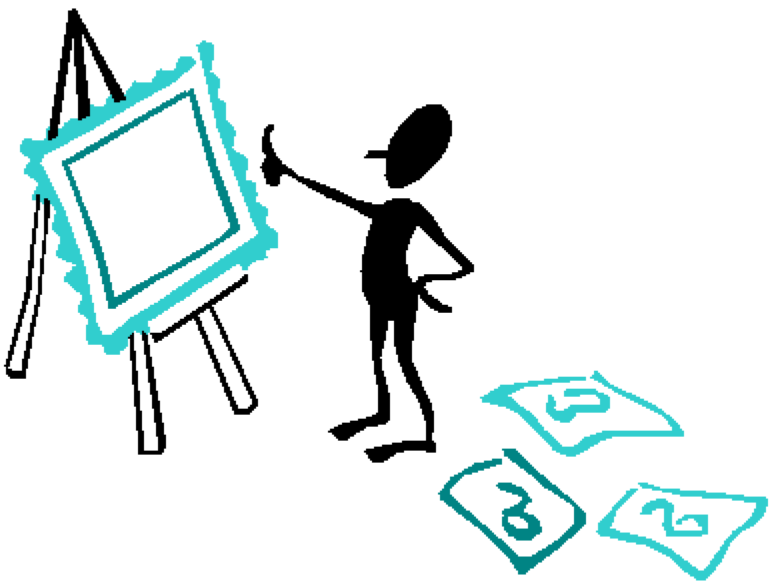
PATH ANALYSIS

Modeling latent variables

Outline

- *Introduction*
- *Definition*
- *Basic assumptions*
- *Decomposition*
- *Path coefficients*
- *Effects*
- *Application*
- *The causality problem*



Path analysis: introduction

- Statistical procedure for analysing the structure using **directly observable** variables
- Introduced in 1934 by Sewall Wright (american genetic scientist)
- In the 60s-70s very popular in sociology
- … is important because principles of path analysis also apply to structural models

Path analysis

- … is a method for investigating an *a priori postulated* path model
- A path model is an explanation of why x and y are related to each other
- … is conducted on the basis of (observed) covariances and correlations
- **Major characteristic:** estimation of directional and non-directional effects

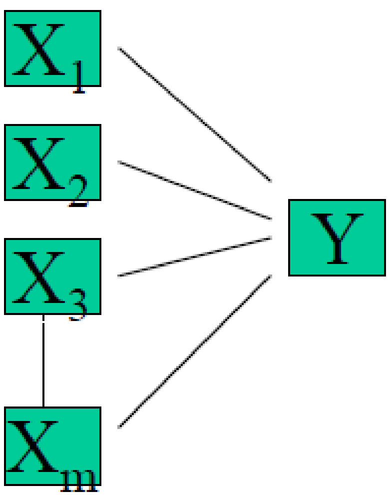
Path analysis

- Specificities of path analysis:
  - … is for *observable* variables
  - … requires *one indicator* only (actually no indicator!)
  - … implicitly assums that there is no error of measurement

Path analysis

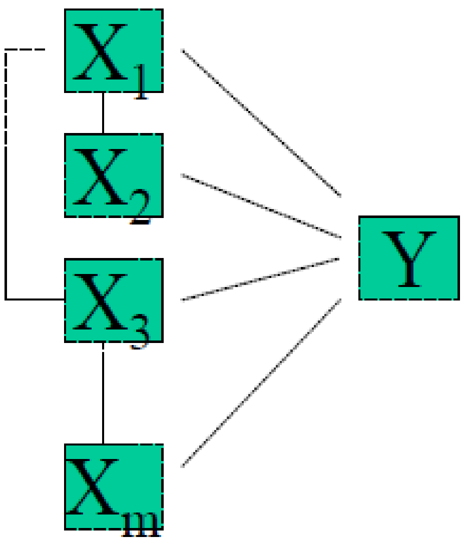
A comparison of path analysis and multiple regression

Multiple regression



$Y = a + \beta_1 X_1 + \ldots + \beta_m X_m + E$

Path analysis

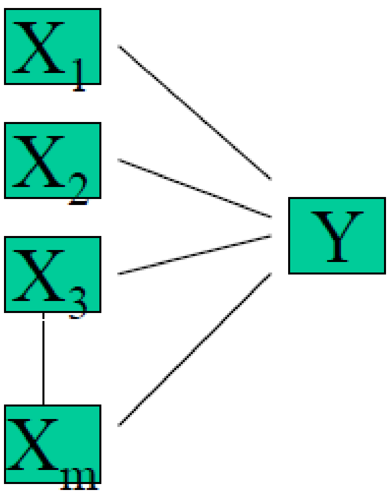


i. e. X_3 can influence Y using different paths

Path analysis

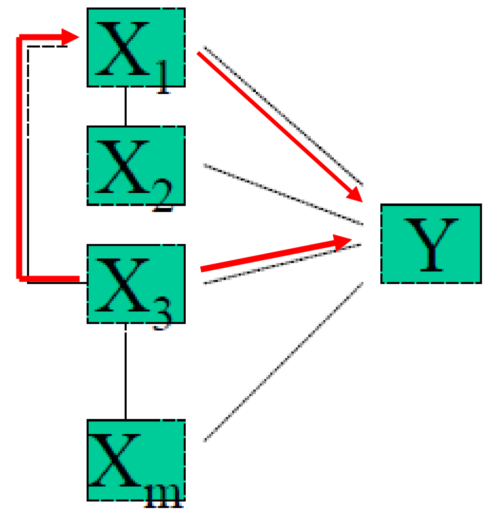
A comparison …

Multiple regression



Y = a + ß$_1$X$_1$ + … + ß$_m$X$_m$ + E

Path analysis



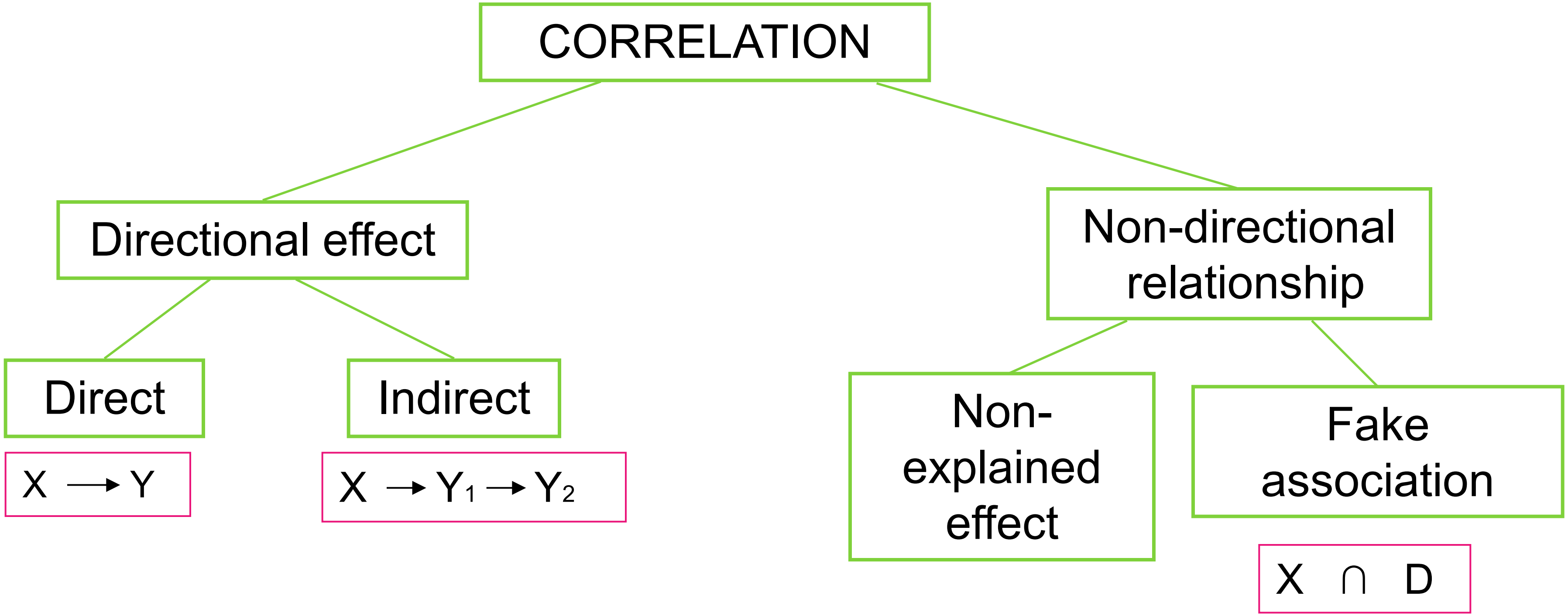
i. e. X_3 can influence Y using different paths (see red lines)

Path analysis: basic assumption

- **A correlation can comprise several directional and non-directional effects**
- Path analysis decomposes a correlation (or covariance) in directional and non-directional components
- → „decomposition of correlations into effects“

