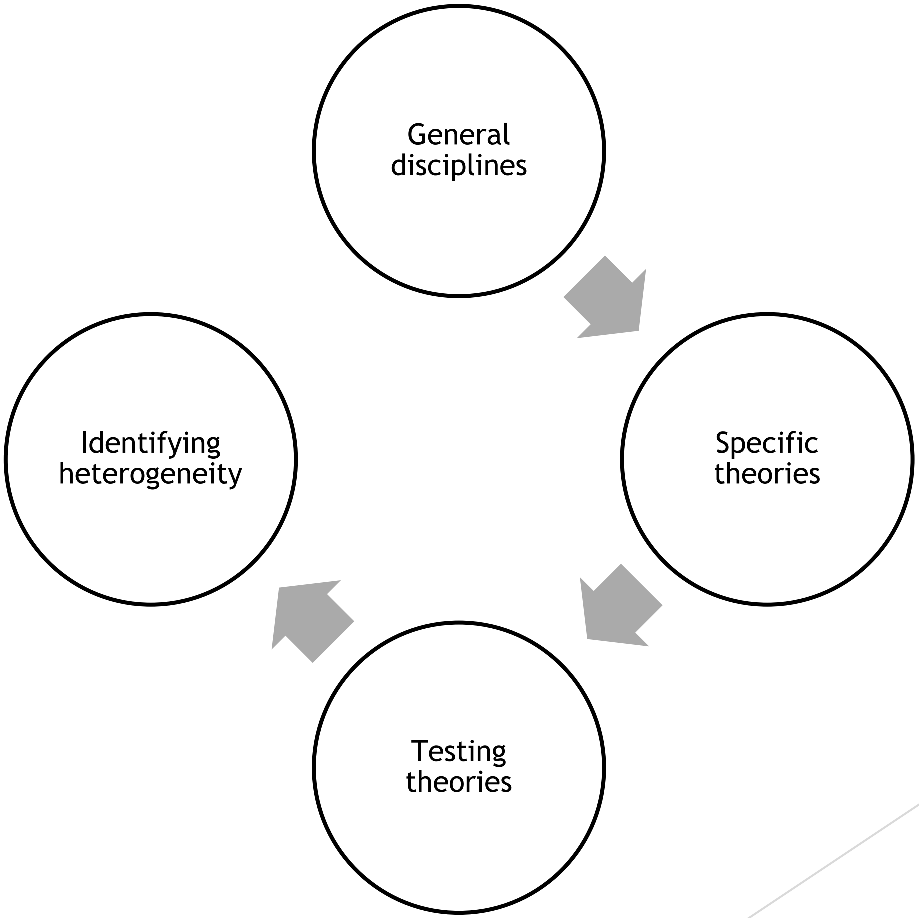
Advanced Theories of Consumer Research



J. Prof. Dr. Dominic Lemken

Academic Career

- PhD at the University of Göttingen:
    - Thesis: „The Adoption of Legumes in Farmer and Consumer Settings“ (2017)
- Professorship for „Socioeconomics of Sustainable Nutrition“, University of Bonn (since 2022)

Research Topics

- Nutrition Policy
- Food choice behavior

Research projects

- Key Food Choices and Climate Change
- Strategien für mehr heimischen Leguminosenkonsum
- MensaPlus

A guide to structural equation modelling

Outline:

- Morning: theory on structural relationships and latent concepts
- Afternoon: SEM-Tutorials 1-6

Course material:

- Software: Adanco 2.3.2 via: https://download.compositemodeling.com/
- Course files on studIP (sub-folder SEM), otherwise: https://github.com/dlemken/Structural-Equation-Modeling

A guide to structural equation modelling

- Adanco licenases are distributed during the first break

Melden Sie sich bitte mit Ihrem Namen an

Wir freuen uns, ihr Gesicht zu sehen!

Schalten Sie Ihr Mikro stumm, wenn Sie nicht sprechen

Nutzen Sie die Chatfunktion für Rückfragen, Meldungen oder Kommentare

A guide to structural equation modelling – how to follow

- Video summary of theoretical part
- Slides are made available
- Tutorials are made available (correct answers in the slides)
- Course conceptualized as a free of math guide
- Often more than one term for the same concept
- Slido.com: #548676

A guide to structural equation modelling – how to follow Bibliography

Fornell, C. (Ed.). (1982). *A Second Generation of Multivariate Analysis: Measurement and Evaluation*. New York: Praeger.

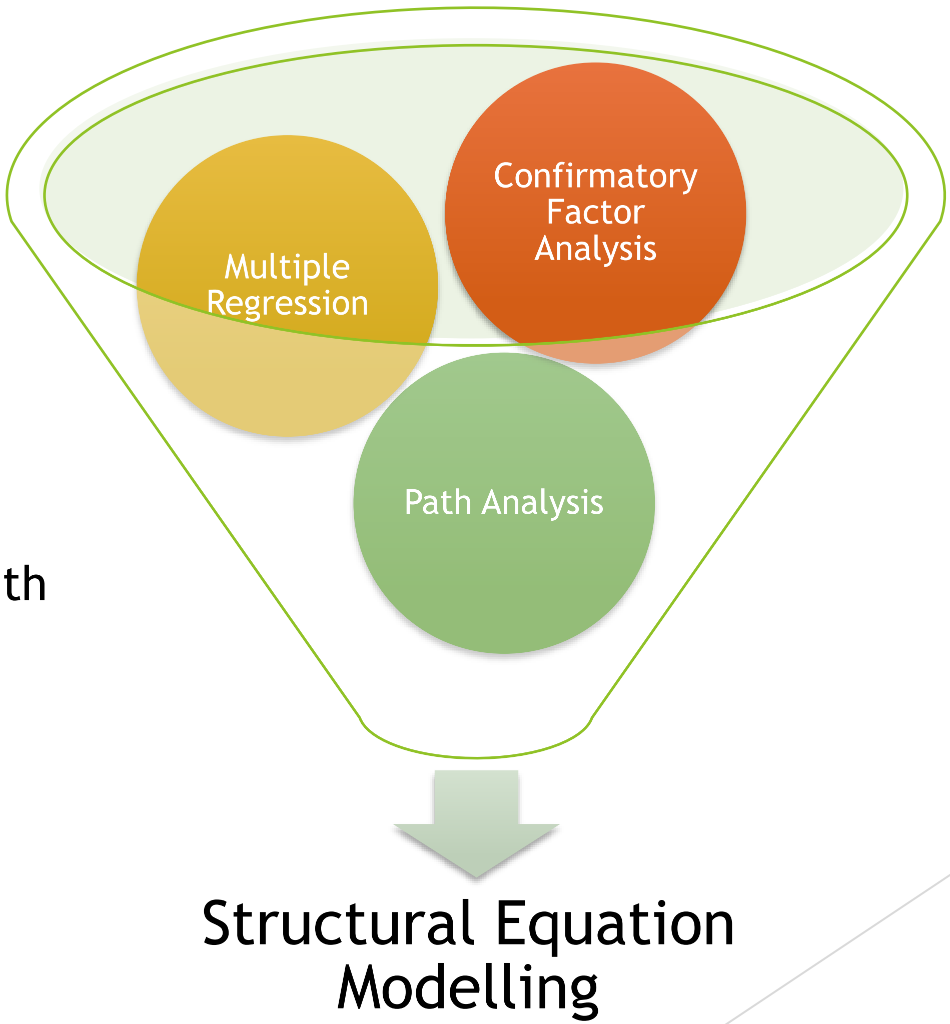
Henseler, J. (2020). *Composite-Based Structural Equation Modeling: Analyzing Latent and Emergent Variables*, New York: Guilford Press.

Ullman, J. B., & Bentler, P.M. (2003). Structural equation modeling. In I. B.Weiner, J. A. Schinka, & W. F. Velicer (Eds.), *Handbook of Psychology: Research Methods in Psychology* (Vol. II, pp. 607–634). Hoboken, NJ: Wiley.

Henseler, J., & Schuberth, F. (2020). Auxiliary Theories. In J. Henseler: *Composite-Based Structural Equation Modeling: Analyzing Latent and Emergent Variables*, New York: Guilford Press, pp. 25–37.

Henseler, J. (2017). Bridging design and behavioral research with variance-based structural equation modeling. *Journal of Advertising*, 46 (1), 178–192.

Structural Relationships and Latent Concepts



Experience with each?

Family Tree of SEM

SEM's Core Task:
Theory Testing

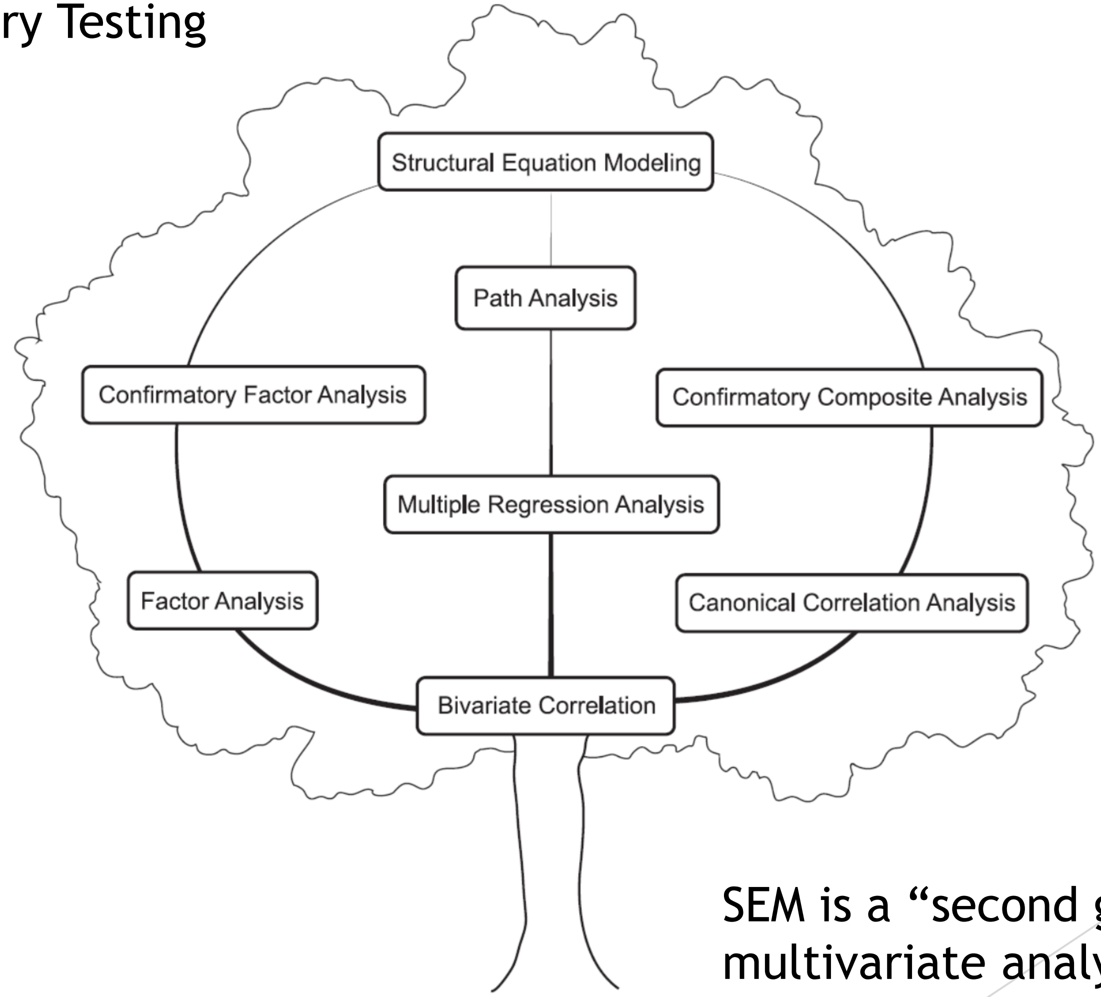
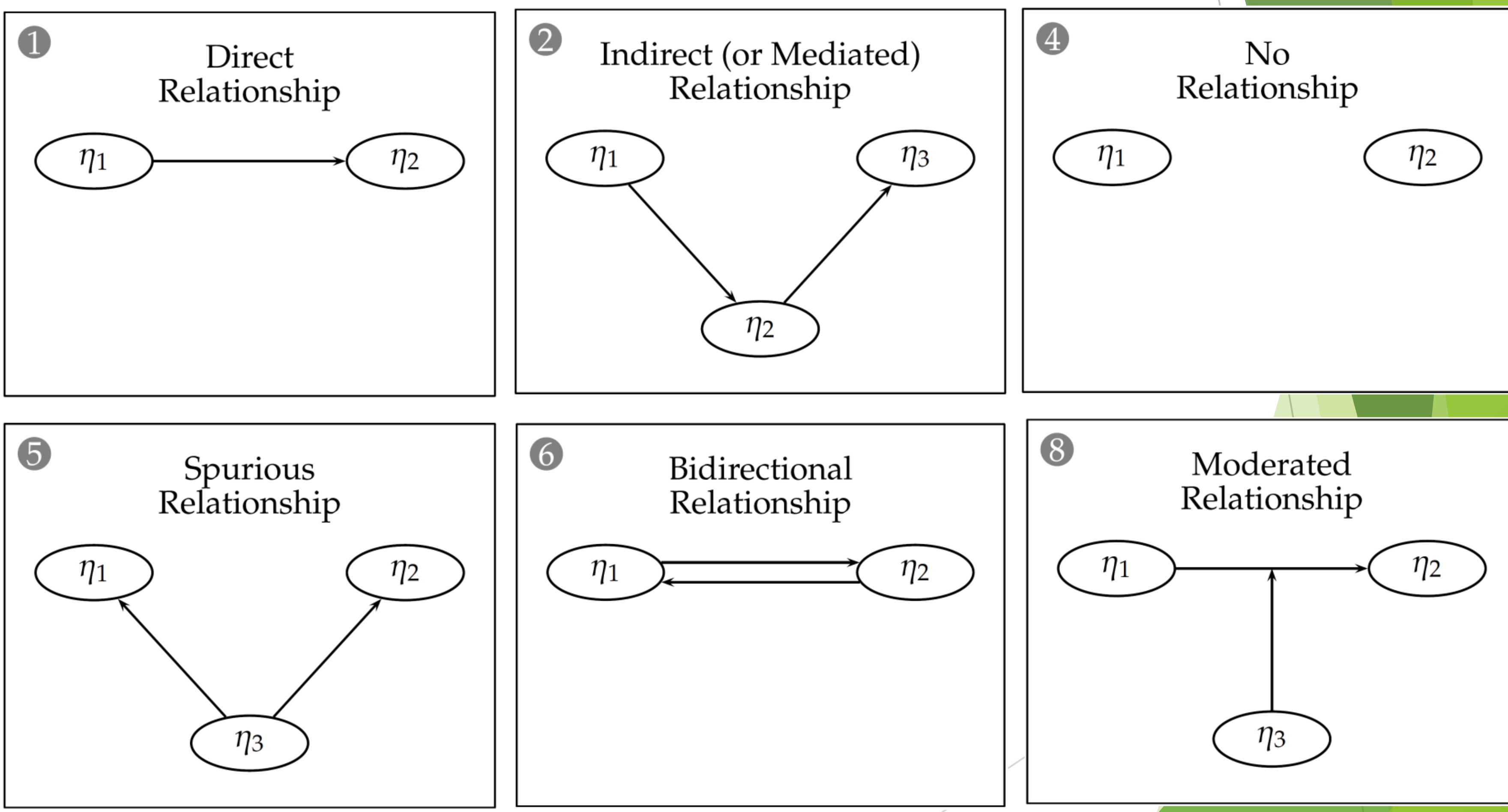


FIGURE 1.1. Family tree of SEM.

