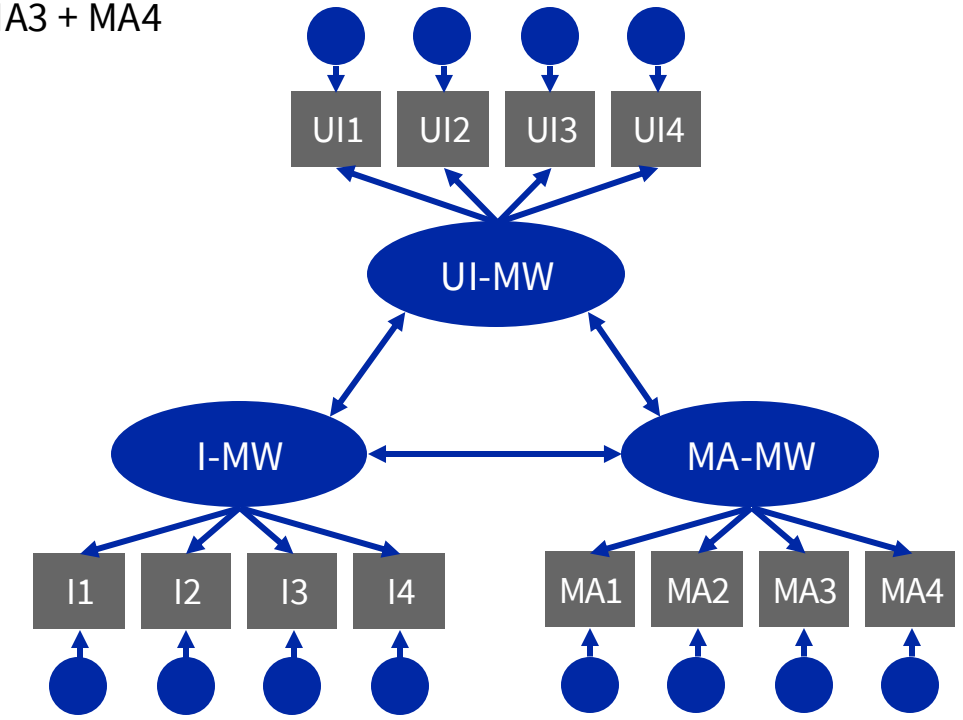
Example assuming equal error variances:

UI1 ~~ e1*UI1

UI2 ~~ e1*UI2

UI3 ~~ e1*UI3

UI4 ~~ e1*UI4



Model Specification

Introduction: lavaan syntax

How to Specify SEM in lavaan?

Model Syntax = Text String

My suggestion:

Create separate Text file

Use `readLines`

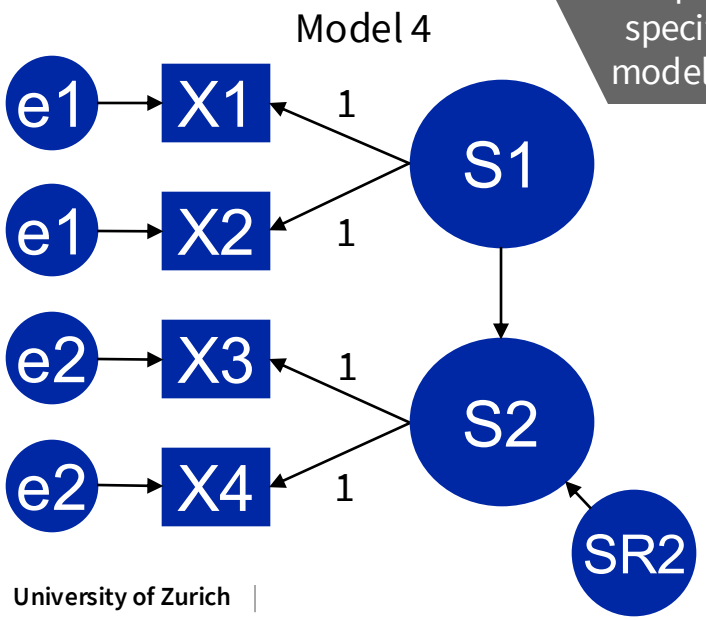
→ import the specified model to R

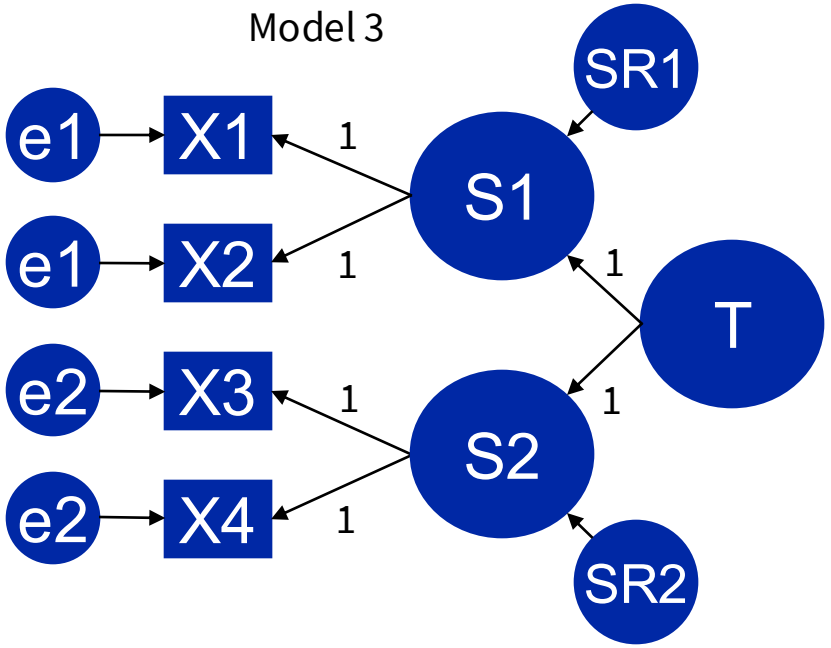
Save them In a separate Folder "SEM" or "models"

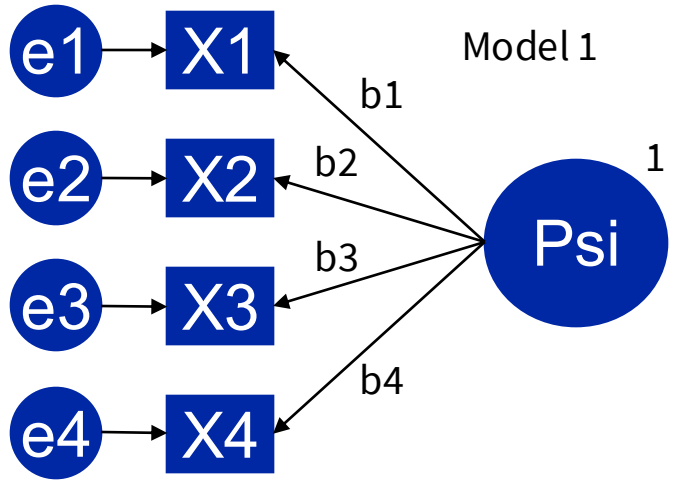
Let's Practice ☺

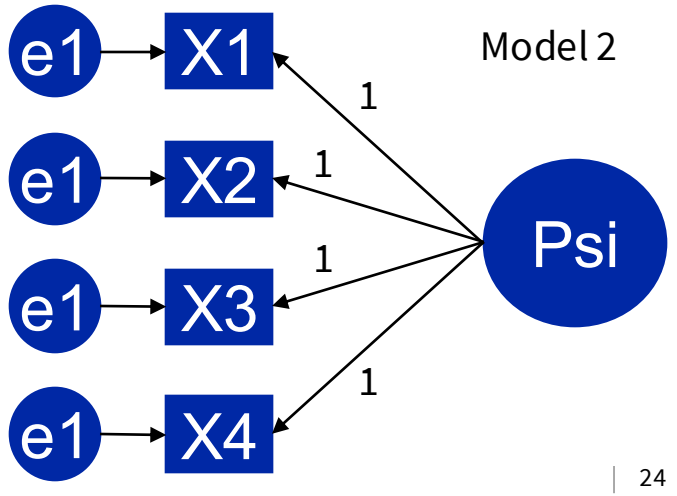
Open RStudio and specify the following SEM

Specifying SEM?