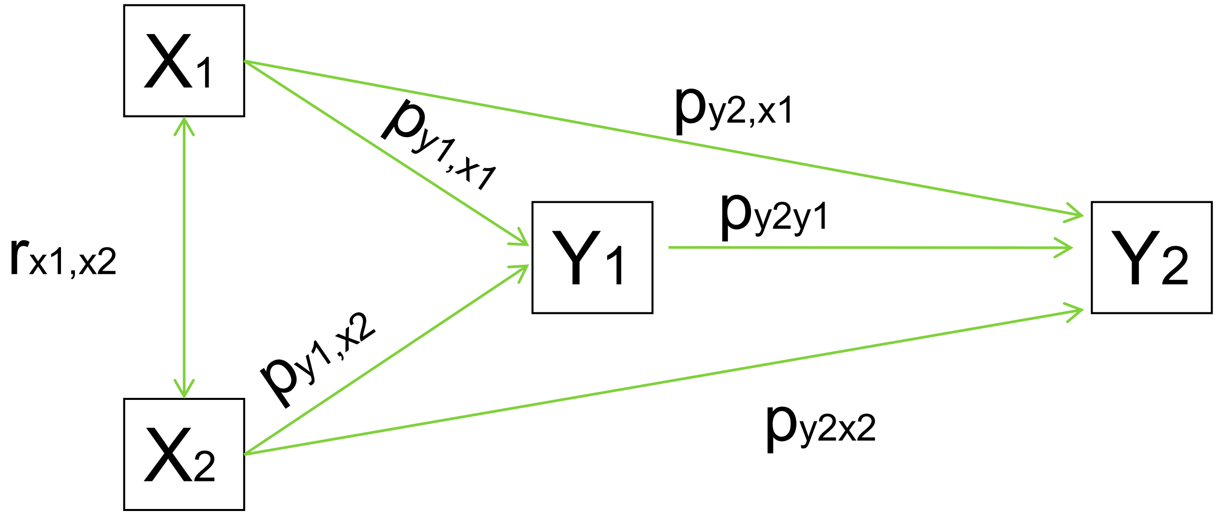
Path analysis

- Decomposition of effects:



Path analysis

- Simple path analysis model with
  - two independent variables X_1 and X_2, a mediator variable Y_1 and a dependent variable Y_2

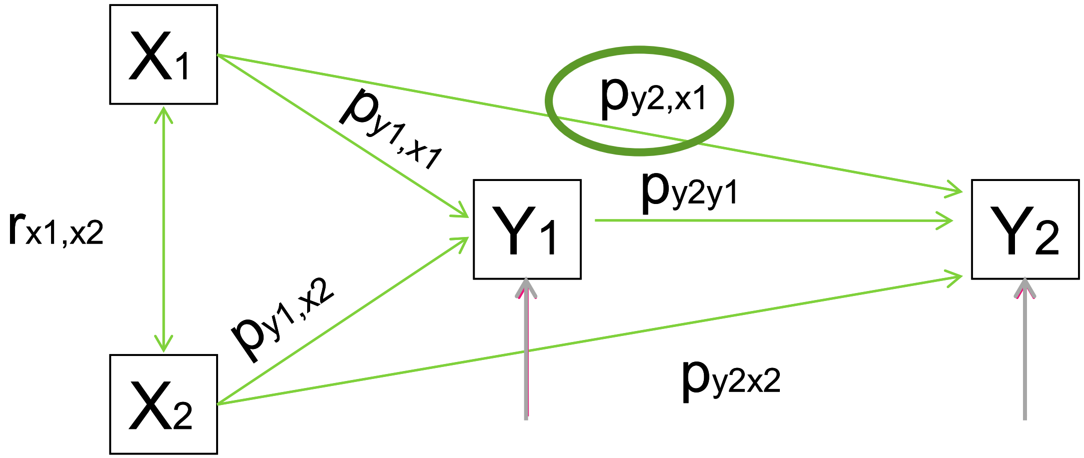


  - … the variables give rise to six correlations

Path analysis: decomposition

- Decomposition the correlation $\mathbf{r_{x1,y2}}$.

DE = Direct effect
IE = Indirect effect
NE = by the model not explained effect



$\mathbf{r_{x1,y2}} = p_{y2,x1} + \ldots.$

(DE) +

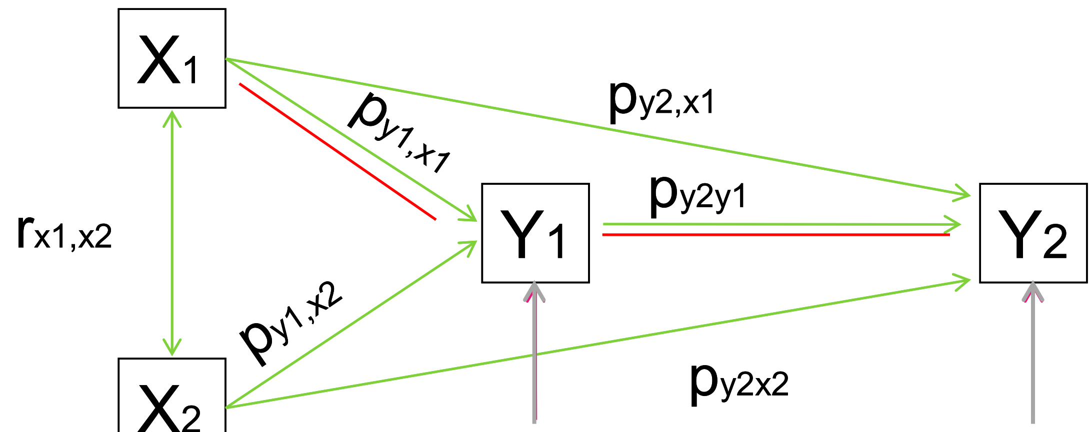
To be estimated parameters of a path model:

- [Product-moment Korrelations **(r)**]
- Path coefficients **(p)**

Path analysis

- Decomposition the correlation $\mathbf{r_{x1,y2}}$.

DE = Direct effect
IE = Indirect effect
NE = by the model not explained effect



$\mathbf{r_{x1,y2}} = p_{y2,x1} + (p_{y1,x1}) \cdot (p_{y2,y1}) + \ldots$

(DE) + (IE) + …

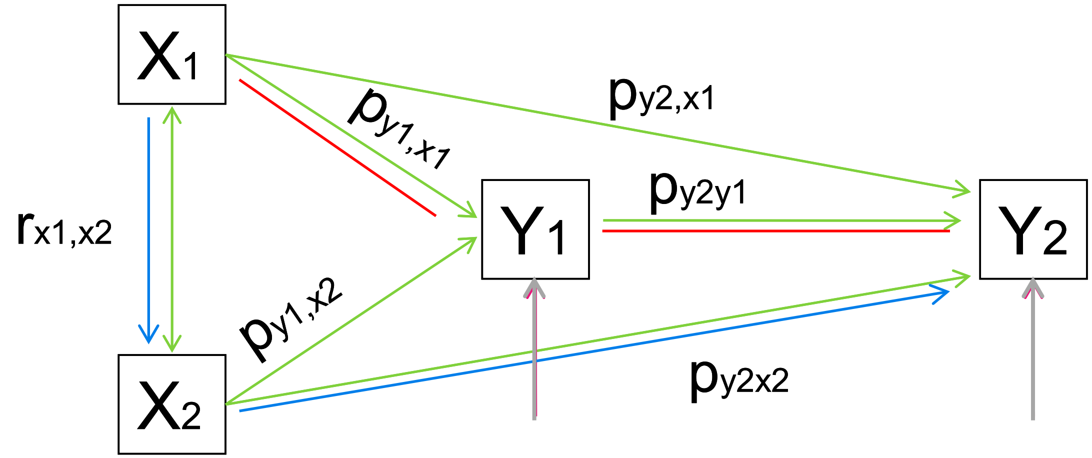
To be estimated parameters of a path model:

- [Product-moment Korrelations **(r)**]
- Path coefficients **(p)**

Path analysis

- Decomposition the correlation $\mathbf{r_{x1,y2}}$.

DE = Direct effect
IE = Indirect effect
NE = by the model not explained effect



$\mathbf{r_{x1,y2}} = p_{y2,x1} + (p_{y1,x1})\cdot(p_{y2,y1}) + (r_{x1,x2})\cdot(p_{y2,x2}) + \ldots$

(DE) + (IE) + (NE) +

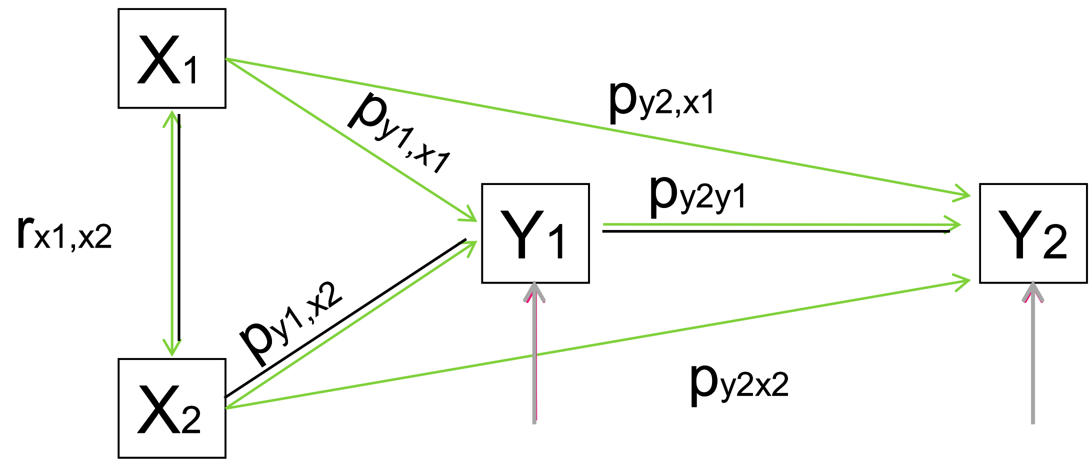
To be estimated parameters of a path model:

- [Product-moment Korrelations **(r)**]
- Path coefficients **(p)**

Path analysis

- Decomposition the correlation $\mathbf{r_{x1,y2}}$.