SEM is a "second generation of multivariate analysis" (Fornell, 1982).

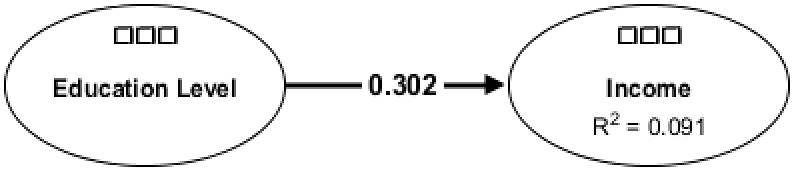
Structural Relationships

Path Analysis

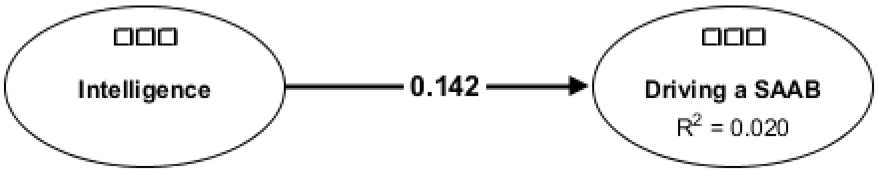


Structural Relationships

Research starts with a conscious or unconscious hypothesis:

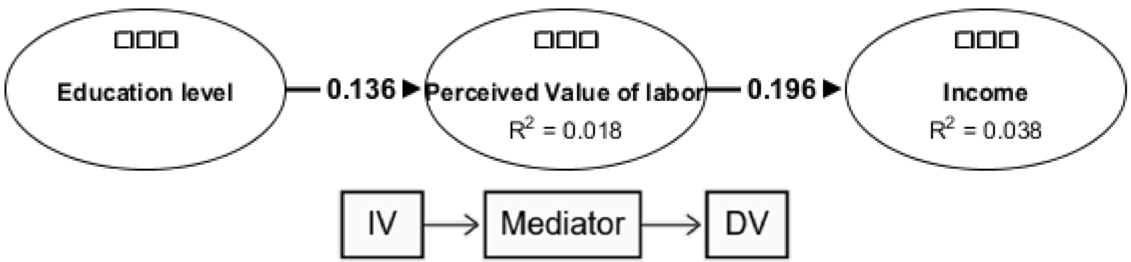


Recursive?



Causality?

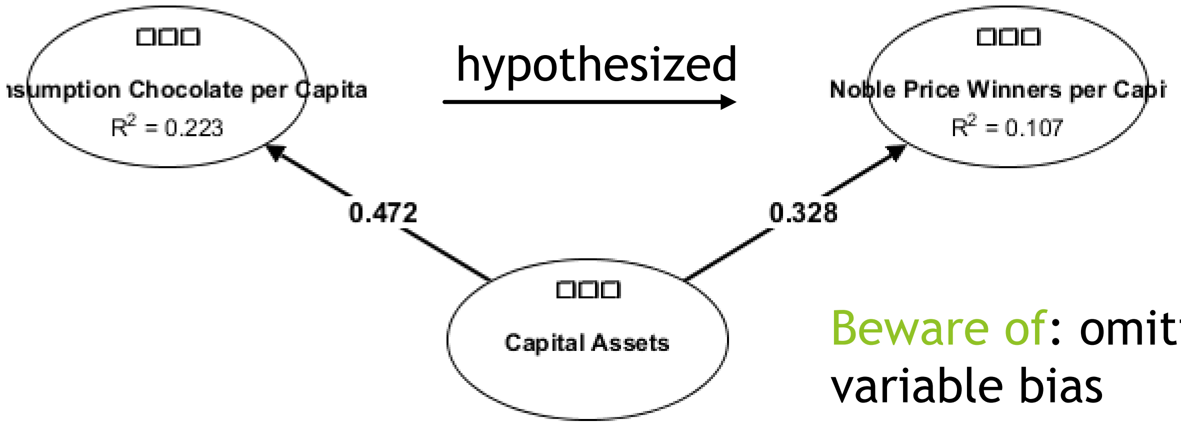
The effect mechanism / mediator



Structural Relationships

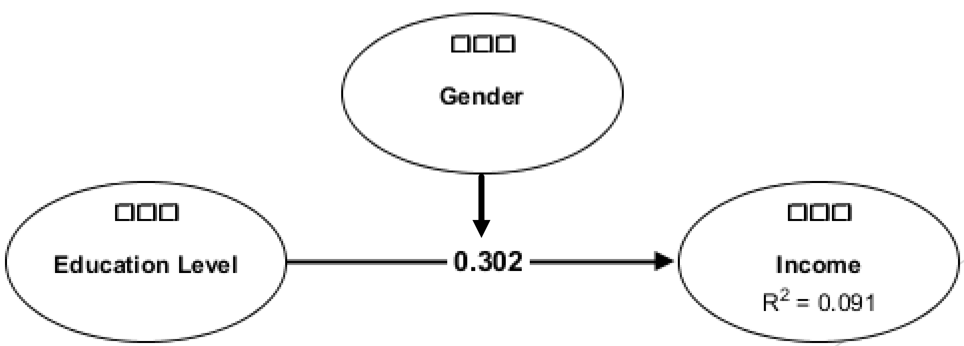
Path Analysis

Phantom or spurious Relationship: Revealing the need for a theoretical foundation



Beware of: omitted variable bias

Moderated Relationship: Interaction Effects



An Example of a Moderated Relationship

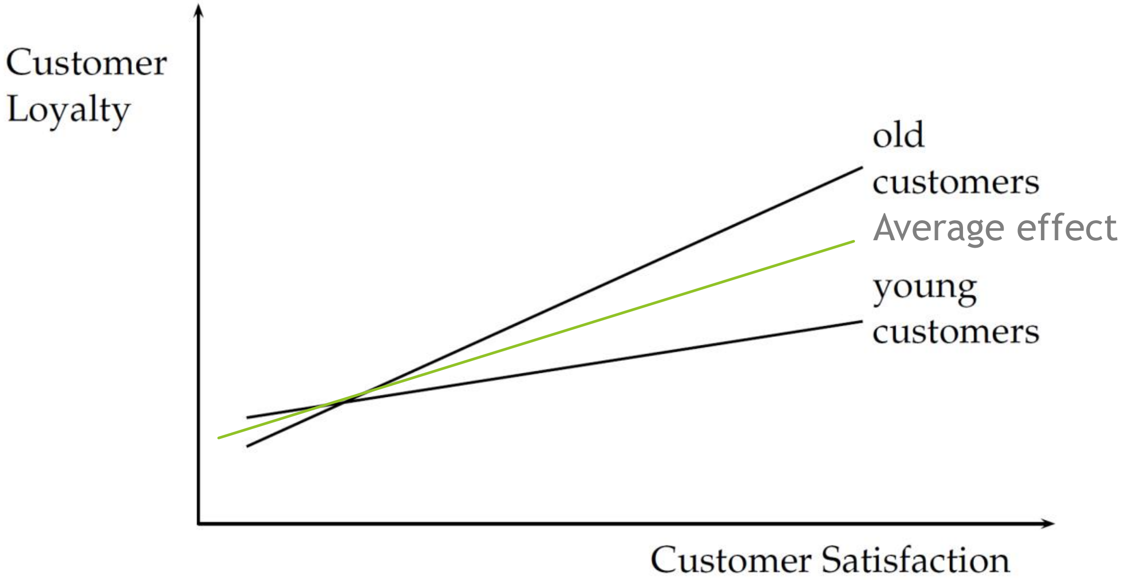


FIGURE 11.1. The moderating role of age on the relationship between customer satisfaction and customer loyalty.

Alternative Representation

Operationalizing a moderation effect

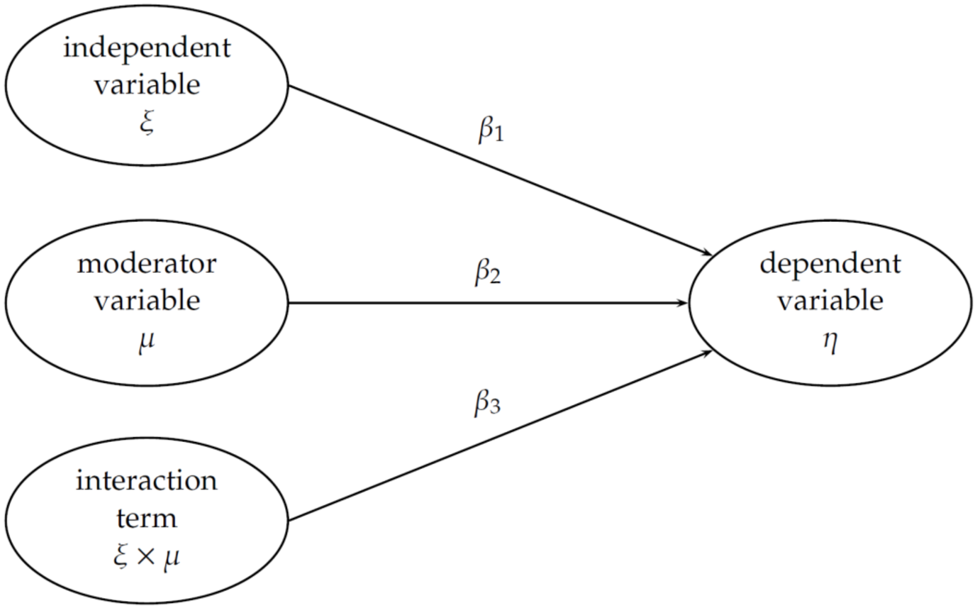
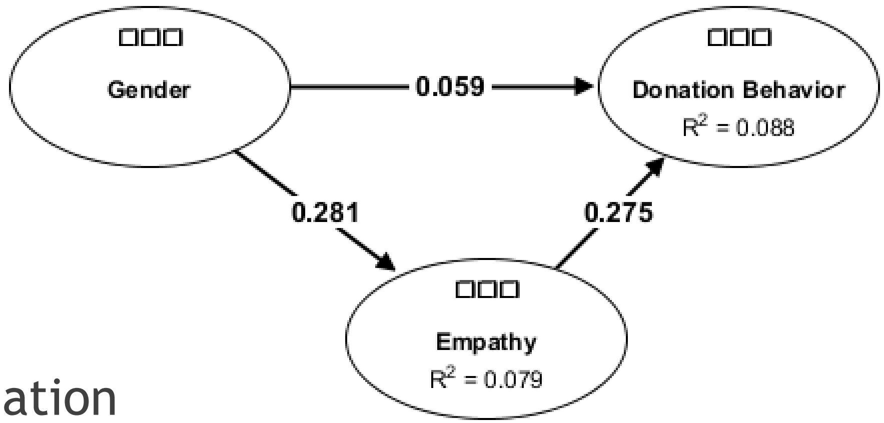


FIGURE 11.3. Alternative representation of an inner model with an interaction effect.

Indirect effect: Triangle of Mediation

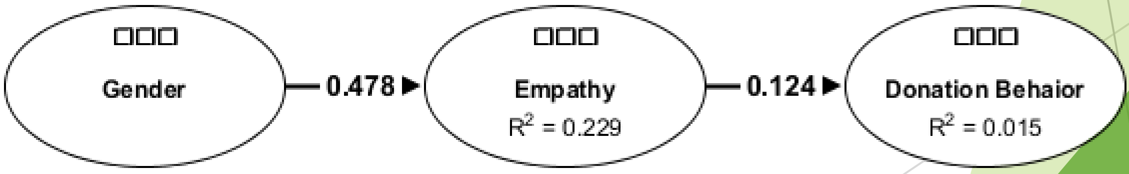
Path Analysis

Direct and Indirect Relationship: Indirect Relationships, the backbone of theories

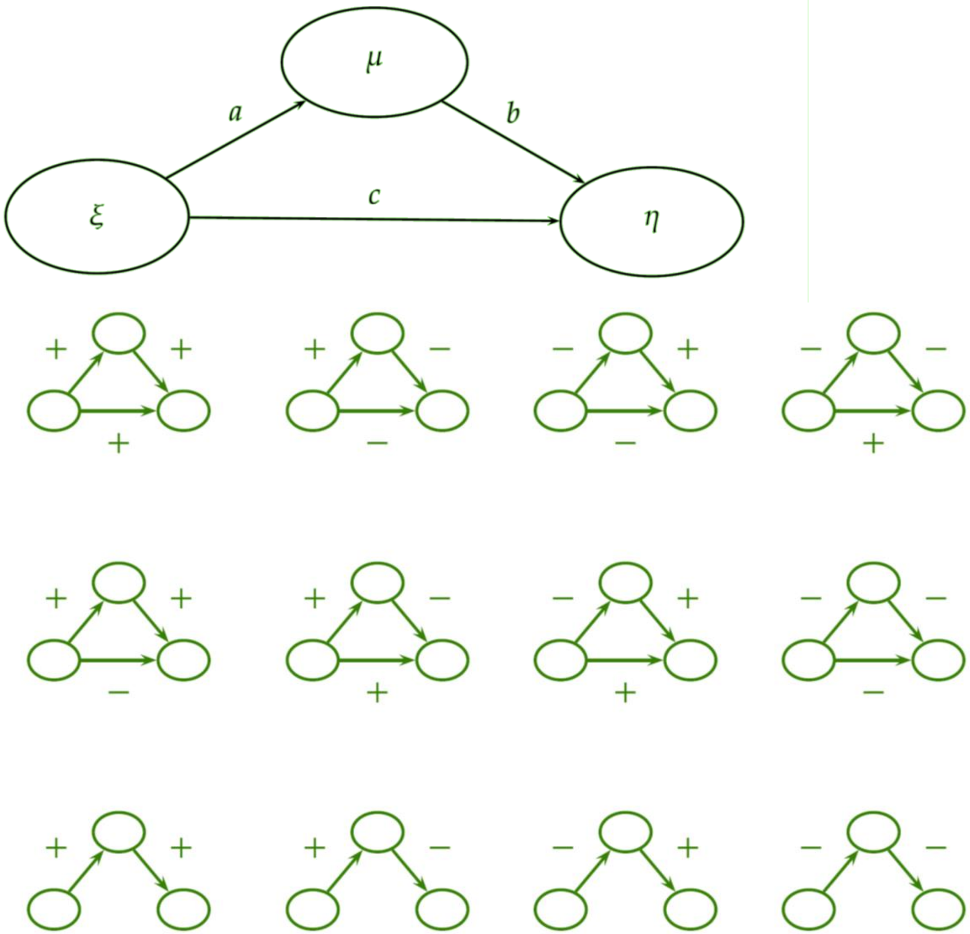


- Partial Mediation
  - Complementary
  - Competitive
- Full mediation
- No mediation