What are your questions so far?

Model Specification

Simulating Data: lavaan

Specified SEM

Given a set of parameters → Simulate data

Implies covariance structure for observed vars.

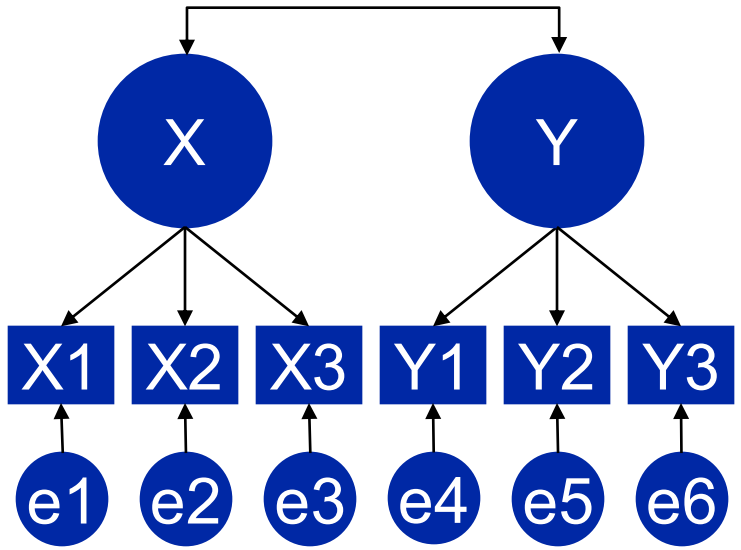
Variances of Exogenous variables

Loadings / Regression Weights

	X1	X2	X3	Y1	Y2	Y3
X1	Var(X) + Var(e1)	Var(X)	Var(X)	Cov(X,Y)	Cov(X,Y)	Cov(X,Y)
X2		Var(X) + Var(e2)	Var(X)	Cov(X,Y)	Cov(X,Y)	Cov(X,Y)
X3			Var(X) + Var(e3)	Cov(X,Y)	Cov(X,Y)	Cov(X,Y)
Y1				Var(Y) + Var(e4)	Var(Y)	Var(Y)
Y2					Var(Y) + Var(e5)	Var(Y)
Y3						Var(Y) + Var(e6)

Assuming all loadings = 1

Specifying SEM?




```
# lavaan model
X =~ b1*X1 + b2*X2 + b3*X3
Y =~ b4*Y1 + b5*Y2 + b6*Y3

X ~~ covXY*Y

X ~~ varX*X
Y ~~ varY*Y

X1 ~~ e1*X1
X2 ~~ e2*X2
X3 ~~ e3*X3
Y1 ~~ e4*Y1
Y2 ~~ e5*Y2
Y3 ~~ e6*Y3
```

Exercise: Specify SEM with lavaan syntax

Label all parameters that could be free

Model Specification

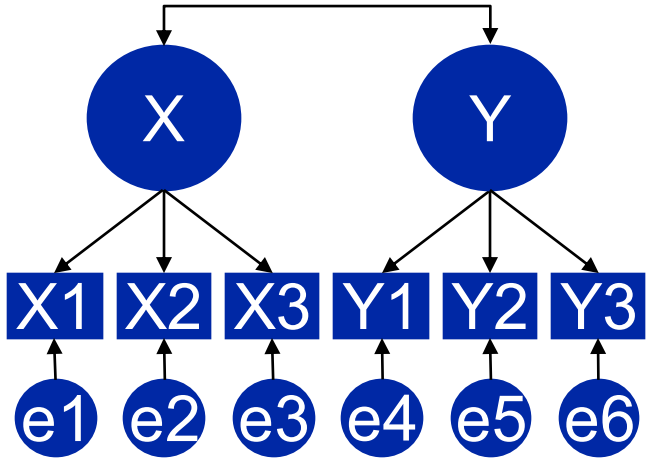
Simulating Data: lavaan

```
# lavaan model
X =~ b1*X1 + b2*X2 + b3*X3
Y =~ b4*Y1 + b5*Y2 + b6*Y3

X ~~ covXY*Y

X ~~ varX*X
Y ~~ varY*Y

X1 ~~ e1*X1
X2 ~~ e2*X2
X3 ~~ e3*X3
Y1 ~~ e4*Y1
Y2 ~~ e5*Y2
Y3 ~~ e6*Y3
```