DE = Direct effect
IE = Indirect effect
NE = by the model not explained effect



$$\mathbf{r_{x_1,y_2}} = p_{y_2,x_1} + (p_{y_1,x_1})\cdot(p_{y_2,y_1}) + (r_{x_1,x_2})\cdot(p_{y_2,x_2}) + (r_{x_1,x_2})\cdot(p_{y_1,x_2})\cdot(p_{y_2,y_1})$$

(DE) + (IE) + (NE) + (NE)

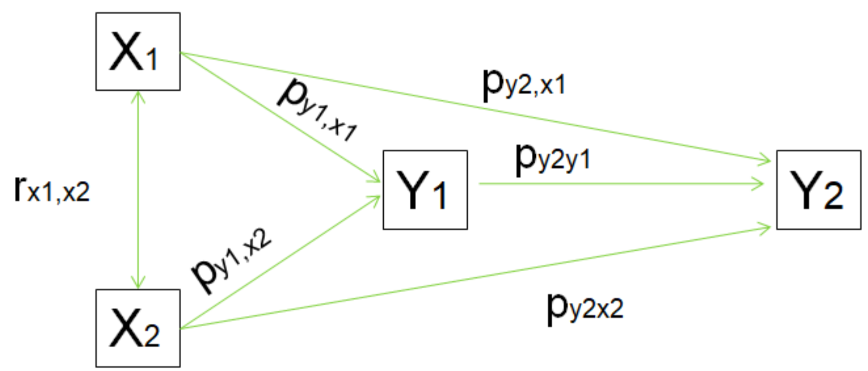
To be estimated parameters of a path model:

- [Product-moment Korrelations **(r)**]
- Path coefficients **(p)**

Path analysis

- Decomposition:
- $r_{x1y1} = p_{y1x1} + (r_{x1x2}) \cdot (p_{y1x2})$
- $r_{x2y1} = p_{y1x2} + (r_{x1x2}) \cdot (p_{y1x1})$
- $r_{x1y2} = p_{y2x1} + (p_{y1x1}) \cdot (p_{y2y1}) + (r_{x1x2}) \cdot (p_{y2x2}) + (r_{x1x2}) \cdot (p_{y1x2}) \cdot (p_{y2y1})$

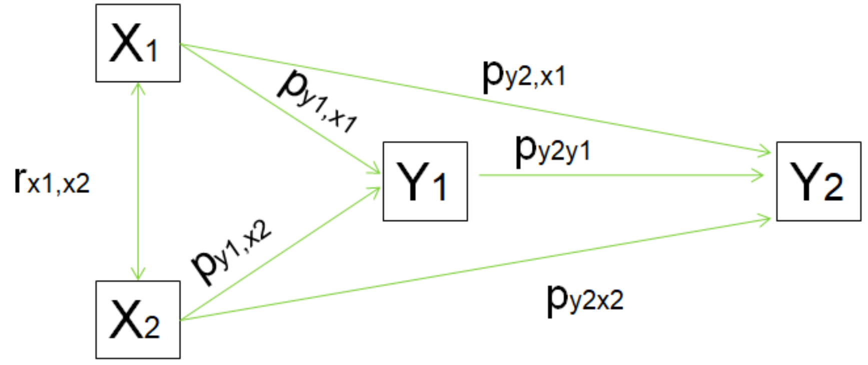
  ....



Path analysis

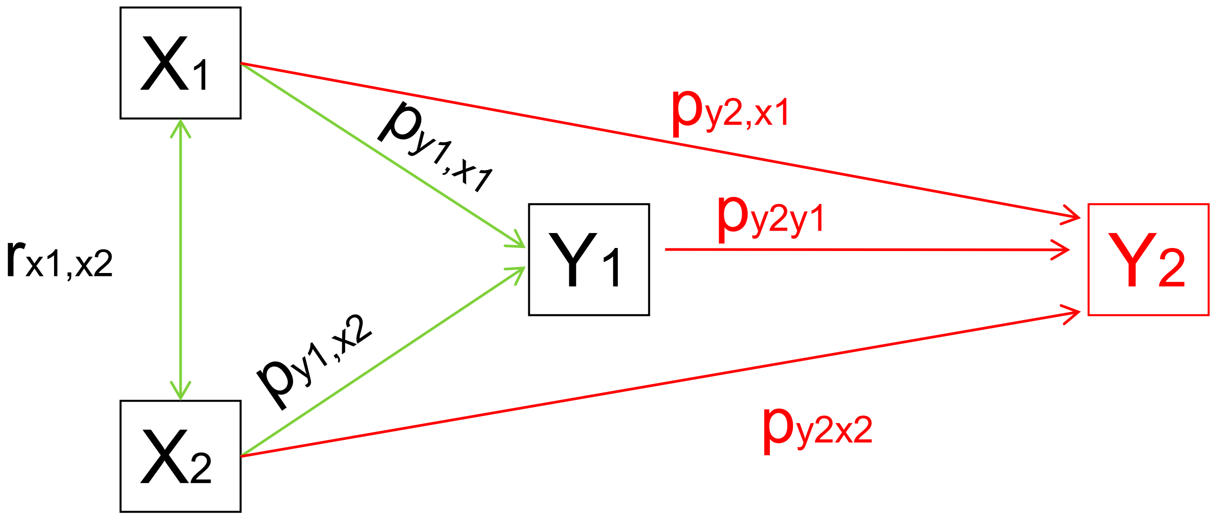
- **Decomposition:**
- $r_{x1y1} = p_{y1x1} + (r_{x1x2}) \cdot (p_{y1x2})$
- $r_{x2y1} = p_{y1x2} + (r_{x1x2}) \cdot (p_{y1x1})$
- $r_{x1y2} = p_{y2x1} + (p_{y1x1}) \cdot (p_{y2y1}) + (r_{x1x2}) \cdot (p_{y2x2}) + (r_{x1x2}) \cdot (p_{y1x2}) \cdot (p_{y2y1})$
- $r_{x2y2} = p_{y2x2} + (p_{y1x2}) \cdot (p_{y2y1}) + (r_{x1x2}) \cdot (p_{y2x1}) + (r_{x1x2}) \cdot (p_{y1x1}) \cdot (p_{y1y2})$
- $r_{y1y2} = p_{y2y1}$

…..



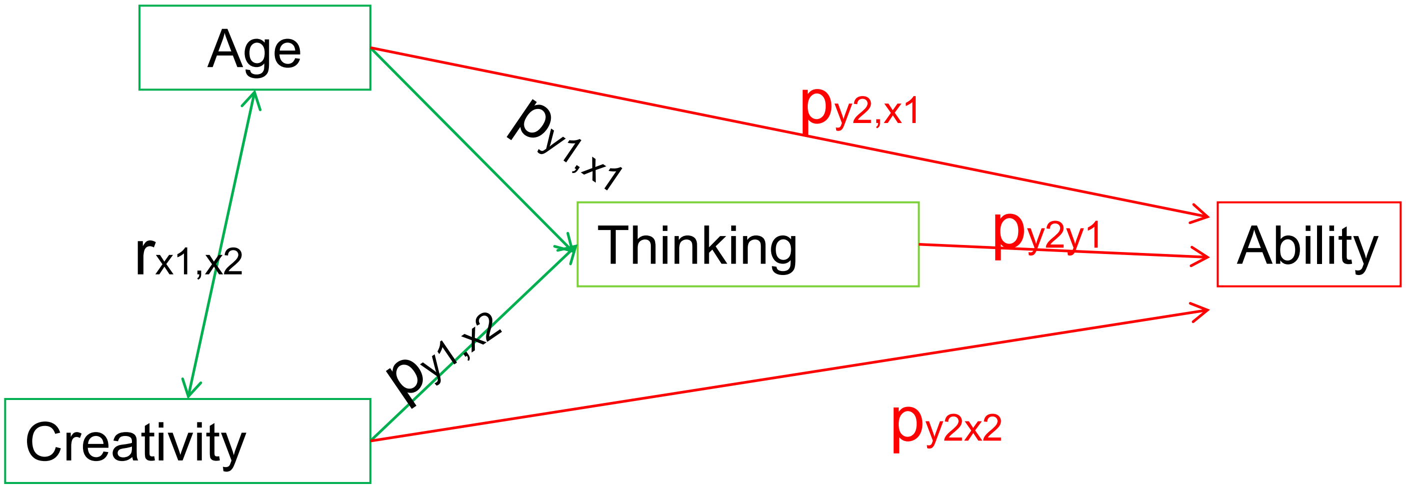
Path analysis

- <u>The new parts of the structure model</u>



Path analysis

- **Example**
  - Age (X1) ; creativity (X2)
  - Thinking (Y1) ; problem solving ability (Y2)