Willer, R. et al. (2015): What drives the gender gap in charitable giving? Lower empathy leads men to give less to poverty relief. In: *Social Science Research* 52, S. 83–98. DOI: 10.1016/j.ssresearch.2014.12.014



Classifying Mediation



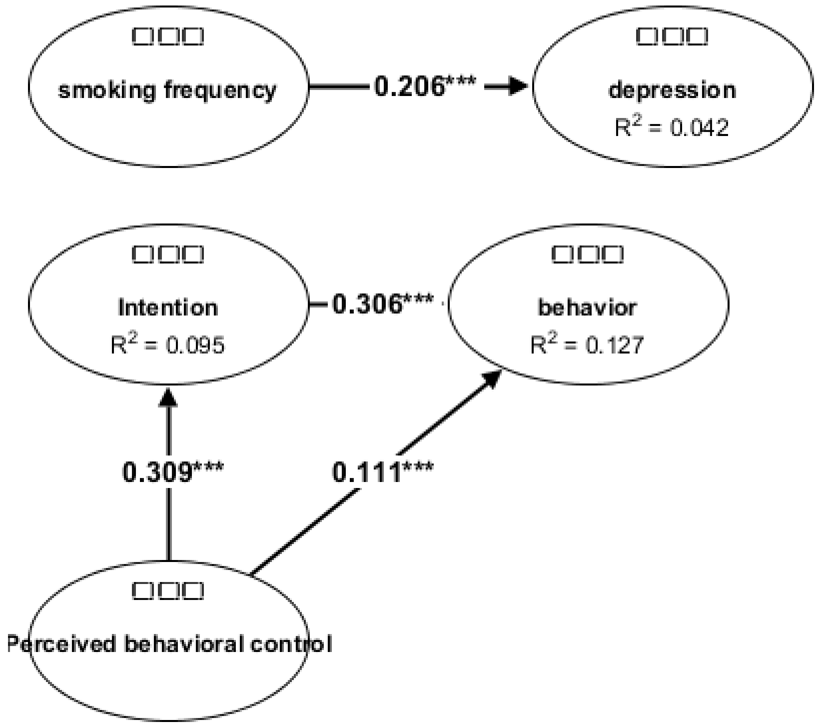
Patterns of complementary mediation:

Patterns of competitive mediation:

Patterns of full mediation:

Structural relationships Quiz

Quiz



Reverse Causality

Partial mediation

Quiz

Can you name an example for

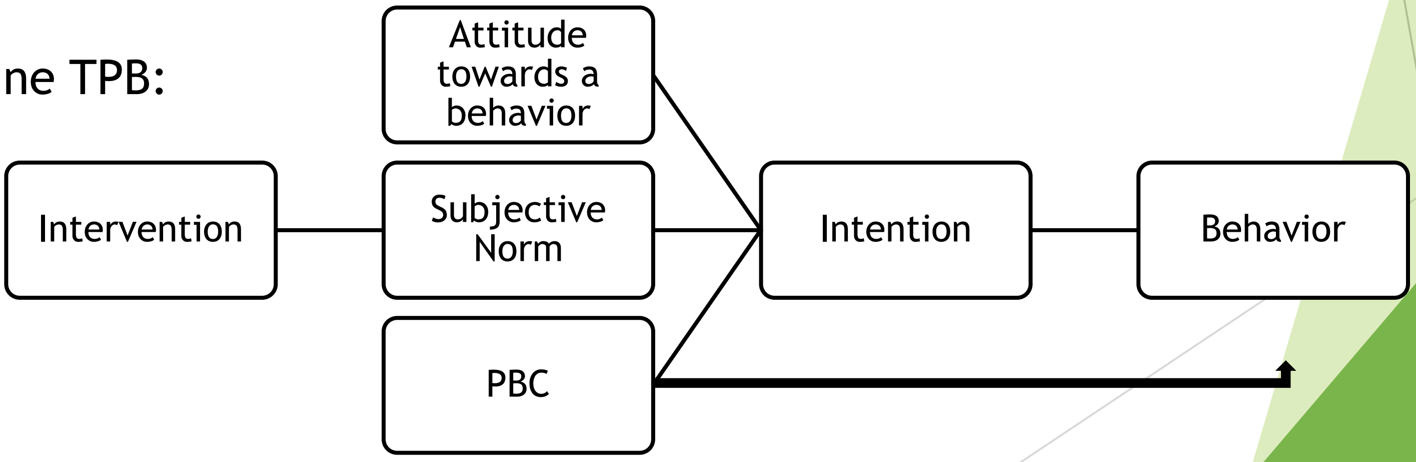
1. Mediation
2. Competitive mediation

Reason for SEM

Advantages of Path analysis:

- Identify initial effects, even though final outcome may not be significantly affected
- Understand effect mechanism
- Understand parallel relationships
- Understand mediator and moderator
- Testing entire theories

Imagine TPB:



Measuring Concepts: The Role of Auxiliary Theories

- Concepts (abstract ideas) are the building blocks of theories in business and social science.
- Auxiliary theories make concepts operational.
- Concepts are not directly observable:
  - Hidden variables e.g. Racism, quality of life, intelligence, etc.
  - Abstract concepts or mental states, e.g. confidence, extraversion, wisdom, Sustainable behavior in the food domain, welfare

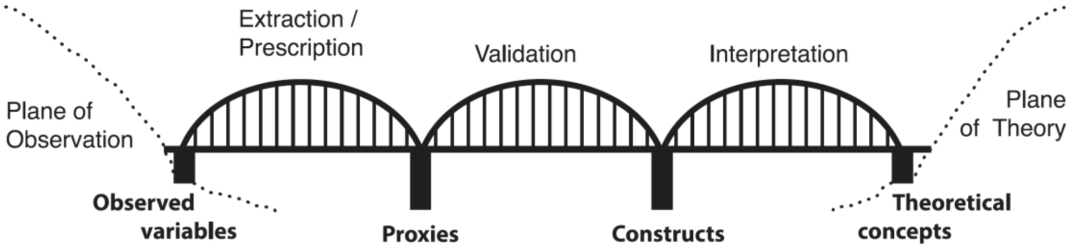


FIGURE 2.1. An auxiliary theory bridges a concept and its observed variables.

The concept of latent variables (LV)

Confirmatory Factor Analysis

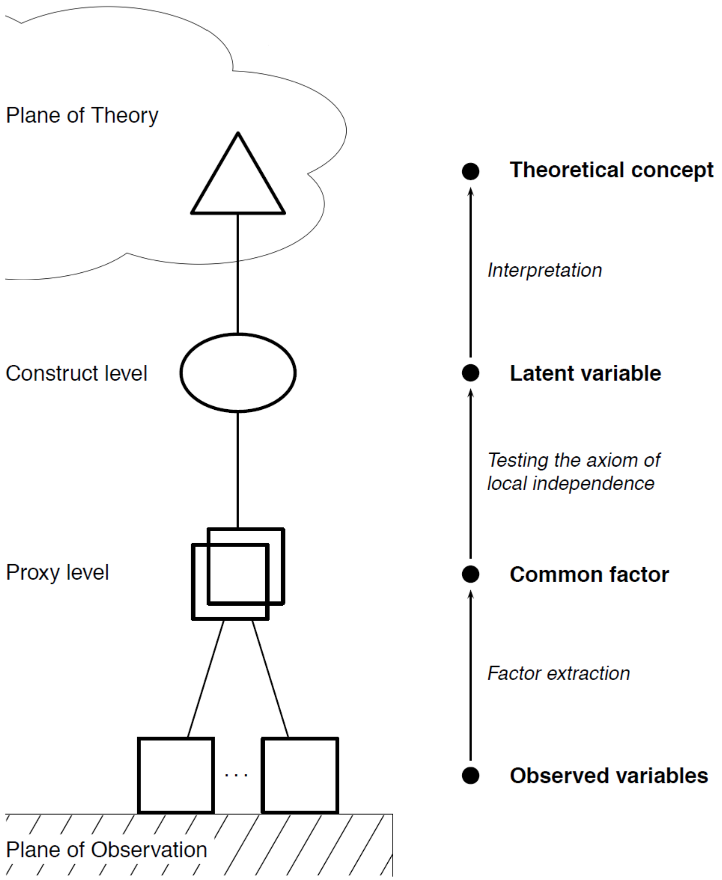


FIGURE 2.2. Anatomy of the measurement theory.

- Postulates the existence of a latent variable.
- The latent variable is the underlying cause of a block of observed variables.
- Realization: Reflective measurement models, in which latent variables are modeled as common factors.

Measuring Concepts

Confirmatory Factor Analysis

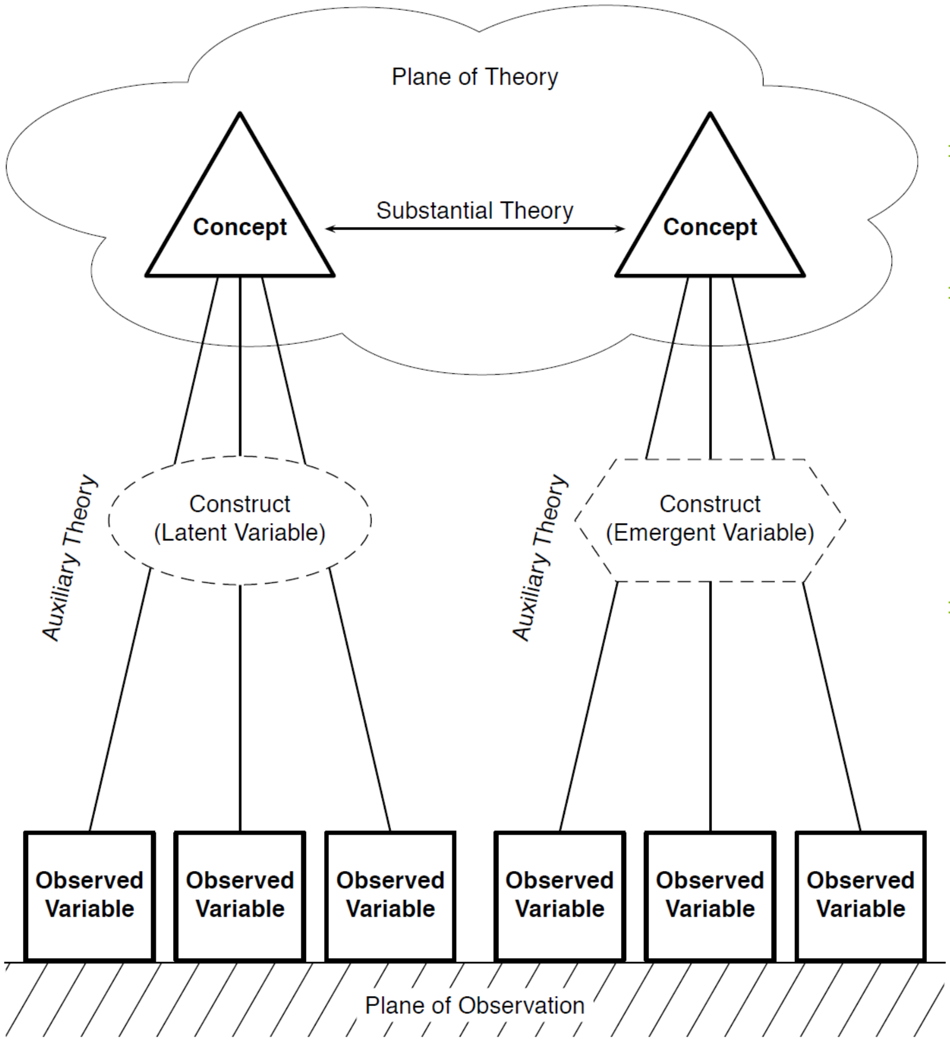
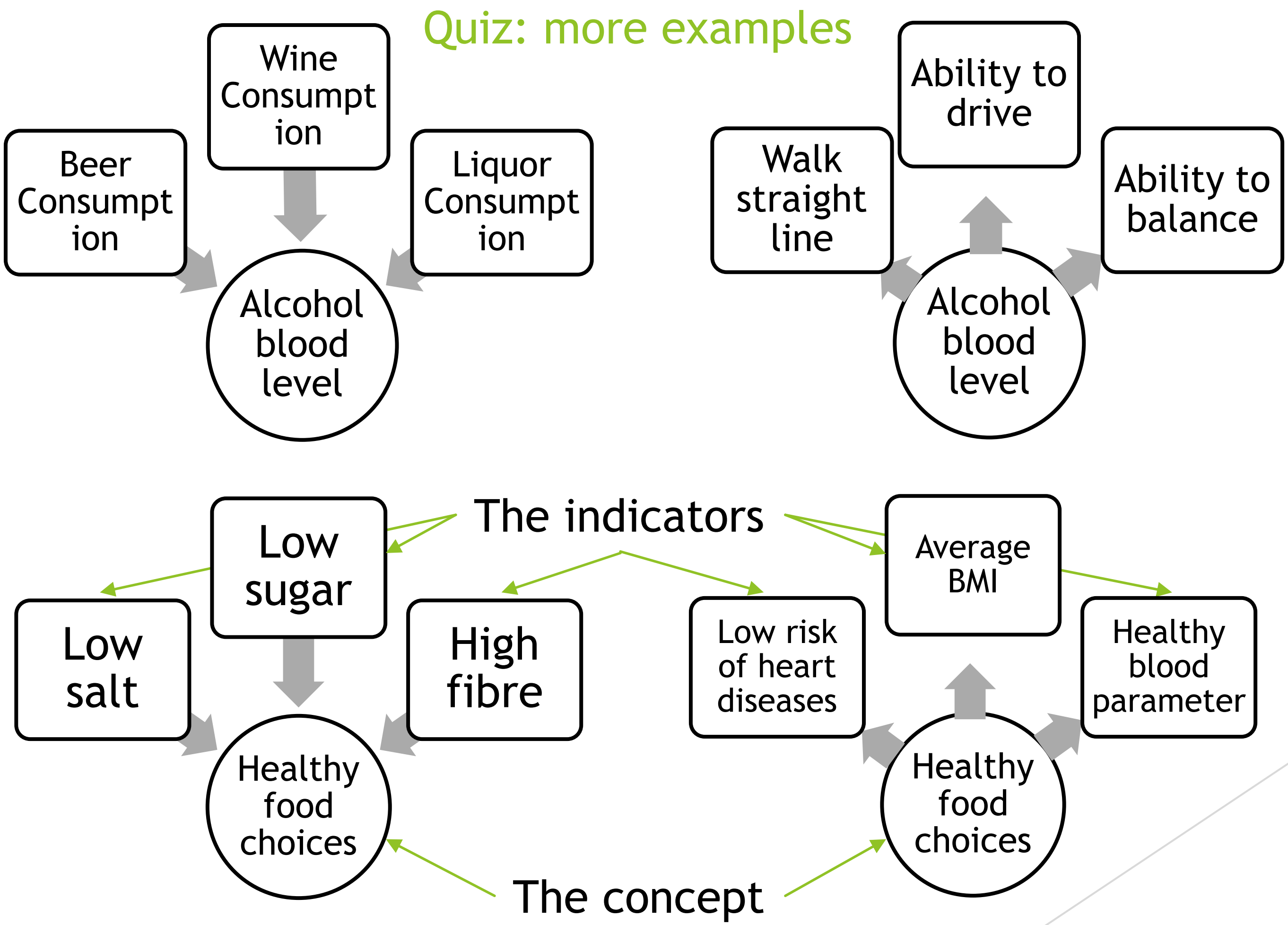


FIGURE 1.2. The structure of theory.