Specifying SEM?

Exercise: Specify SEM with lavaan syntax

2) Simulate Data for 200 subjects

```
# load model
my_model <- readLines("my-model-syntax.txt")

# simulate data
my_data <- simulateData(model, sample.nobs)
```

3) Calculate The empirical covariance matrix

4) Obtain The model implied covariance matrix

```
# implied cov.structure
fitted(sem(model))
```

1) Fix parameters To these values

b1	b2	b3	b4	b5	b6
0.8	0.7	0.9	0.6	0.5	0.7
e1	e2	e3	e4	e5	e6
0.3	0.5	0.2	0.65	0.75	0.5
covXY	varX	varY			
.60	1.00	1.00			

5) Correlate the aggregated X and Y values

Model Specification

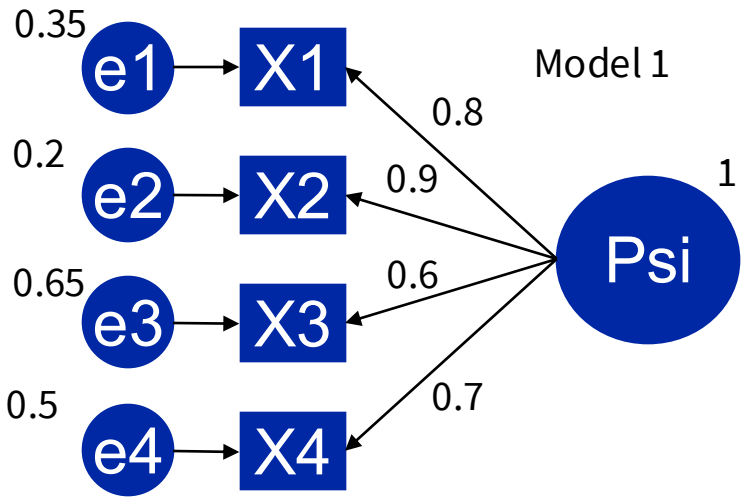
Simulating Data: lavaan

Exercise: Compare the implied covariance structures

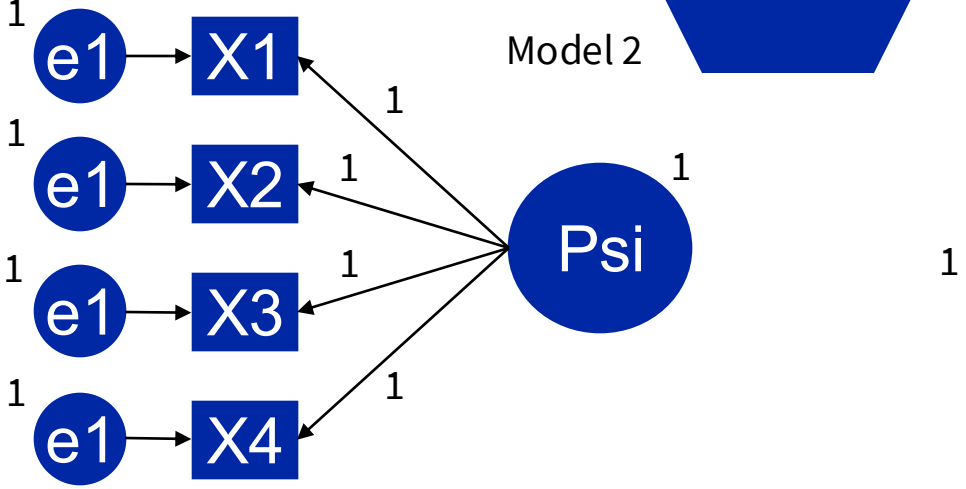
```
# implied cov.structure
fitted(sem(model))
```