Path analysis

Decomposition for example:

- 1 *age & creativity* $(r_{x1x2}) = r_{x1x2}$
- 2 *age & thinking* $(r_{x1x2}) = p_{y1x1} + (r_{x1x2})\cdot(p_{y1x2})$
- 3 *creativity & thinking* $(r_{x2y1}) = p_{y1x2} + (r_{x1x2})\cdot(p_{y1x1})$

Path analysis

- 4 *age & problem solving ability* (r_{x1x2}) =

$$p_{y2x1} + (p_{y1x1})\cdot(p_{y2y1}) + (r_{x1x2})\cdot(p_{y2x2}) + (r_{x1x2})\cdot(p_{y1x2})\cdot(p_{y2y1})$$

- 5 *creativity & problem solving ability* (r_{x2y2}) =

$$p_{y2x2} + (p_{y1x2})\cdot(p_{y2y1}) + (r_{x1x2})\cdot(p_{y2x1}) + (r_{x1x2})\cdot(p_{y1x1})\cdot(p_{y1y2})$$

Path analysis

- 6 *thinking & problem solving ability* (r_{y1y2}) = p_{y2y1}

Path analysis: *the unkowns*

- Decomposition:
- $r_{x1y1} = p_{y1x1} + (r_{x1x2}) \cdot (p_{y1x2})$
- $r_{x2y1} = p_{y1x2} + (r_{x1x2}) \cdot (p_{y1x1})$
- $r_{x1y2} = p_{y2x1} + (p_{y1x1}) \cdot (p_{y2y1}) + (r_{x1x2}) \cdot (p_{y2x2}) + (r_{x1x2}) \cdot (p_{y1x2}) \cdot (p_{y2y1})$
- $r_{x2y2} = p_{y2x2} + (p_{y1x2}) \cdot (p_{y2y1}) + (r_{x1x2}) \cdot (p_{y2x1}) + (r_{x1x2}) \cdot (p_{y1x1}) \cdot (p_{y1y2})$

…..

→ Series of linear equations

Path analysis: path coefficients

- Decomposition:
- $r_{x1y1} = p_{y1x1} + (r_{x1x2})\cdot(p_{y1x2})$
- $r_{x2y1} = p_{y1x2} + (r_{x1x2})\cdot(p_{y1x1})$
- $r_{x1y2} = p_{y2x1}+(p_{y1x1})\cdot(p_{y2y1})+(r_{x1x2})\cdot(p_{y2x2})+(r_{x1x2})\cdot(p_{y1x2})\cdot(p_{y2y1})$
- $r_{x2y2} = p_{y2x2}+(p_{y1x2})\cdot(p_{y2y1})+(r_{x1x2})\cdot(p_{y2x1})+(r_{x1x2})\cdot(p_{y1x1})\cdot(p_{y1y2})$

  …..

→ Series of linear equations

- with path coefficients as unknowns
- the path coefficients can be determined if their number corresponds to the number of equations
- the path coefficients can be estimated if their number is smaller than the number of equations

Path analysis

- Demonstration of how to find estimates of the path coefficients

  Given are r_{x1y1} , r_{x2y1} and r_{x1x2} and the following decompositions:

- $r_{x1y1} = p_{y1x1} + (r_{x1x2})\cdot(p_{y1x2})$ (1)
- $r_{x2y1} = p_{y1x2} + (r_{x1x2})\cdot(p_{y1x1})$ (2)

(1) -> (3) $-p_{y1x1} = -r_{x1y1} + r_{x1x2} \times p_{y1x2}$

(3) -> (4) $p_{y1x1} = r_{x1y1} - r_{x1x2} \times p_{y1x2}$

(2) -> (5) $r_{x2y1} = p_{y1x2} + r_{x1x2} \times (r_{x1y1} - r_{x1x2} \times p_{y1x2})$

(5) -> (6) $r_{x2y1} = p_{y1x2} + r_{x1x2} \times r_{x1y1} - r_{x1x2} \times r_{x1x2} \times p_{y1x2}$

Path analysis

- Demonstration of how to find estimates of the path coefficients

…..

(2) -> (5) $r_{x2y1} = p_{y1x2} + r_{x1x2} \times (r_{x1y1} - r_{x1x2} \times p_{y1x2})$

(5) -> (6) $r_{x2y1} = p_{y1x2} + r_{x1x2} \times r_{x1y1} - r_{x1x2} \times r_{x1x2} \times p_{y1x2}$

(6) -> (7) $r_{x2y1} = p_{y1x2} \times (1 - r_{x1x2} \times r_{x1x2}) + r_{x1x2} \times r_{x1y1}$

(7) -> (8) $p_{y1x2} \times (1 - r_{x1x2} \times r_{x1x2}) = r_{x2y1} - r_{x1x2} \times r_{x1y1}$

(8) -> (9) $$p_{y1x2} = \frac{r_{x2y1} - r_{x1x2} \times r_{x1y1}}{1 - r_{x1x2} \times r_{x1x2}}$$

Path analysis

- Demonstration of how to find estimates of the path coefficients

  …..

(7) -> (8) $$p_{y1x2} = \frac{r_{x2y1} - r_{x1x2} \times r_{x1y1}}{1 - r_{x1x2} \times r_{x1x2}}$$

- … it corresponds to the standardized partial regression coefficent:

$$\beta_{y1x2.x1} = \frac{r_{x2y1} - r_{x1x2} \times r_{x1y1}}{1 - r_{x1x2}^2}$$

Path analysis

Application to the example of regression analysis

Predictors

Criterion

Estimate by p_{yx1} ? .61

Age x_1

r=0.7

r=0.8

y Thinking

Creativity x_2

r=0.6

Estimate by p_{yx2} ? .11

$r_{x1,x2} = .8$

$r_{x1,y} = .7$

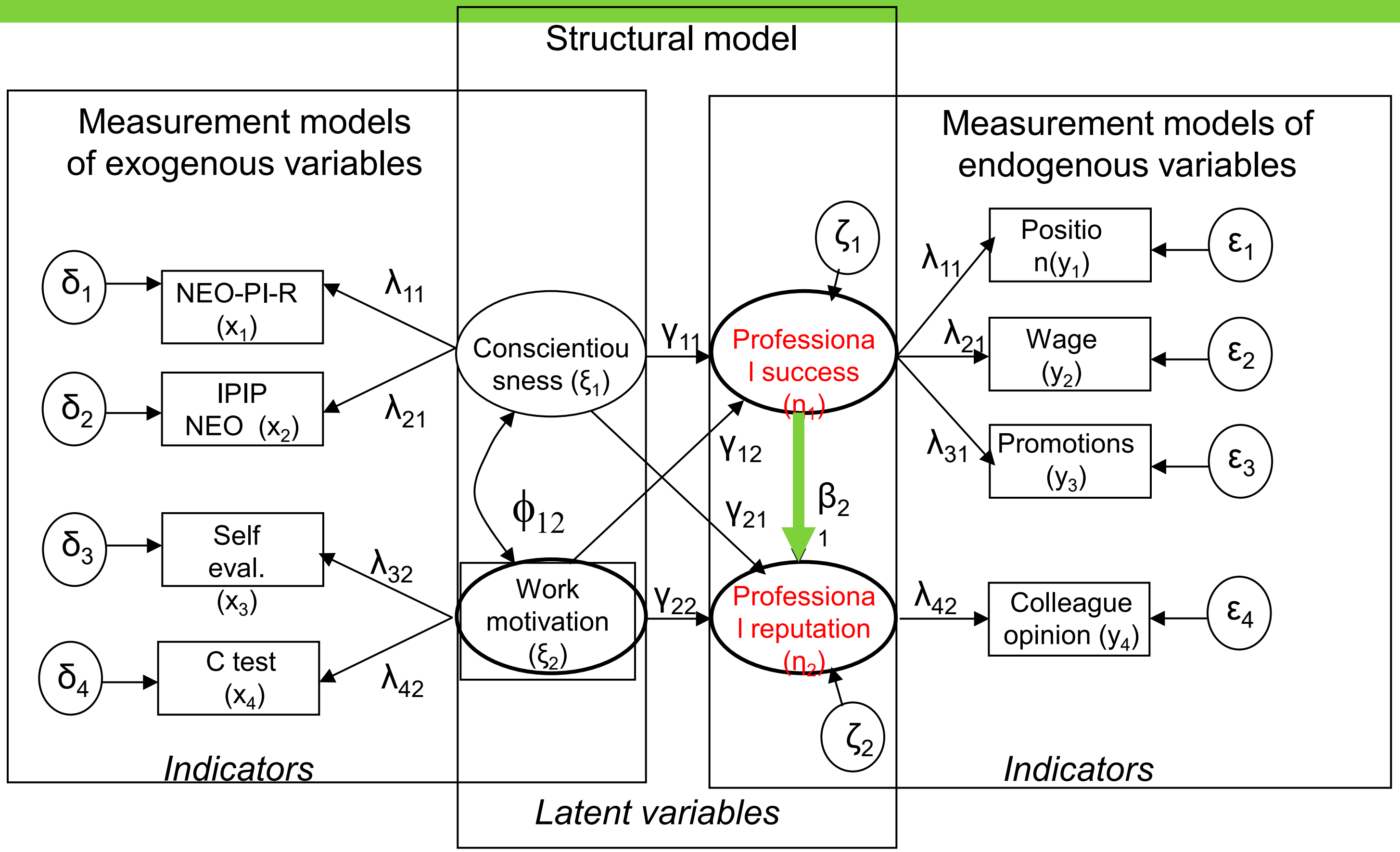
$r_{x2,y} = .6$

Path analysis

- Note:

Formally the contribution of path analysis to SEM is apparent as a new parameter matrix:

the BETA matrix


Structural model
Measurement models of exogenous variables
Measurement models of endogenous variables
δ1
δ2
δ3
δ4
NEO-PI-R (x1)
IPIP NEO (x2)
Self eval. (x3)
C test (x4)
λ11
λ21
λ32
λ42
Conscientiousness (ξ1)
Work motivation (ξ2)
φ12
γ11
γ12
γ21
γ22
ζ1
ζ2
Professional success (η1)
Professional reputation (η2)
β21
λ11
λ21
λ31
λ42
Position (y1)
Wage (y2)
Promotions (y3)
Colleague opinion (y4)
ε1
ε2
ε3
ε4
Indicators
Latent variables
Indicators

Path analysis

- Note 1:

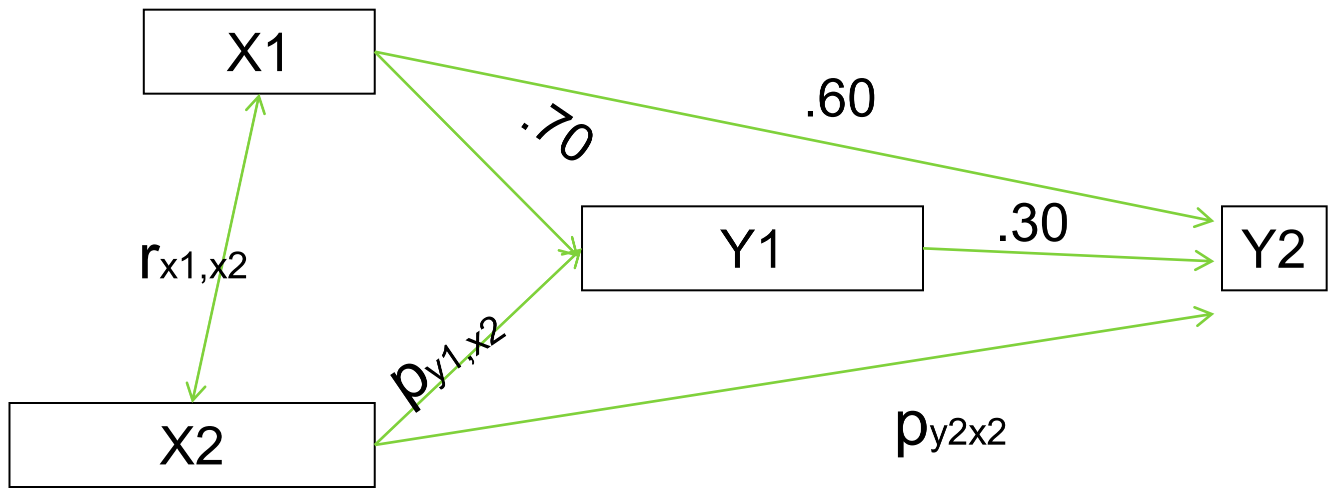
 In simple structural models multiple regression and path analysis do equally well.

But a complex path model can includes structural unites that <u>cannot be realized </u>as a multiple regression model.

Because of path analysis new parameters become available: these are measures of indirect effects.

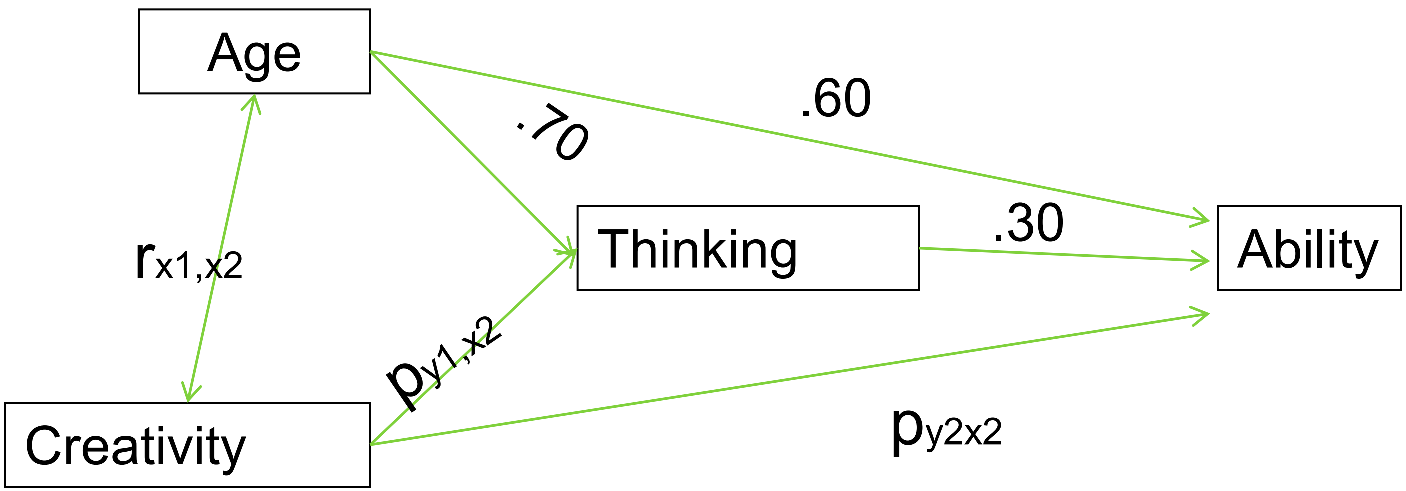
Path analysis: effects

- **Effect estimation :**
  - Direct effects *correspond* to path coefficients
  - Indirect effects are estimated by *multiplying* path coefficients



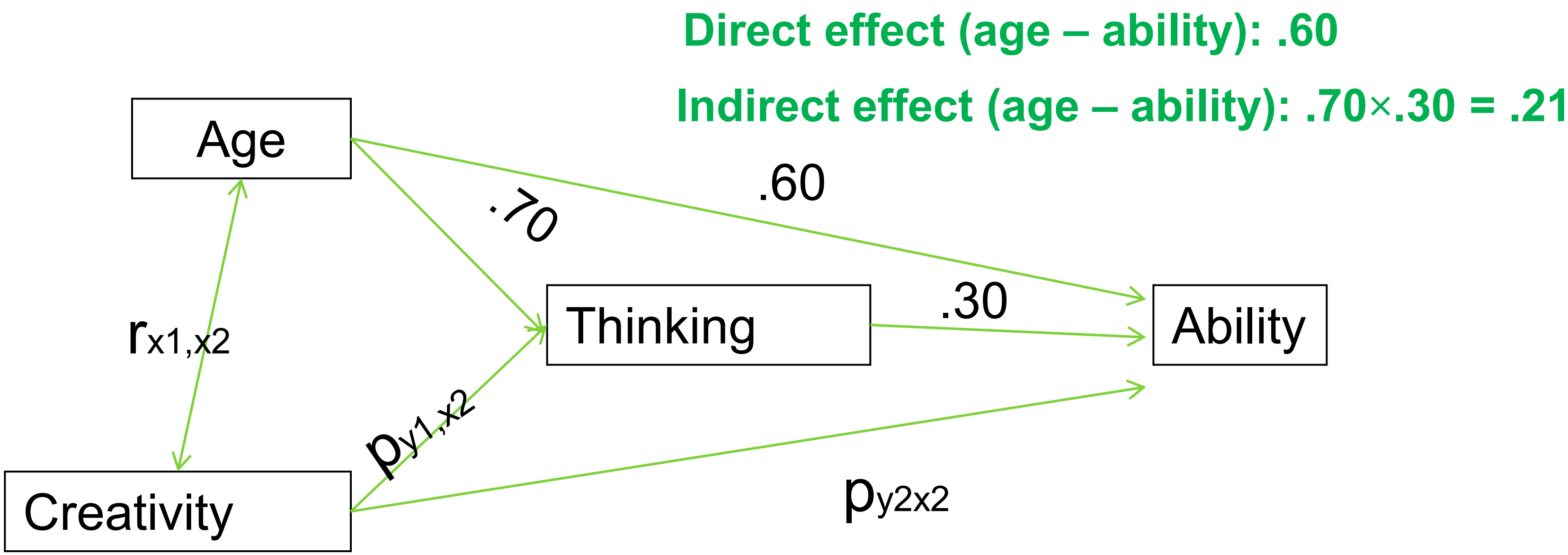
Path analysis

- **<u>Example with path coefficients added:</u>**
  - Age (X1) ; creativity (X2)
  - Thinking (Y1) ; problem solving ability (Y2)



Path analysis

- **Example with path coefficients added:**
  - Age (X1) ; creativity (X2)
  - Thinking (Y1) ; problem solving ability (Y2)