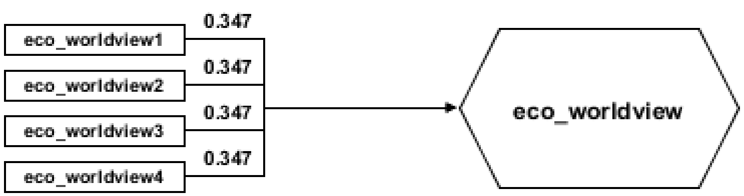
- ⇒ LV reduces dimensionality
- ⇒ LV lead to improved concept measurements, i.e. less measurement error
- ⇒ LV leads to inter-individual, i.e. generalizable concepts

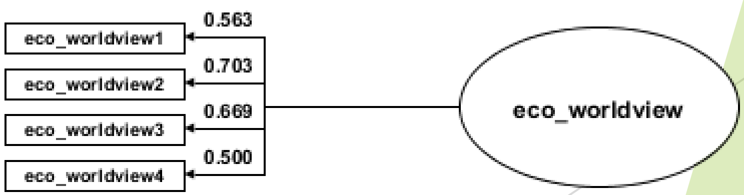
formative vs. reflective concepts



Latent Variables: formative vs. reflective

Formative	Reflective
Direction of causality is from indicator to LV	Direction of causality is from LV to indicator
No reason to expect indicators to be correlated	Indicators (at least 3 for identification) should be correlated
Dropping an indicator from the measurement model increases error of LV measurement	Dropping an indicator may not induce error





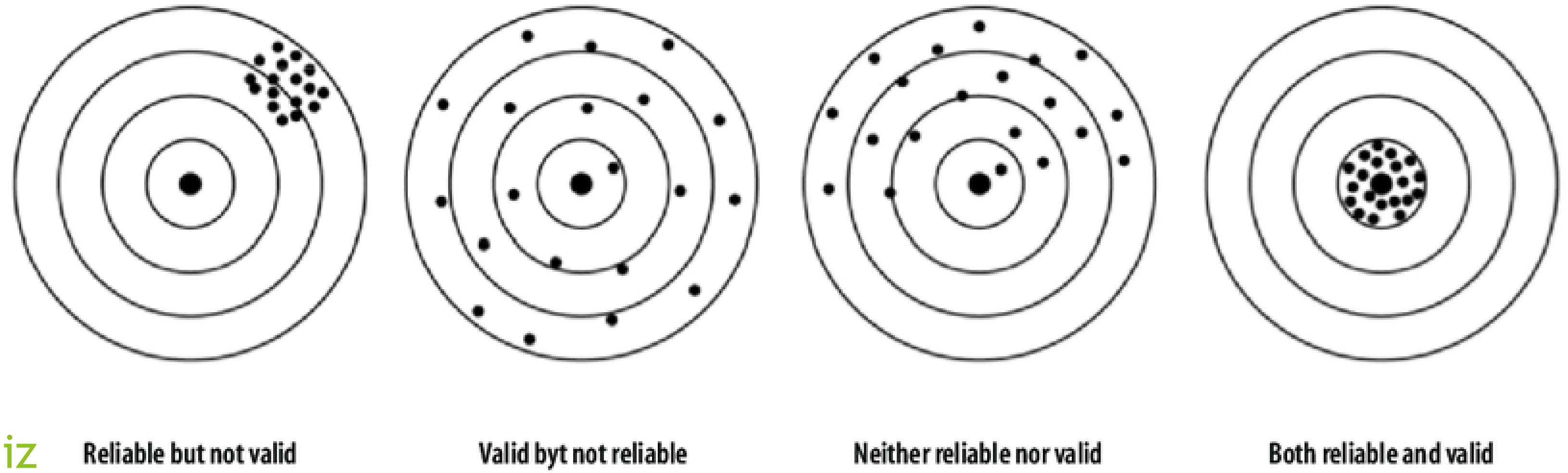
Latent Variables: formative vs. reflective

Confirmatory Factor Analysis

	Formative	**Reflective**
In Empirical Research	Rarely found in consumer Research (e.g. Reputation or Success factor research)	The common approach in Consumer Research
The measurement model	statistical tests for reliability of indicators do not make sense	Considers **measurement error** at the indicator level
Special Cases	Composite measurements (ingredients, weights can be predefined)	

Assessing Reflective Measurement Models

Reliability and Validity (Measurement error):



Quiz

- Estimate Reliability and Validity

Confirmatory Factor Analysis

Assessing Reflective Measurement Models

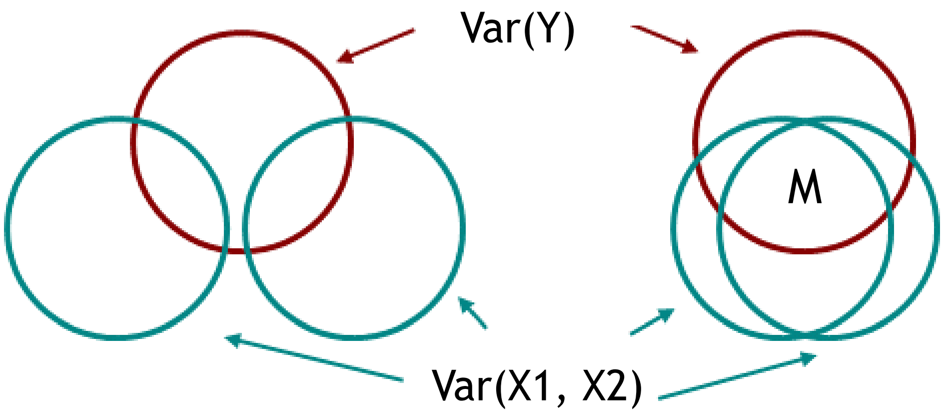
Reliability and Validity – capture measurement error:

1. **Concept Reliability: Cronbach's alpha**, **Dijkstra-Henseler's rho**)
2. Concept Validity
   - **Convergence Validity:** Average Variance extracted, i.e. **AVE** (similar to explained variance concepts or LV explanatory power for indicator variance)
   - **Discriminant Validity** (Heterotrait-monotrait ratio of correlations, i.e. **HTMT**, early assessments with Fornell-Larcker criterion: An LV should explain variance of its own indicators better than indicators of other LVs
3. **Indicator Reliability** (squared factor loadings): an indicator's explainatory power for an LV

Quiz: Why can we theoretically keep an indicator with a factor loading <0.1?

Assessing Formative Measurement Models

- Indicator Relevance: factor loadings
- Multicollinearity: Variance Inflation factor (VIF)



- External validity: empirical confirmation

Structural Equation Models

A structural equation model is essentially a set of structural equations that puts constraints on the model-implied variance-covariance matrix of observed variables.

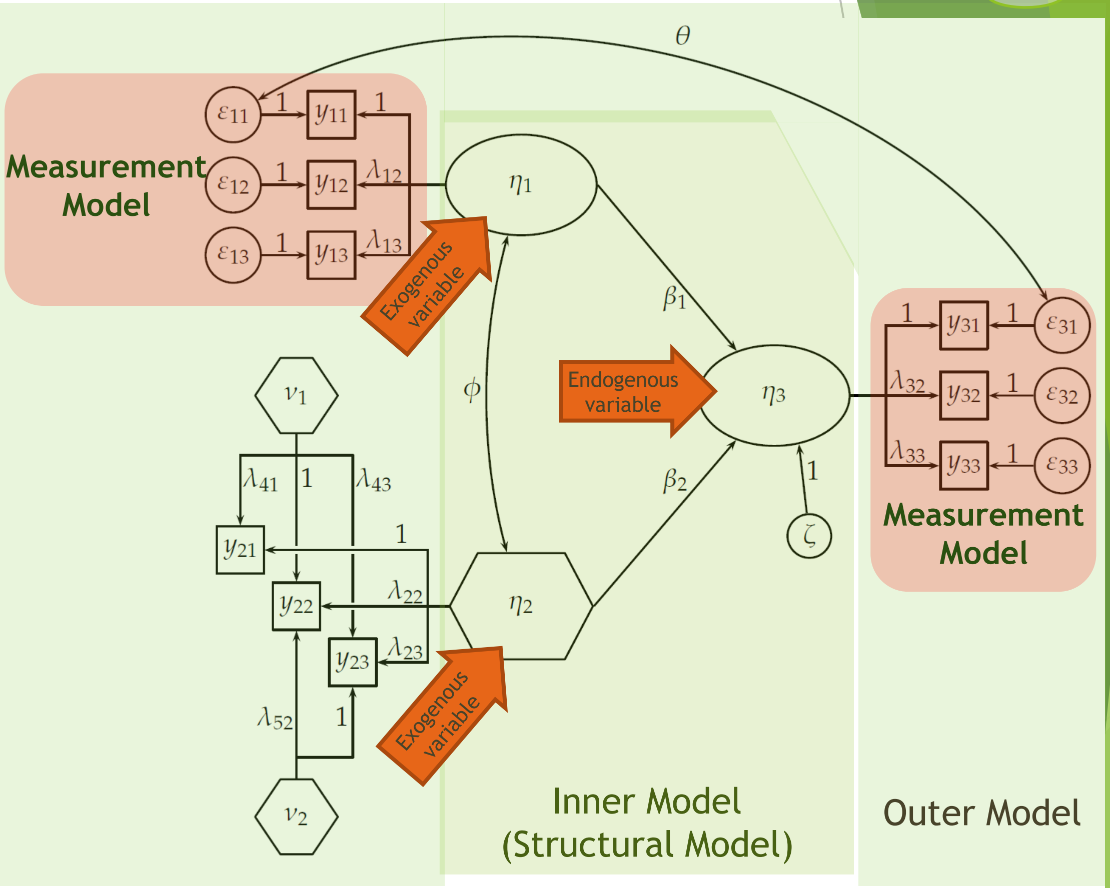
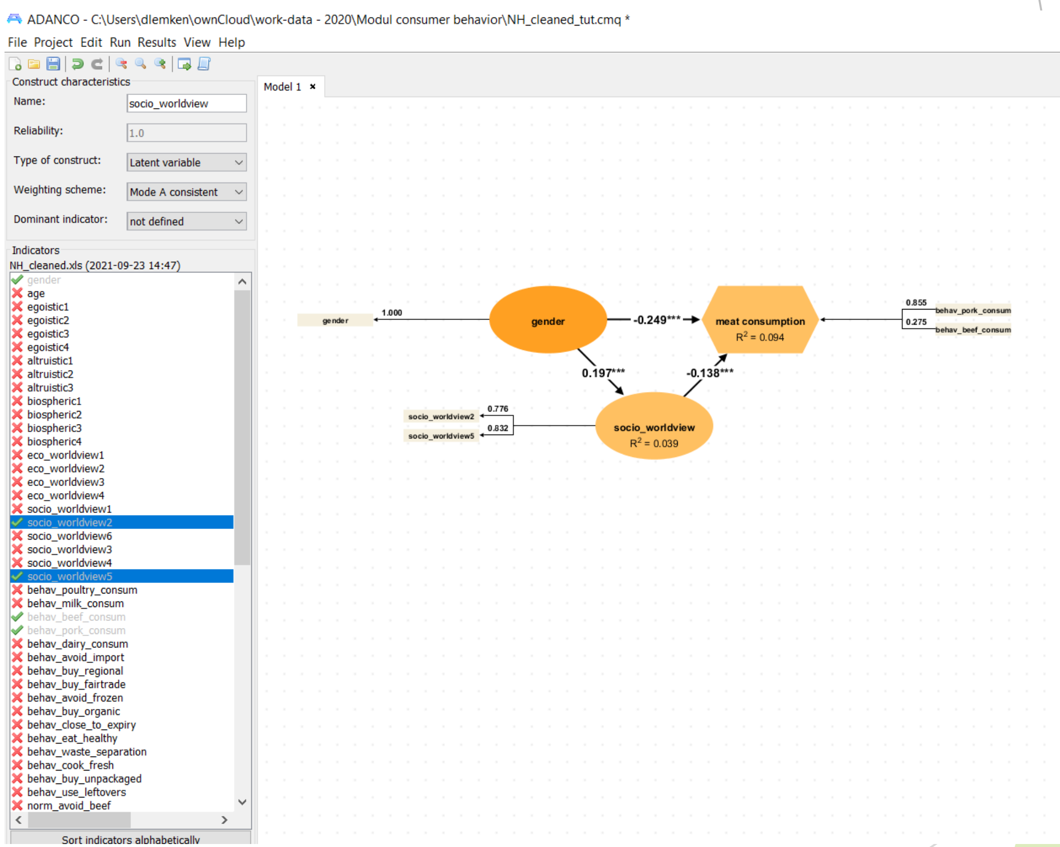


FIGURE 3.1. An exemplary structural equation model with three constructs.