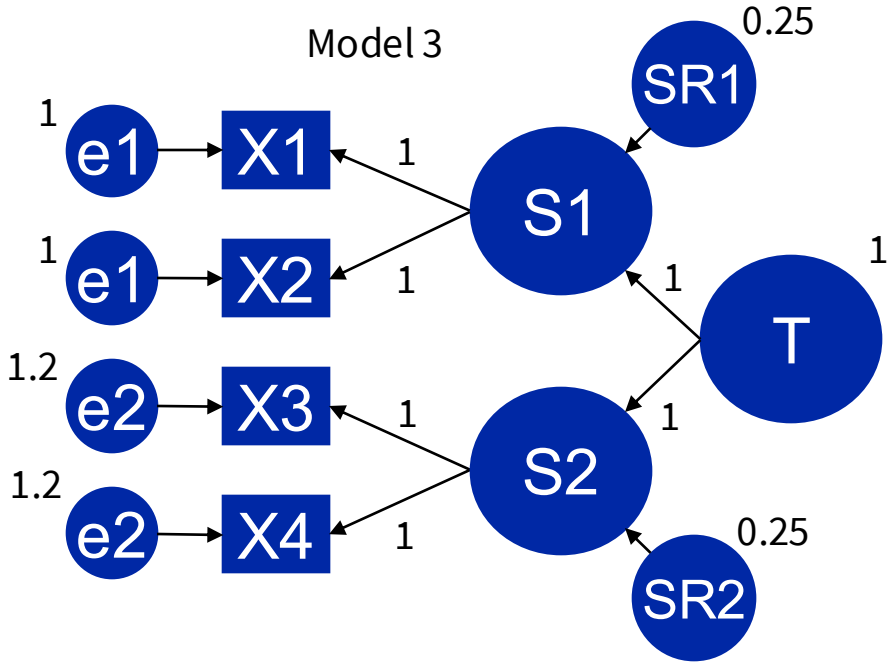
Defaults for simulation

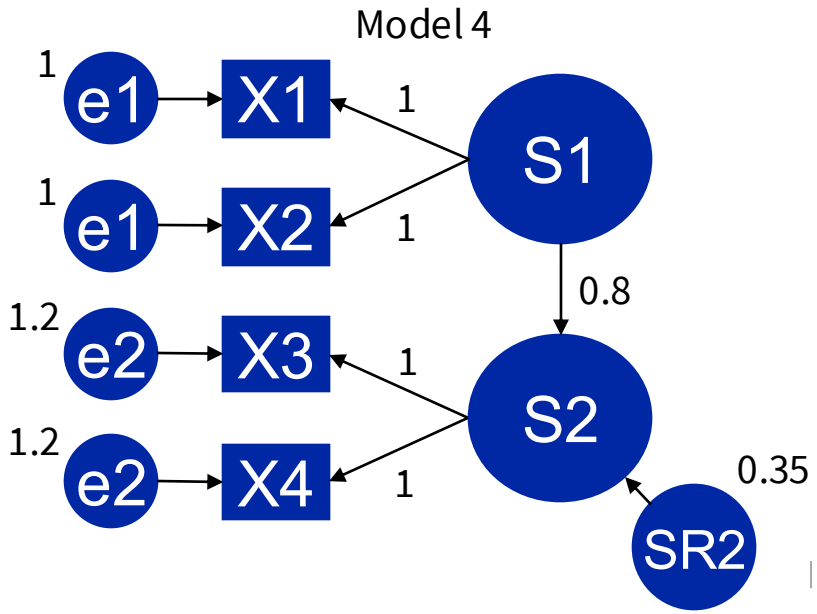
Variances = 1 If not specified



Specifying SEM?







What are your questions so far?