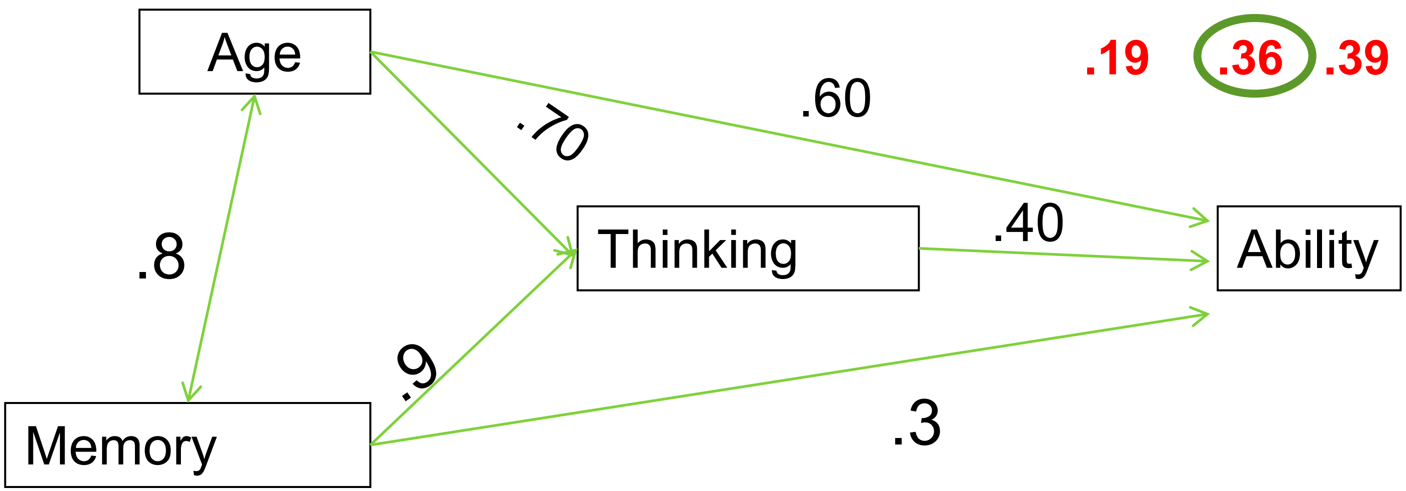


Path analysis *practice*

- <u>Example with path coefficients added:</u>
  - Age (X1) ; memory (X2)
  - Thinking (Y1) ; problem solving ability (Y2)

Determine the indirect effect of memory on <u>ability!</u>

.19 .36 .39 .41



Path analysis

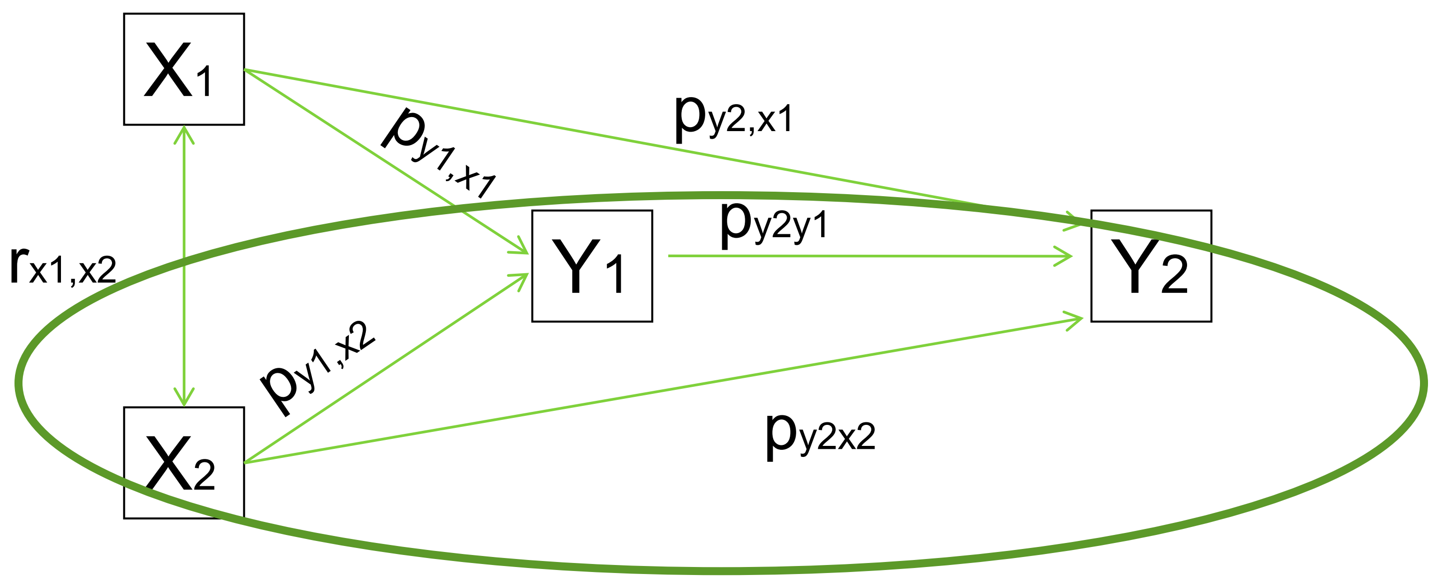
Consequences for parameter estimation:

- Parameter estimation can be realized as search for the solution(s) of a system of linear equations (i.e. in a somewhat different way as in regression analysis)

- Even in complex structures parameter estimation can be conducted on the basis of correlations

Path analysis: application

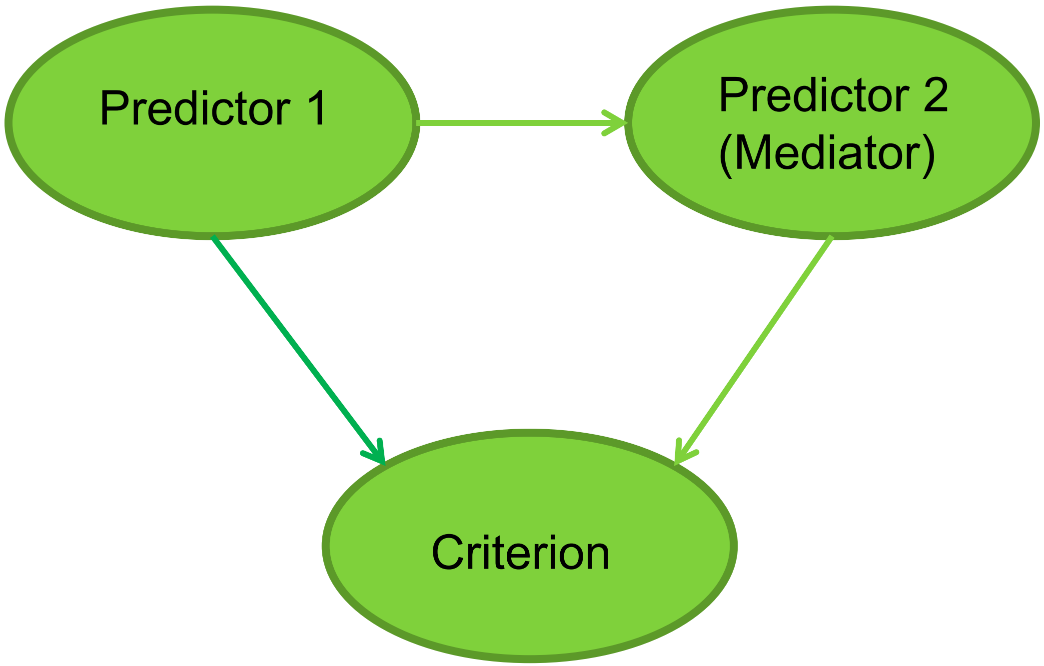
- In application frequently the focus is on *parts* of the complete simple path model:



Path analysis: application

- In application frequently the focus is on *parts* of the complete simple path model:

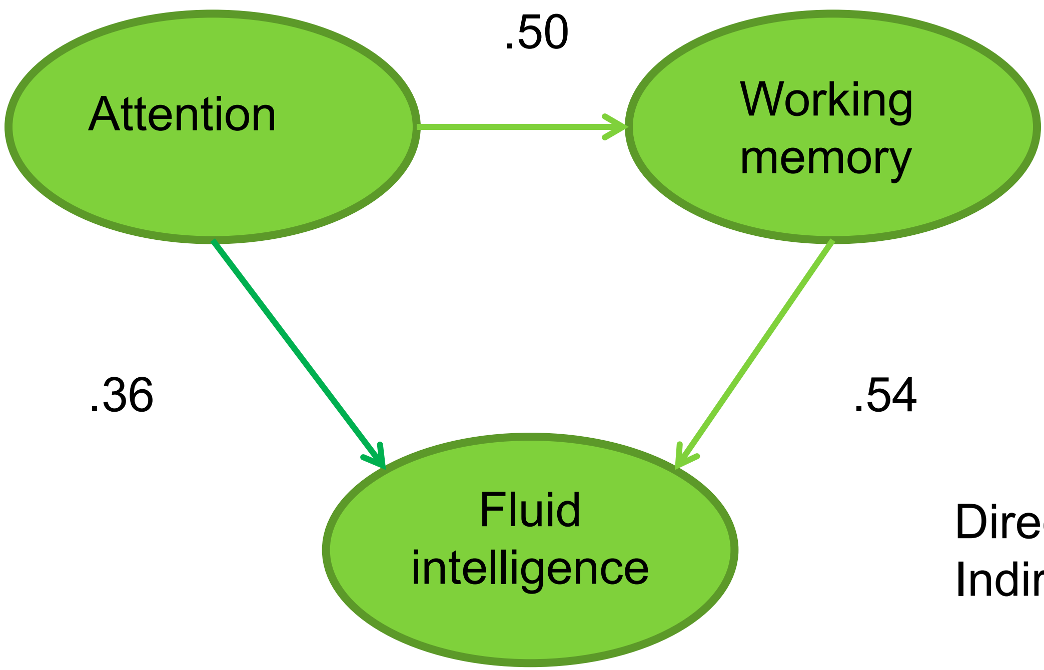
Model for testing the hierarchy of predictors