Introduction to Adanco

Why Adanco:

- Cheap
- User friendly and fast to learn
- Developed by Henseler
- PLS-SEM



Steps of an SEM

STEP 6
Interpretation

STEP 5
Importance-Performance Map

STEP 4
Parameter Estimation

STEP 3
Inner Model Specification

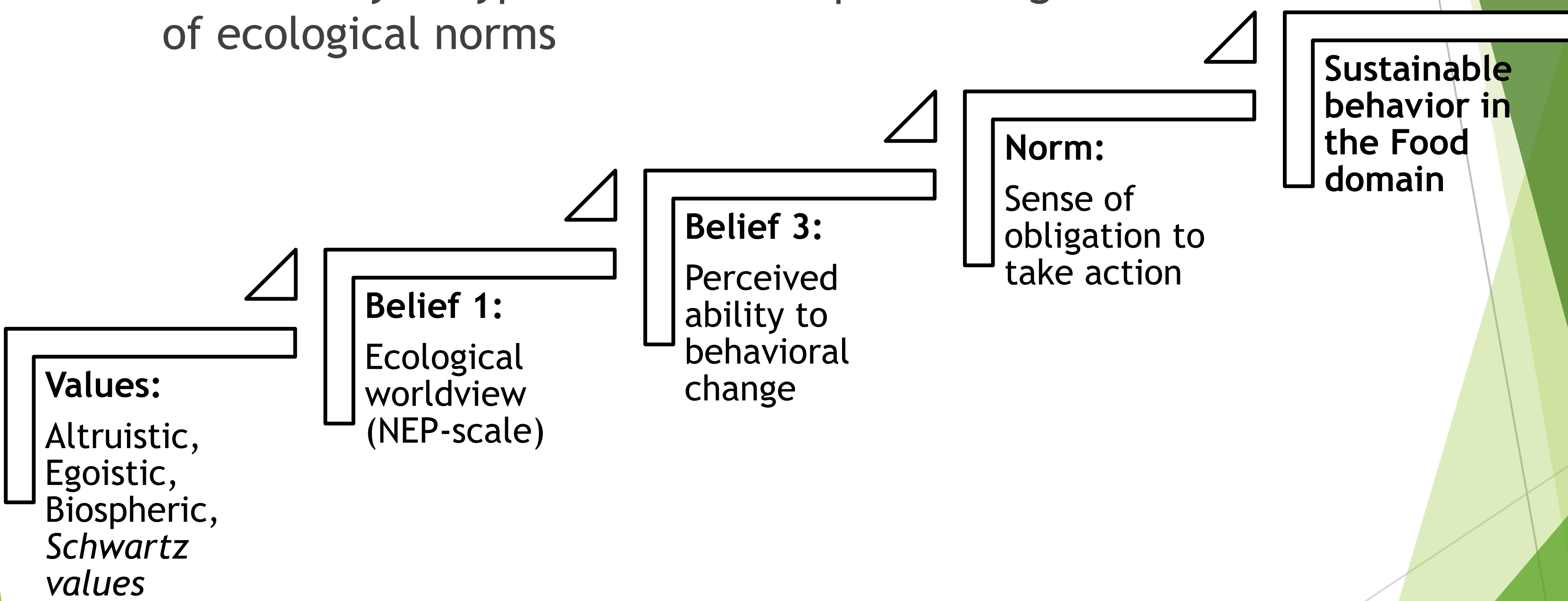
STEP 2
Outer Model Specification

STEP 1
Data Preparation

FIGURE 12.2. Steps of an importance-performance analysis.

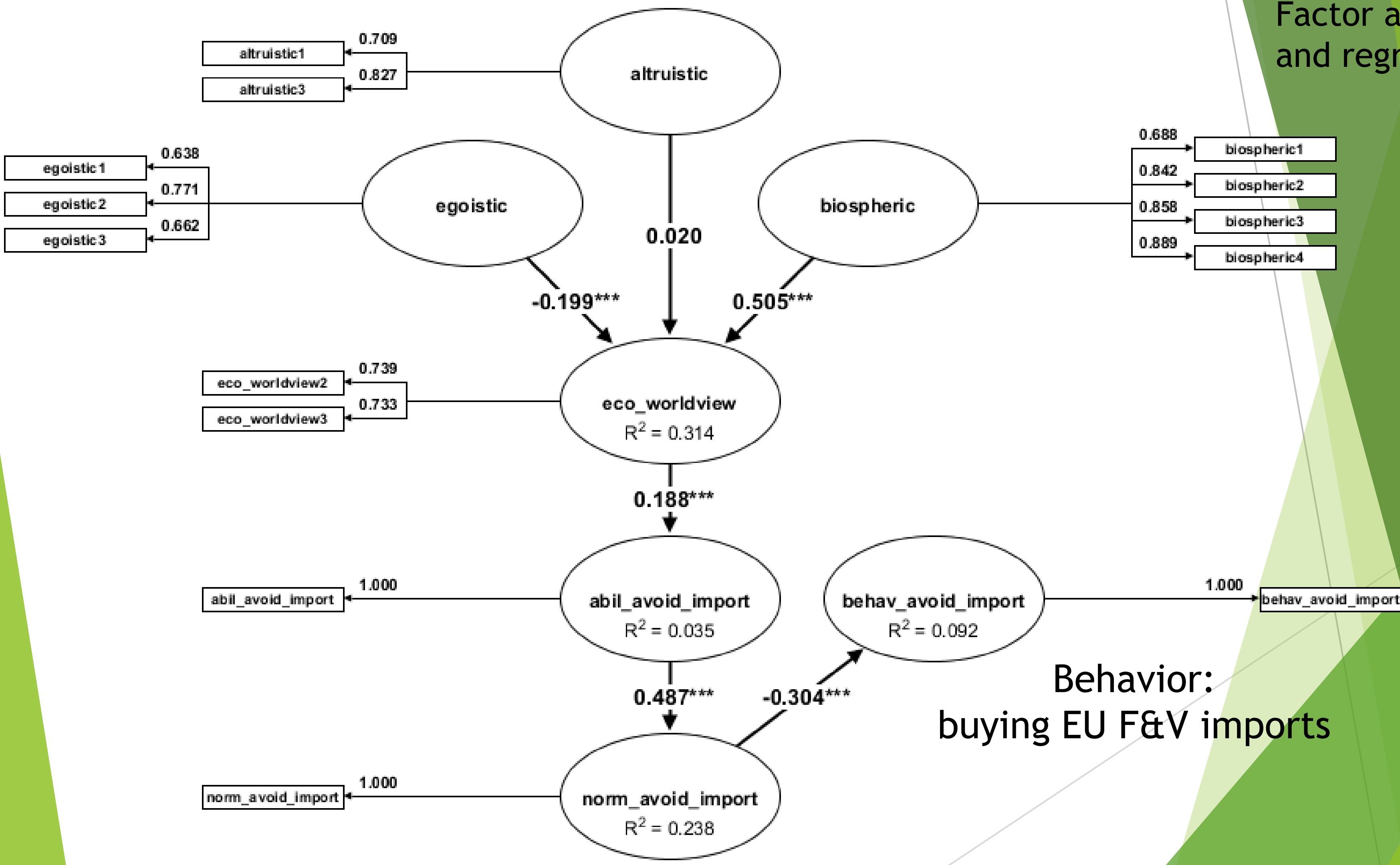
Value Belief Norm Theory (VBN)

- VBN-theory is hypothesized to explain the generation of ecological norms



Empricial Data: VBN-Theory (M1)

Factor analysis and regressions!



Behavior:
buying EU F&V imports

Assessing Reflective Measurement Models: Construct Reliability

Confirmatory Factor Analysis

STEP 2
Outer Model Specification

The model after recoding eco_worldview 1

Construct	Dijkstra-Henseler's rho (ρ_A)	Jöreskog's rho (ρ_c)	Cronbach's alpha(α)
egoistic	0,7397	0,7330	0,7390
altruistic	0,7512	0,7435	0,7391
biospheric	0,8999	0,8926	0,8926
eco_worldview	0,7027	0,7027	0,7027
abil_avoid_import	1,0000	1,0000	
behav_avoid_import	1,0000	1,0000	
norm_avoid_import	1,0000	1,0000	

DHR> 0.7, Cα>0.7 is advised

Assessing Reflective Measurement Models: Convergence Validity

Confirmatory Factor Analysis

STEP 2
Outer Model Specification

Construct	Average variance extracted (AVE)
egoistic	0,4796
altruistic	0,5932
biospheric	0,6770
eco_worldview	0,5417
abil_avoid_import	1,0000
behav_avoid_import	1,0000
norm_avoid_import	1,0000

Egoistic 4 dropped

AVE > 0.5 advised

Quiz: Does the model violate the assumption of convergence validity?

Assessing Reflective Measurement Models: Discriminant Validity

Confirmatory Factor Analysis

STEP 2
Outer Model Specification

Construct	egoistic	altruistic	biospheric	eco_worldview	abil_avoid_import	behav_avoid_import	norm_avoid_import
egoistic	0,4796						
altruistic	0,0005	0,5932					
biospheric	0,0010	0,3599	0,6770				
eco_worldview	0,0463	0,1072	0,2738	0,5417			
abil_avoid_import	0,0151	0,0000	0,0055	0,0354	1,0000		
behav_avoid_import	0,0197	0,0029	0,0001	0,0049	0,0789	1,0000	
norm_avoid_import	0,0309	0,0104	0,0497	0,0714	0,2376	0,0924	1,0000

Forner Larcker Criterion or HTMT: An LV should explain variance of its own indicators better than indicators of other LVs

Solution: Merging or Dropping of constructs or indicators

Discriminant Validity: The Heterotrait-Monotrait Ratio of Correlations (HTMT)

Confirmatory Factor Analysis

STEP 2
Outer Model Specification

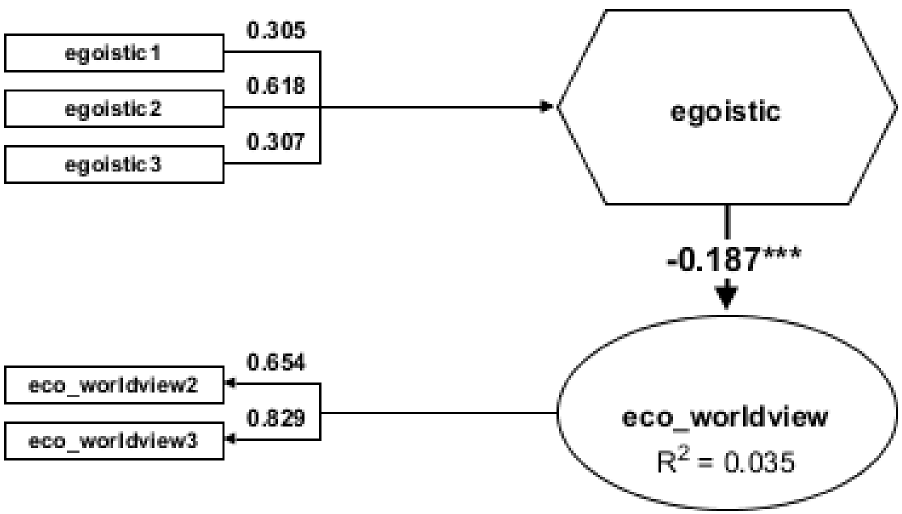
- A latent variable should explain better the variance of its own indicators than the variance of other latent variables.
- Consequently, a construct's indicators should correlate more highly with each other than with the indicators of other constructs.
- Realization: Heterotrait-monotrait ratio of correlations (HTMT)
- Thresholds:
  - $HTMT_{.85}$: $HTMT < 0.85$
  - $HTMT_{.90}$: $HTMT < 0.90$
  - $HTMT_{inference}$: $HTMT < 1.00$ (significantly)

Assessing formative Measurement models

Confirmatory Factor Analysis

STEP 2
Outer Model Specification

- You may argue Egoistic Values based on Schwartz are not reflected by importance of authority, wealth, social power and influence, but egoistic values consist of these 4 attributes:

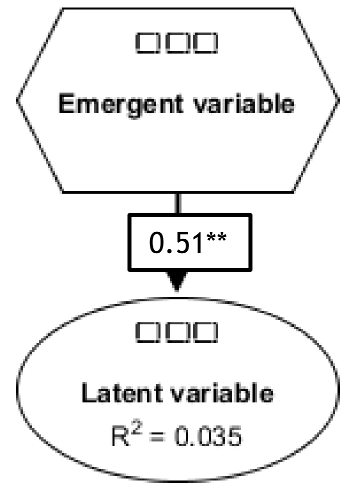


Confirmatory Factor Analysis

Assessing formative Measurement models: External validity

STEP 2
Outer Model Specification

- Strong and significant relationship between formative concept and direct inquiring of concept



Beware: At the data collection stage you need to ask a direct question (reflective) on the latent concept

- Reflective concept suffers from measurement error
- Reflective concept should be correlated with the formative concept

Assessing formative Measurement models: Collinearity

Confirmatory Factor Analysis

STEP 2
Outer Model Specification

Variance inflation factor (VIF) < 5

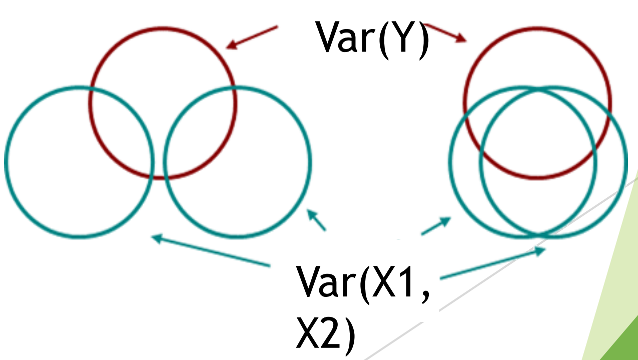
- Conservative measure

Indicator	egoistic
egoistic1	1,7225
egoistic2	1,2645
egoistic3	1,7827

Variance inflation factors (VIF)

- Statistical threshold: **Estimates are affected long before collinearity**
- Empirical Correlation

	egoistic1	egoistic2	egoistic3
egoistic1	1,0000	0,3943	0,6332
egoistic2	0,3943	1,0000	0,4290
egoistic3	0,6332	0,4290	1,0000



Confirmatory Factor Analysis

STEP 2
Outer Model Specification

Assessing formative Measurement models: Indicator relevance

- Empirical Relevance
  - Loadings > 0,4 can be included
  - Loadings > 0,7 must be included

Loadings

Indicator	egoistic
egoistic1	0,7433
egoistic2	0,8704
egoistic3	0,7655

- Significance based on bootstrap:
  - T-Values > 1,645 (α=0,05)

T-Values

Indicator	egoistic
egoistic1	6,3081
egoistic2	9,4835
egoistic3	5,9920

Multiple Regression

Assessing the Structural models: The path coefficient

How to interpret the path coefficient?

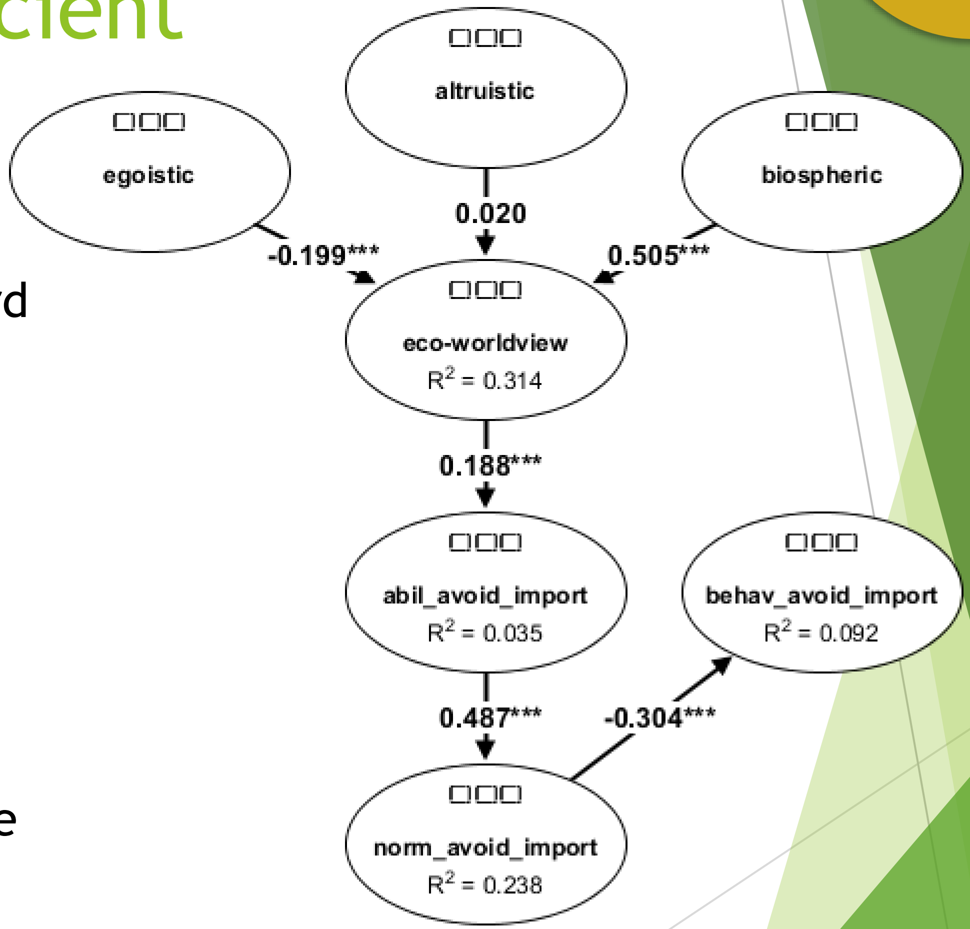
1. PLS standardizes each var
2. If X changes by one standard deviation Y changes by b standard deviations (with b being the path coefficient)
3. Perfection correlation results in b=1, reverse b=-1

⇒ **b is perfectly comparable within a model**

⇒ Interpret sign

⇒ Interpret relative effect size

holds if everything else is constant



STEP 3

Inner Model Specification

Quiz: What is the range of values b can obtain?

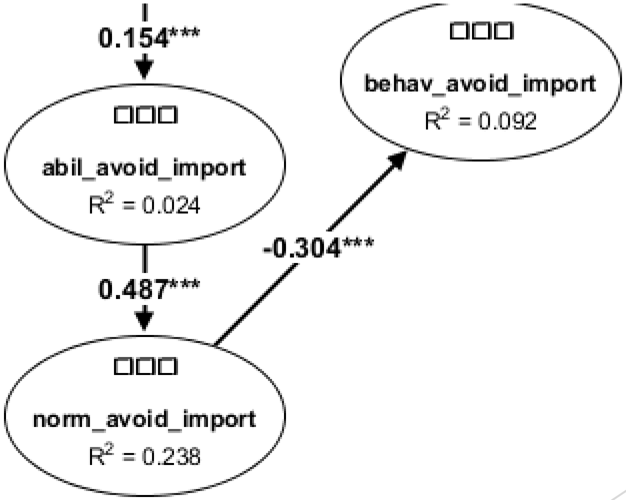
Assessing the Structural models: the “objective” effect size (f^2)

Multiple Regression

f^2-statistic:

Effect	Beta	Indirect effects	Total effect	Cohen's f^2
abil_avoid_import -> norm_avoid_import	0,4875		0,4875	0,3117
abil_avoid_import -> behav_avoid_import		-0,1482	-0,1482	
norm_avoid_import -> behav_avoid_import	-0,3039		-0,3039	0,1018

- Effect size based on f^2 (threshold may depend on discipline):
  - f^2>0,35 strong effect
  - f^2>0,15 moderate effect
  - f^2>0,02 weak effect



STEP 3
Inner Model Specification

Assessing the Structural models: Significance and Bootstrapping

Multiple Regression

Should we assume our data to be normally distributed?

- Resampling method with replacement

⇒ draw multiple samples to create a population

- Bootstrapping estimates a sampling distribution => standard errors =>
  - Confidence intervals, variance, prediction error and so on
  - No parametric tests, no normality assumptions
- Most recommend a bootstrap with repetitions >= 999 OR >N

Sample	N	X -> Y
1	1000	0.51
2	1000	0.46
3	1000	0.43
4	1000	0.47
5	1000	0.3
6	1000	0.62
7	1000	0.5
8	1000	0.55
9	1000	0.38
10	1000	0.44
11	1000	0.46
12	1000	0.44
…	…	…
k	1000	0.52

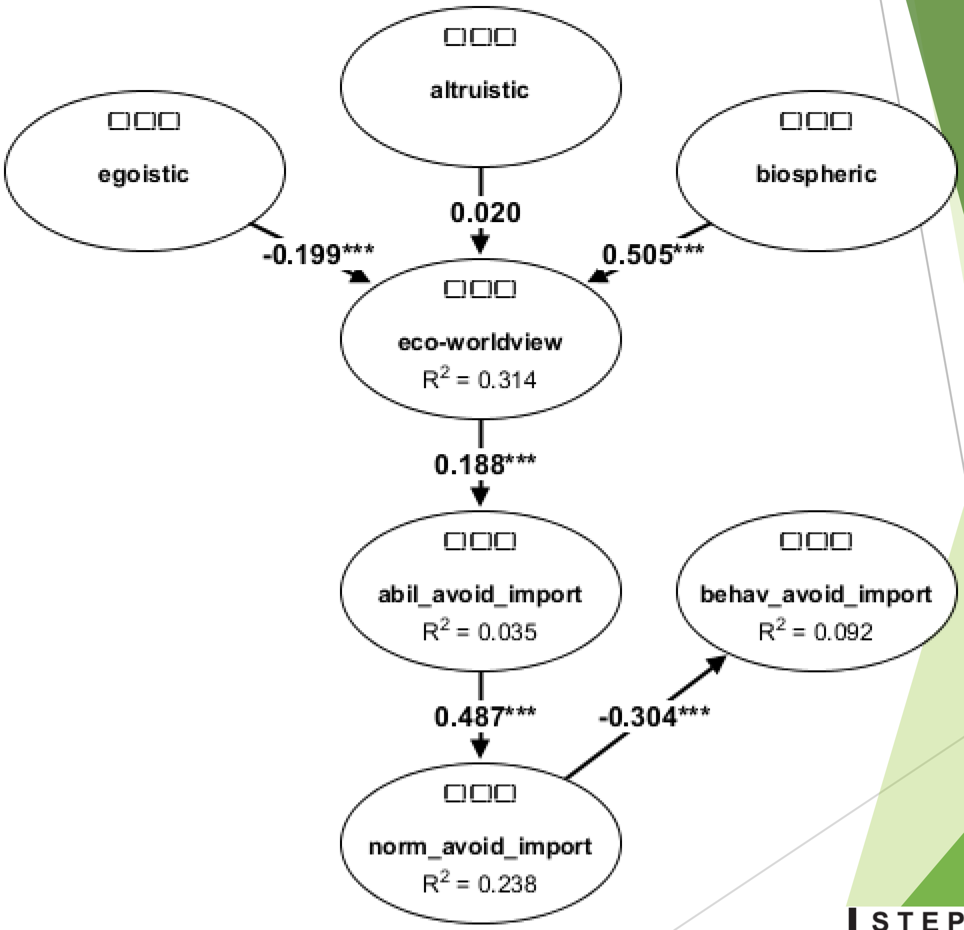
STEP 3
Inner Model Specification

Assessing the Structural models: R^2

Multiple Regression

(Adj.) R^2: measures within sample prediction

$\Rightarrow$ Here R^2 is not a global model fit criteria



STEP 3
Inner Model Specification

Assessing the Structural models: Overall model fit and prediction power

Compare different models

	Value	HI95	HI99
SRMR	0,0575	0,0328	0,0402

- SRMR = standardized root mean square residual, the lower the better, SRMR<0,08 is considered a good model fit (estimated model)
- Q^2-predictive relevance (not in Adanco, yet)
  - Blindfolding procedure: Blindfolding is a sample re-use technique, which systematically deletes data points and provides a prognosis of their original values

Assessing the Structural models: Exposing a Model to Reality

$$\Sigma \quad = \quad \Sigma(\theta) \quad ?$$

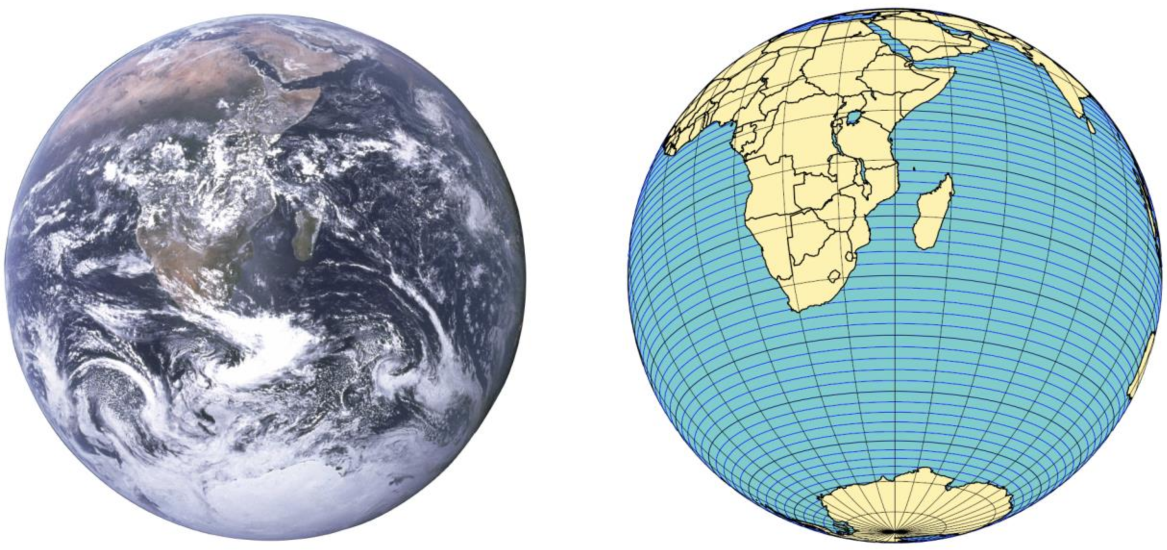


FIGURE 6.1. The nature of fit: An analogy.

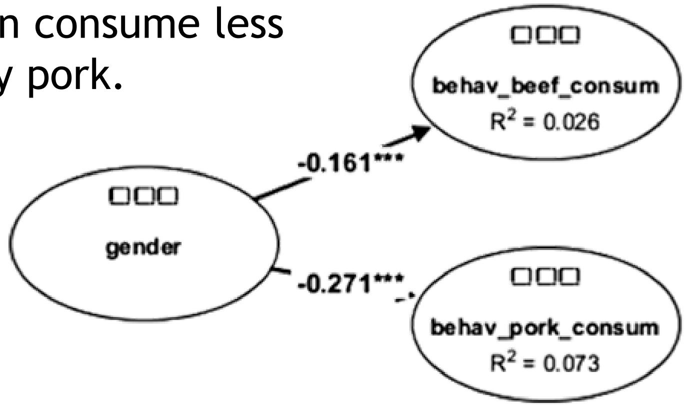
Now it's your turn: Tutorials 1 to 6

- Open a new project with Adanco
- Import data: NH_cleaned and open tutorials: https://github.com/dlemken/Structural-Equation-Modeling or studIP
  - NH_cleaned is university survey data (GF financed) on sustainable behavior in the food domain among German consumers 2018 (VBN based)
- Sustainable behavior is based on Geiger, Fischer, Schrader (2018):

Sustain-ability dimension	**Acquisition**	**Usage**	**Disposal**
Ecological	-I eat meat (pork, beef, poultry) for the main meals -I eat dairy products -I buy certified organic products -I buy imported fruits and vegetables	-I buy frozen foods and meals	-I actively separate waste -I refrain from foods with excessive packaging -I use left-overs for the next meal
Socio-economic	-I buy fair trade food products -I buy regional food products	-I eat healthy -I cook my own meals with fresh ingredients	I buy food close to their expiration date

Tutorial 1

There is a gender gap in red meat consumption. Women consume less beef and particularly pork.

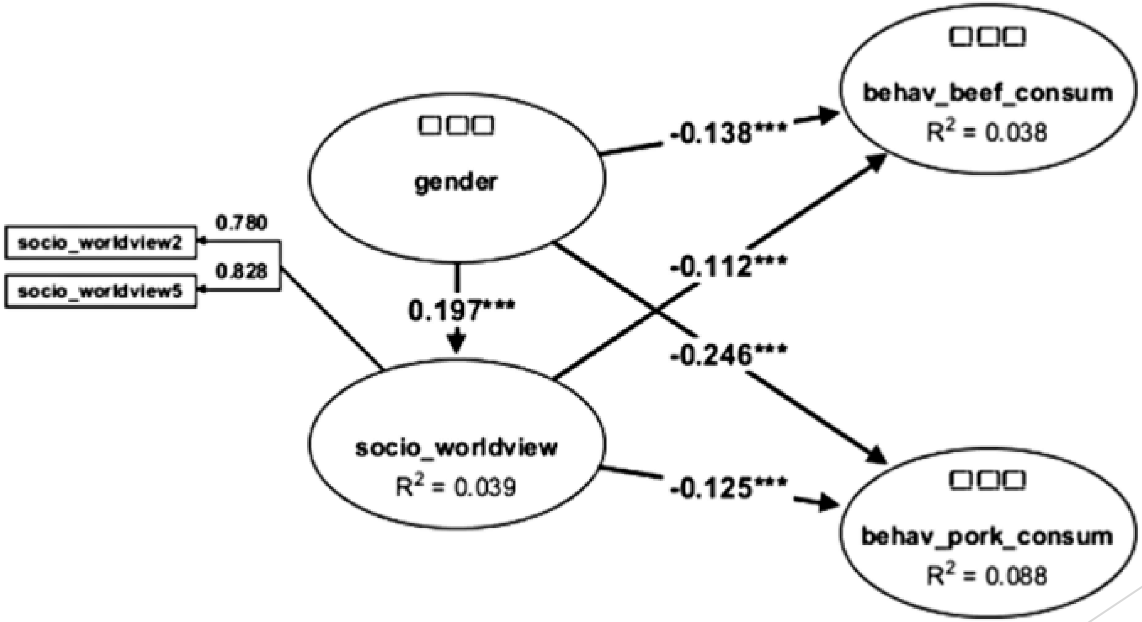


1. Can animal welfare related attitudes (here: socio_worldview2 and 5) explain the gender gap (mediate) in red meat consumption?
2. What type of mediation relationship has to be concluded?