Path analysis: application

- In application frequently the focus is on *parts* of the complete simple path model:

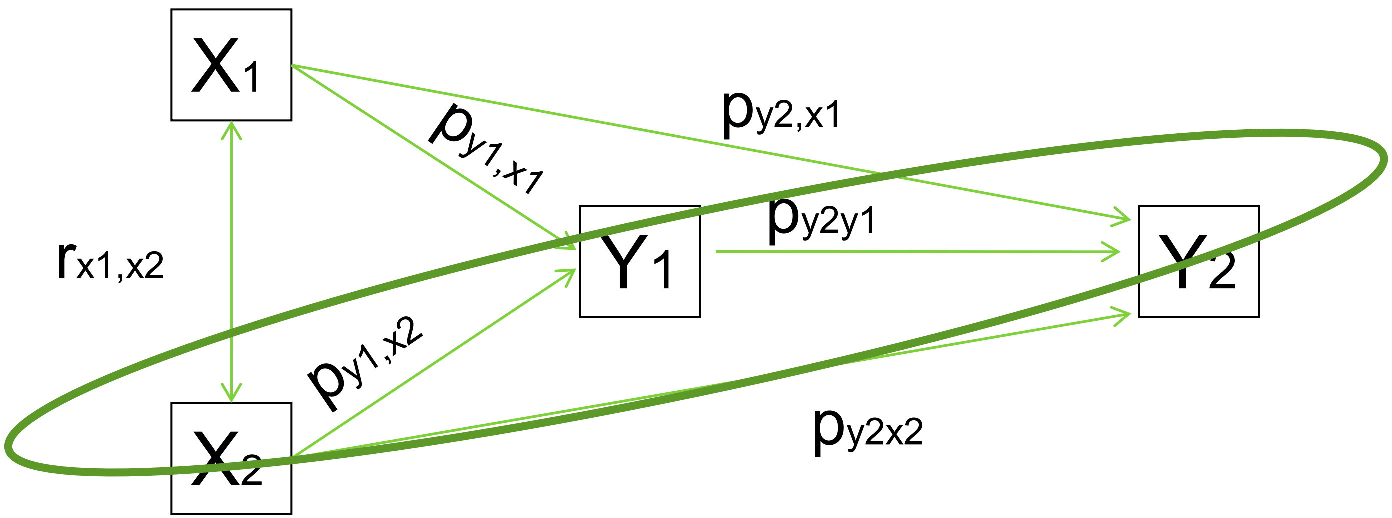
e.g.



Direct effect: .36
Indirect effect: .27

Path analysis: application

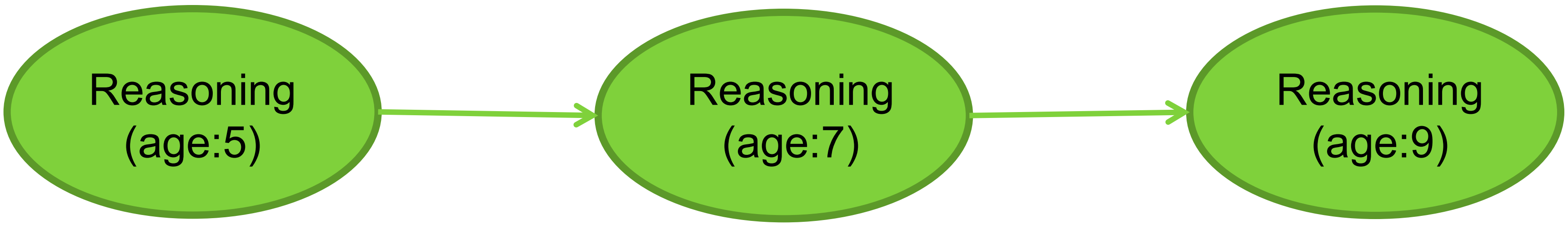
- In application frequently the focus is on *parts* of the complete simple path model:



Path analysis: application

- In application frequently the focus is on *parts* of the complete simple path model:

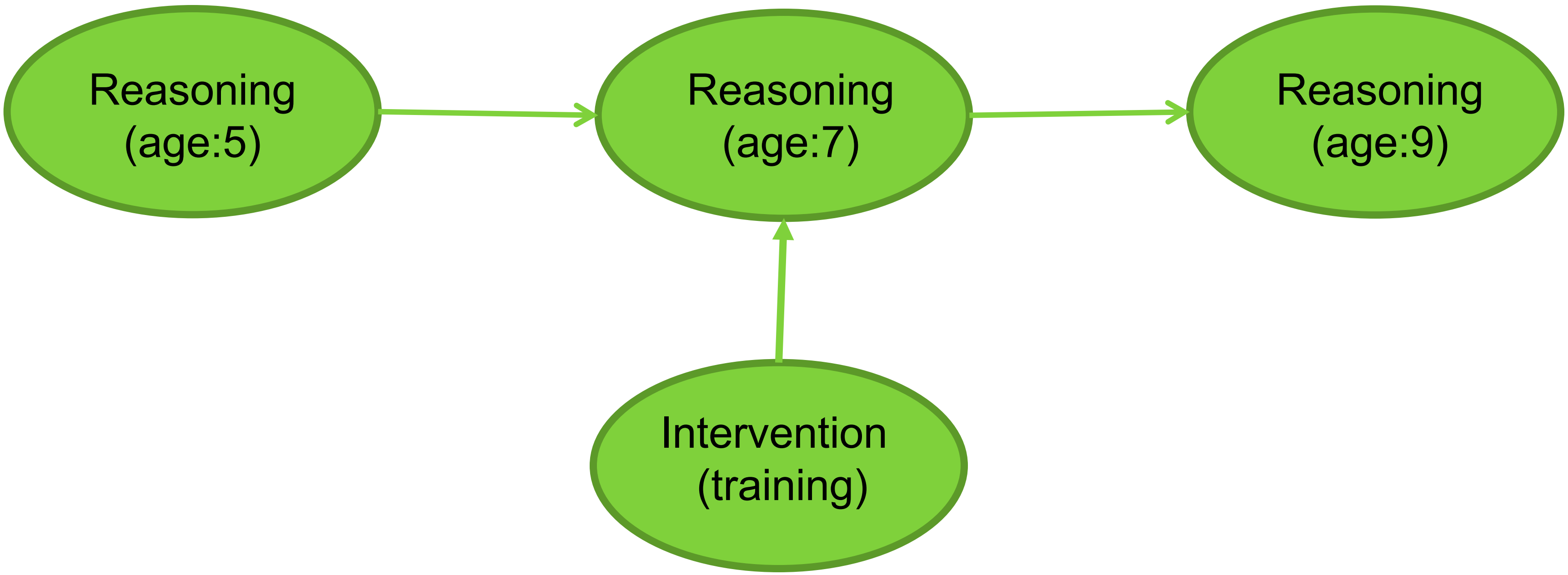
Developmental model (mediation model)



Path analysis: application

- In application frequently the focus is on *parts* of the complete simple path model:

Developmental model (mediation model)



In sum

Path analysis brings about new structural feature

.....

... increases the complexity of models that can possibly be investigated

.... enable investigations on the basis of covariances and correlations

The causality problem

Path analysis: the causality problem

- The „causality“ problem of SEM
  - There is disagreement among SEM researchers regarding the status of the outcome of SEM:

    Does SEM inform us of

    - a causal consequences

    or - effects / influences?

    e.g. .. high working memory capacity ***causes*** high intelligence

    e.g. .. working memory capacity ***influences*** the degree of intelligence

Path analysis: the causality problem

- Problem „causality“
  - The terms „causality“ and „causal“ are occasionally used for the relationships at the latent level

  … ***but what does „causality“ mean?***

  - turning the switch for the light „causes“ the light bulb to illuminate the room

It is correct, that the „turning of the switch“ causes this specific event!

But taking a *general perspective minimizes* the importance of the „turning of the switch“

… because it would not work if there were no **electricity**

Path analysis: the causality problem

- Problem „causality“
  - The terms „causality“ and „causal“ are occasionally used for the relationships at the latent level

… ***but what does „causality“ mean?***

- it may be argued: „the cause for student X to solve an arithmetic task is (a high degree of) intelligence“

But taking a general perspective minimizes the importance of „intelligence “

… because solving the task is not possible without **reading ability, knowledge of rules, …**

Path analysis: the causality problem

- Problem „causality“
  - The terms „causality“ and „causal“ are occasionally used for the relationships at the latent level
  - In science there is the established opinion that **experimentation** is necessary in order to justify the assumption of causality

    - experimentation marked the *beginning of scientific psychology*

    - experimentation includes …
      - the willful manipulation of the phenomenon
      - several treatment levels
      - the control of random influences
      - there is random assignment

Path analysis: the causality problem

- Problem „causality“
  - The terms „causality“ and „causal“ are occasionally used for the relationships at the latent level
  - Poponents of the causality assumptions argue …

    … the absence of error or other influences at the latent level in SEM guarantee that there is only one source of effect that is the exogeneous variable

    … so that experimentation is not necessary

Path analysis: the causality problem

- Problem „causality“
  - The terms „causality“ and „causal“ are occasionally used for the relationships at the latent level

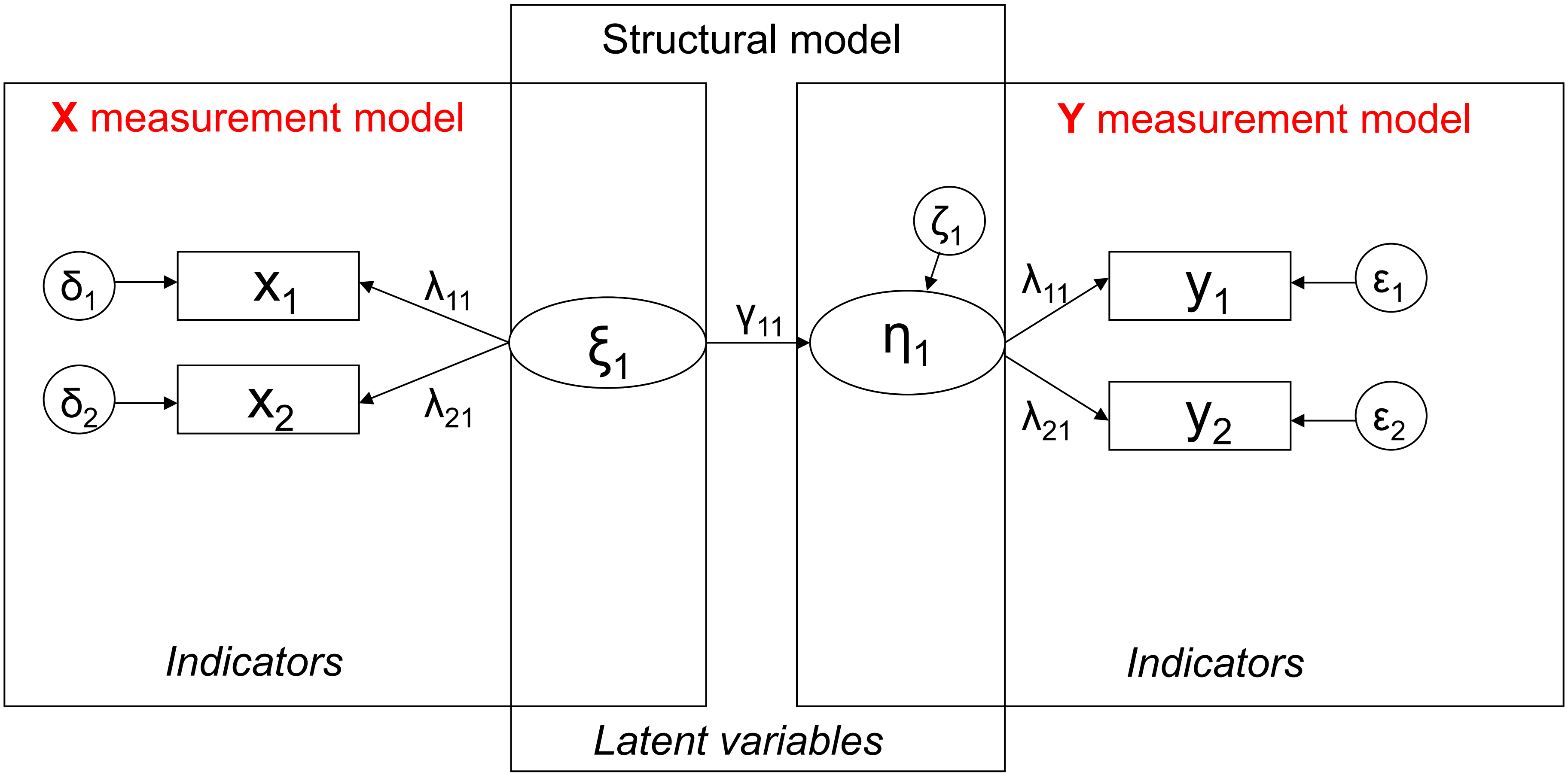
  - Their opponents argue …

    … that the denotation as „**directional influence**“ or „**effect**“ is much more appropriate since there is usually *no perfect relationship* between predictor and criterion variables

Path analysis: the causality problem

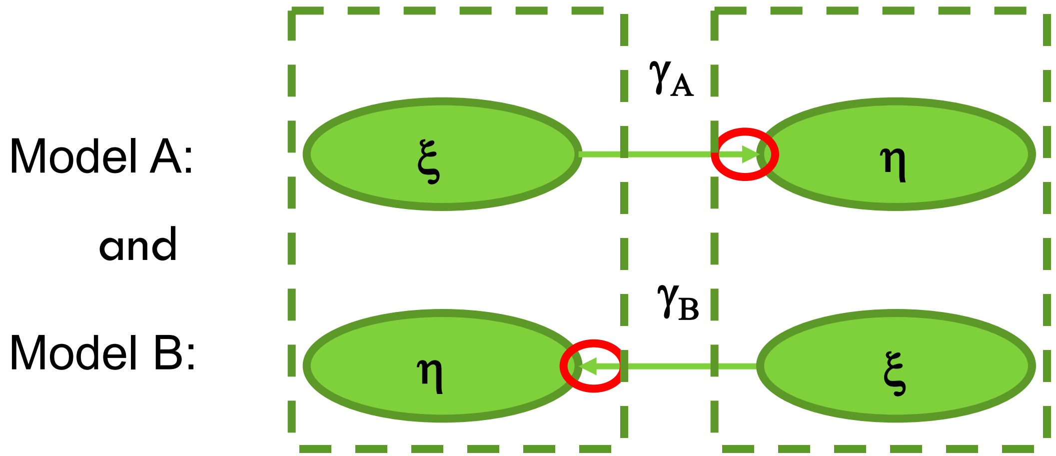
- … the direction of the influence may not even be reflected by the regression weight *correctly* (regarding the question which variable influences the other variable)

Path analysis: the causality problem



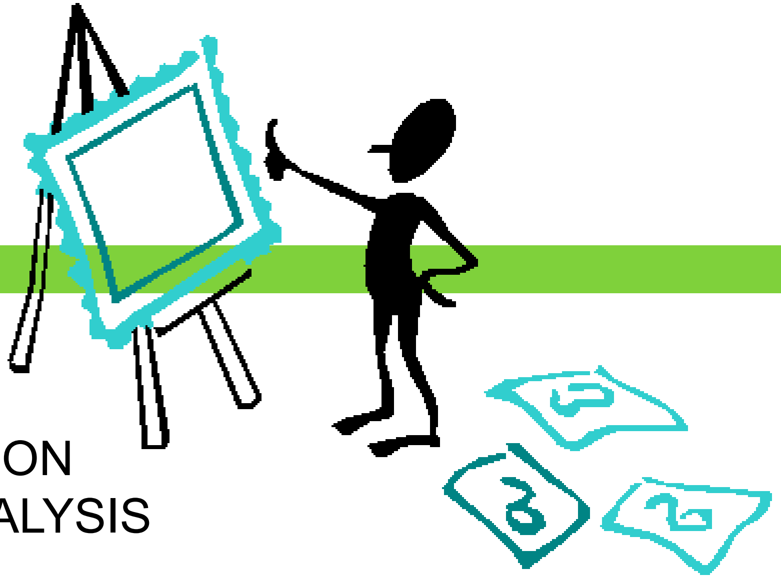
Path analysis: the causality problem

- Further problems in SEM
  - *In some cases there is even the situation that*



i.e. both models lead to the same parameter estimates and model fit: $\gamma_A = \gamma_B$

Summary and brush up:



CORRELATION

PARTIAL CORRELATION

MULTIPLE REGRESSION

PATH ANALYSIS

Contribution to SEM ...?

- 1. ......
- 2. Path … an assumed way of influence
- 3. Mediation … indirect influence
- 4. Decomposition … subdivision into effects
- 5. Effect … influence

QUESTIONS REGARDING COURSE UNIT 7

- What is a path coefficient?
- What is a mediation variable?
- What kinds of effects are considered in path analysis?
- How many indicators are required in path analysis?

Literature

- Kline, R. B. (2011). *Principles and practices of structural equation modeling* (3rd edition) (Chapter 2: Basic statistical concepts: correlations and regression, Chapter 5: Introduction to path analysis). New York, NJ: The Guilford Press.
- Cohen, J., Cohen, P., West, S. G. & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (Partial regression, partial correlation, direct and indirect effects: S. 64-79). Mahwah, N.J.: Lawerence Erlbaum.