Item name	Wording (Likert-Scale)
Socio_worldview2	Tiere sind mit Würde und Respekt zu behandeln
Socio_worldview5	Es sollte der Anspruch der Menschheit sein, dass Tiere zunehmend besser leben können

Solution 1 (Adanco 2.2.1)

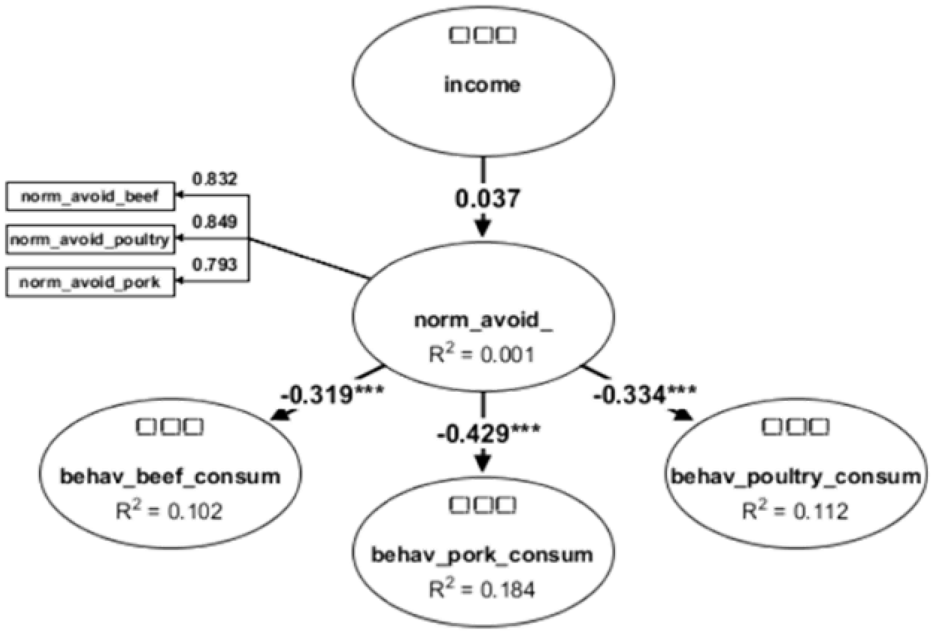
1. Gender can explain some of the variance in the socio_worldview, which in turn can explain some of the variance in red meat consumption
2. Complementary Partial Mediation



Tutorial 2

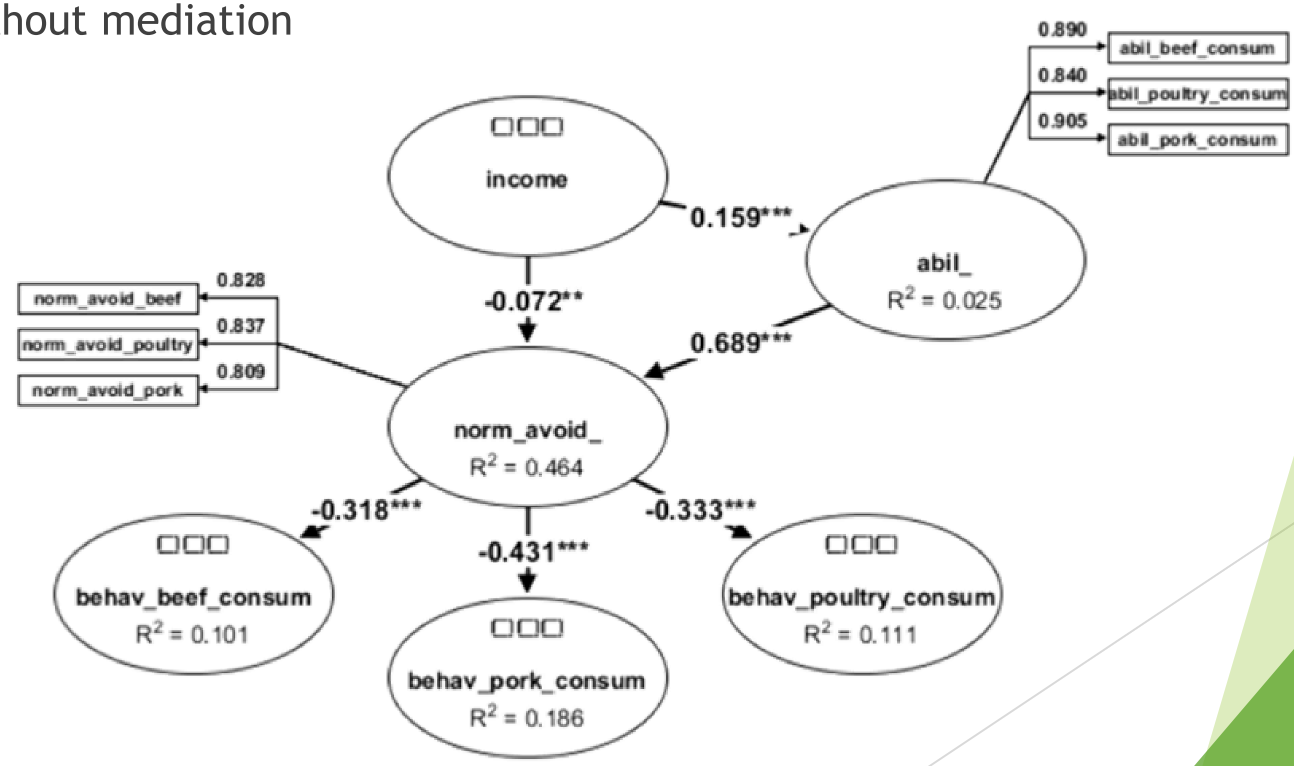
We are interested in income differences for the felt norm to avoid meat consumption. We conclude no significant differences.

1. Does the conclusion change if we mediate the relationship by the perceived ability to avoid meat (abil_beef_consum, _pork, _poultry)?
2. How do you interpret the mediated income effect on the norm?



Solution 2 Adanco 2.2.1

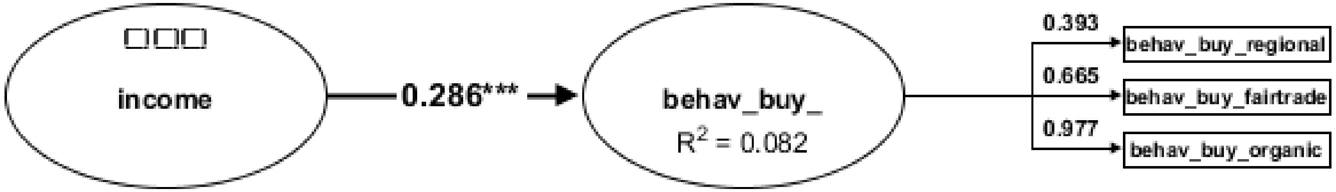
1. competitive mediation relationship. Direct effect b_d=-0.072. Indirect effect of income on norm is b_i= 0.1093
2. Apparently, income decreases the felt norm (additional effect mechanisms) which offsets the positive influence via the perceived ability. The sum of b_i and b_d should resemble the original effect without mediation



Tutorial 3

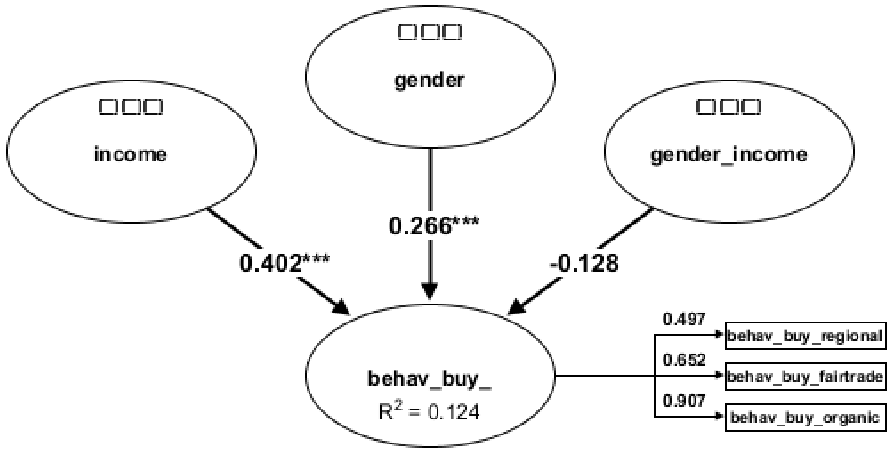
Consumers with a higher income consume more labelled food products (organic, regional, fairtrade). Is the effect of income moderated through gender differences? (report the p-value for the interaction term)

Hint: Use „gender“ and the interaction term „gender_income“ to model the moderation similar to a regression approach



Solution 3

- b=-0.128, a higher income for men leads to an even higher label purchase than for women. However, the variance is high. A significant relationship cannot be concluded, p=0.2409. **We cannot confirm a moderator relationship for the general population**



Reason for SEM

- Identify initial effects, even though final outcome may not be significantly affected
- Understand effect mechanism
- Understand parallel relationships
- Understand mediator and moderator
- Testing entire theories
- Measure structural relationships of latent variables
- Report on multiple hypothesis in 1 model

Reasons for Using PLS

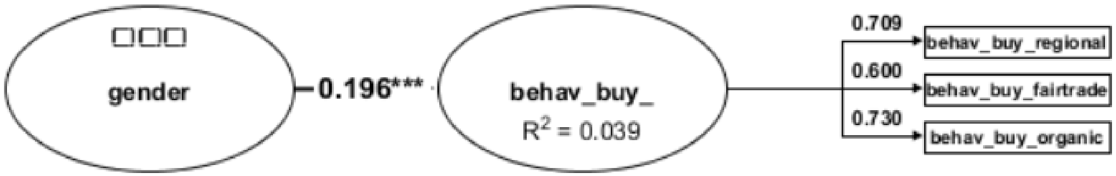
- Robust to small sample sizes (distribution assumption)
- **Robust to non-normal data**
- **Allows for formative measures**
- Prediction value
- Large number of indicators
- Theory development

Conclusions

- Understand the structural relationships you hypothesize
- Measure your concepts as good as possible
- Provide evidence for both with PLS

Tutorial 4

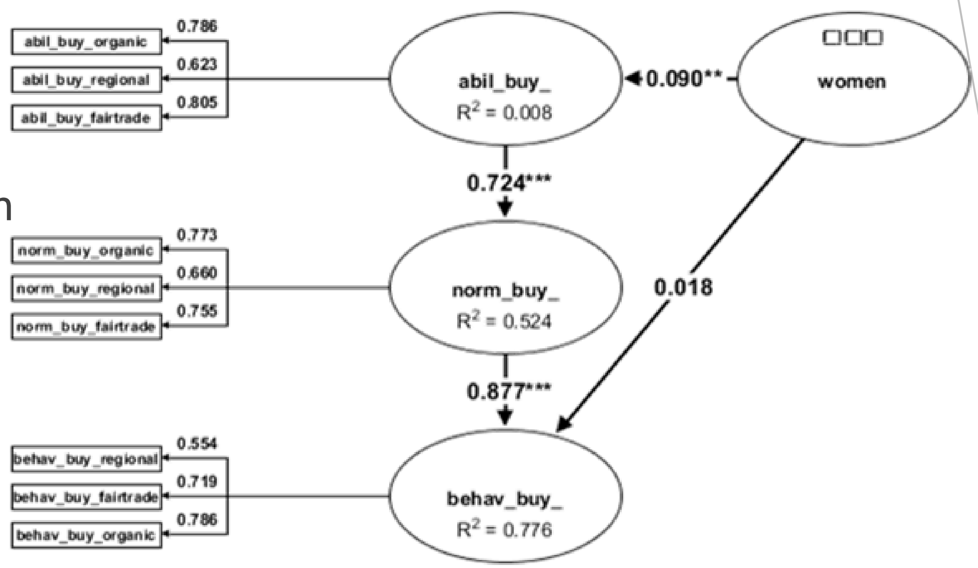
1. What type of measurement problem do we create, if we model the „gender - label purchase“ relationship with a potential mediation via the ability and norm to buy labelled products (abil_, norm_: regional, organic, fairtrade)?
2. What type of mediator relationship would we have concluded?



Solution 4

1. **A violation of discriminant validity:** The Norm to buy labelled products explains the indicators of label purchases better than its own indicators

⇒ Either drop Norm or behavior OR aggregate the concepts OR improve the measurement of label orientation (back to the field)



2. Full complementary mediation

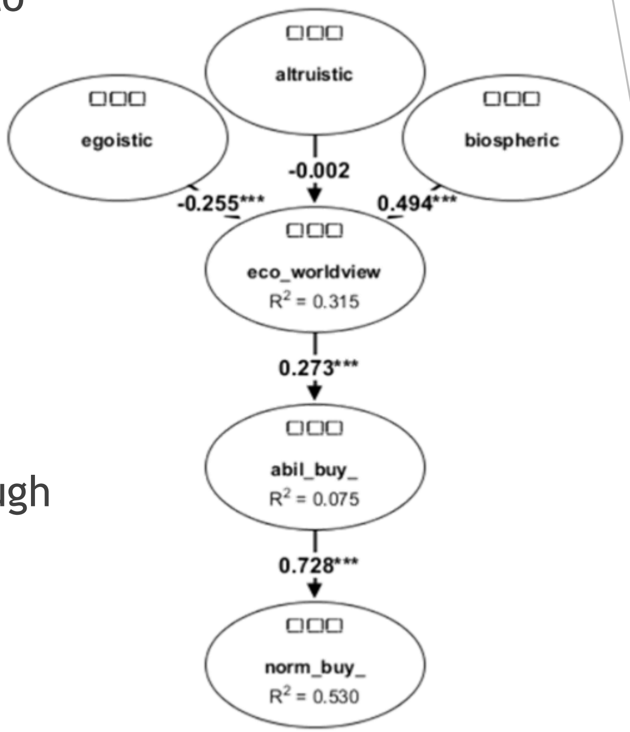
Tutorial 5

Build a VBN model to predict the norm to buy labelled products

Add a socio-economic worldview (socio_worldview 1 and 3) to the VBN-Theory.

1. Which values (egoistic, altruistic, biospheric) influence the socio-worldview?
2. Do altruistic values matter to the norm to buy labelled products through the socio-worldview?

- Keep AVE>0.45 and Cronbach_α >0.7 and do not use single indicator concepts

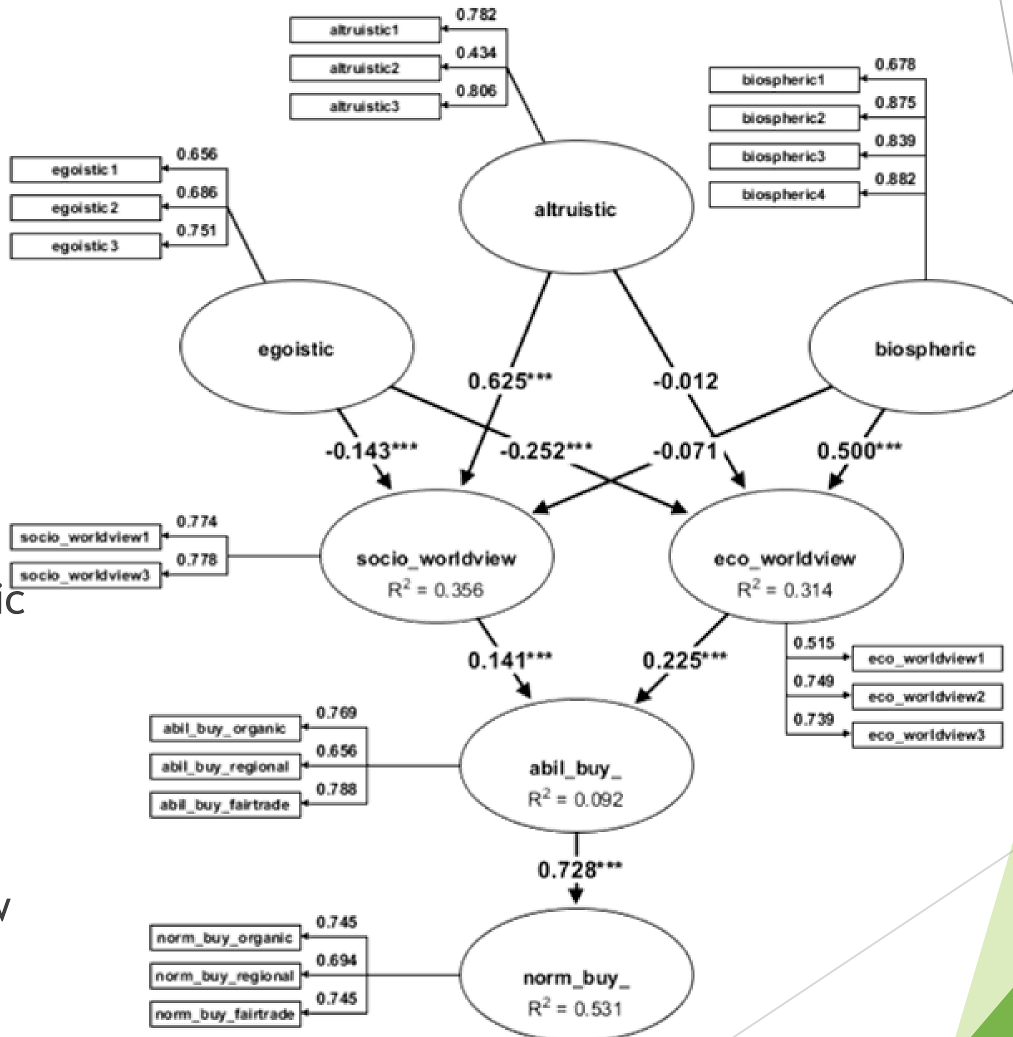


Item name	Wording (Likert-Scale)
Socio_worldview1	Humanity should have the aspiration that no human suffers from hunger
Socio_worldview3	All humans must have the opportunity to fulfil basic needs

Solution 5

- Drop egoistic4, eco-worldview4

1. Alturistic values strongly influence the socio-worldview, to a lesser degree egositic ones
2. b=0.0623, p=0.0043. altruistic values influence the norm to buy labelled food. However, the p-value may slightly change with a new bootstrap.



Tutorial 6

Build one comparative VBN model to predict the Norm to

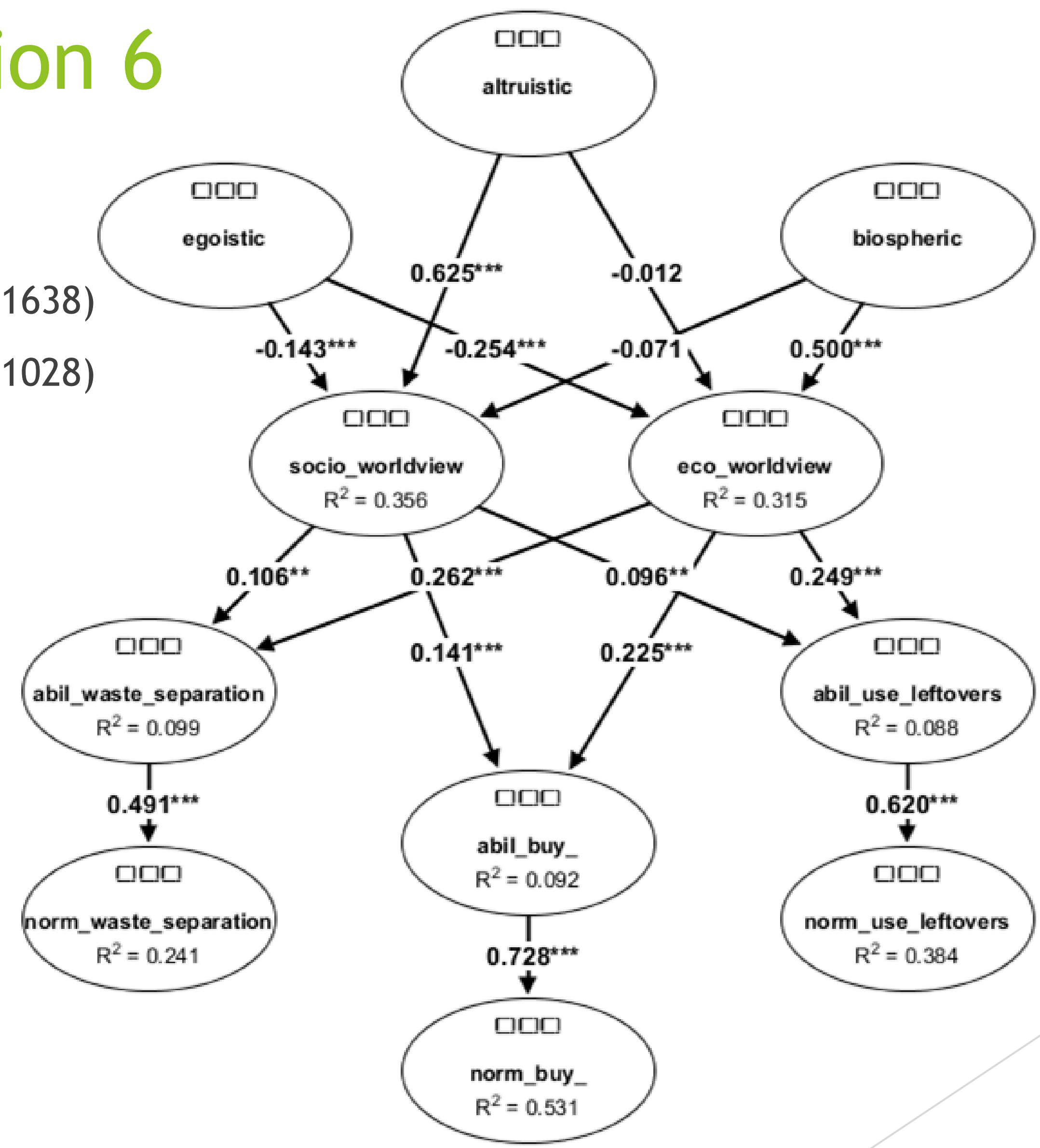
a. buy labelled products
b. to use leftovers
c. separate waste

1. Which of the norms (a,b,c) is explained best by the socio-worldview?
2. Which of the norms (a,b,c) is explained best by the eco-worldview?

Hint: start from the model in tutorial 5

Solution 6

1. a (b_i=0.1638)
2. a (b_i=0.1028)

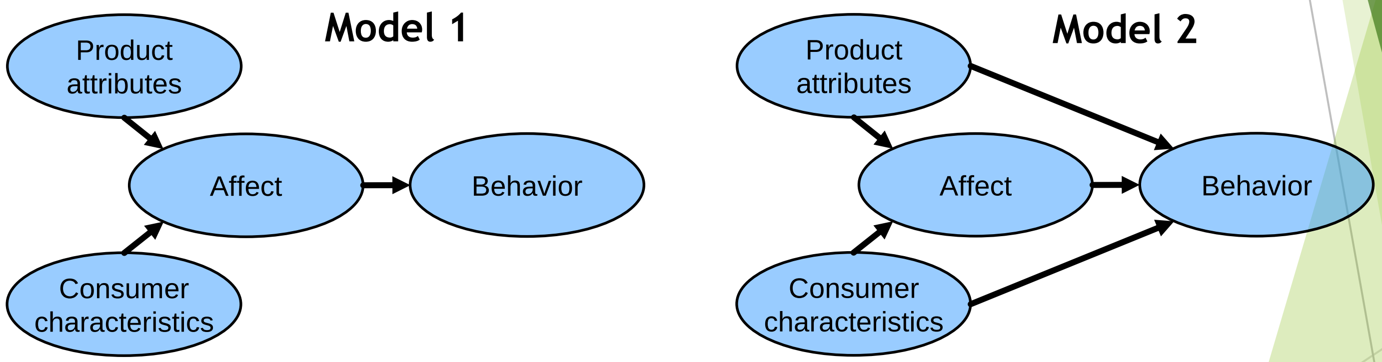


Outlook

- Upload the html-files to the owncloud folder during the week
- Chat will be uploaded
- Slides will be made available
- Contact me: dominic.lemken@ilr.uni-bonn.de

Theory Testing: Alternative Models and fit

- The researcher has specified several alternative models (or competing models) and, on the basis of an analysis of a single set of empirical data, one of the models should be selected.



- Model 1 is supported if its fit is not significantly worse.

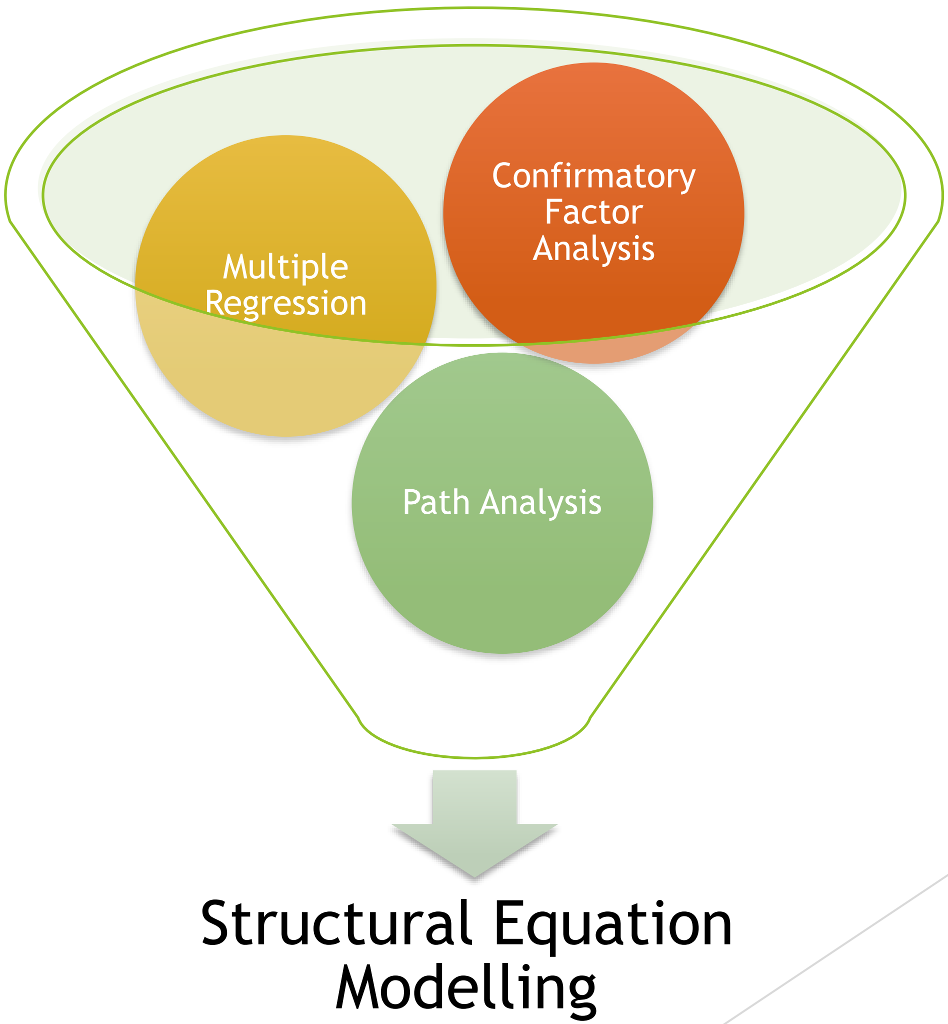
Outlook

- Causality: Group Comparisons allow tests for unobserved heterogeneity (2SLS)
- Creative reporting of the concepts are on the rise: e.g. Importance-Performance mapping
- Latent Variable Scores allow for a use of the variables in other software and analysis tools
- Several options for model comparison and inter model testing (e.g. path coefficients between models

Outlook

- Rekursive Mediator (do not control) and Confounder (do control) can be solved with lag variables in SEM
- Dagitty.net to visualize and evaluate structural relationships
- On model fit: https://www.smartpls.com/documentation/algorithms-and-techniques/model-fit

Good luck with your research!



Step 2: Outer Model Specification

Choosing the outer weighting scheme for IPA

Outer weighting scheme	Negative weights possible ?	Comment
PLS Mode B	yes	Best linear unbiased estimates.
PLS Mode PLS	yes	Somewhat less affected by multicollinearity than PLS Mode B.
BFPI	no	Some attributes can have an importance of 0.
PLS Mode A	**no**	**If correct orientation is ensured.**
PCA	no	If correct orientation is ensured. The weighting ignores the indicators' respective predictive validity.
Preset weights	no	If non-negative weights are chosen. The weighting ignores the indicators' respective predictive validity.

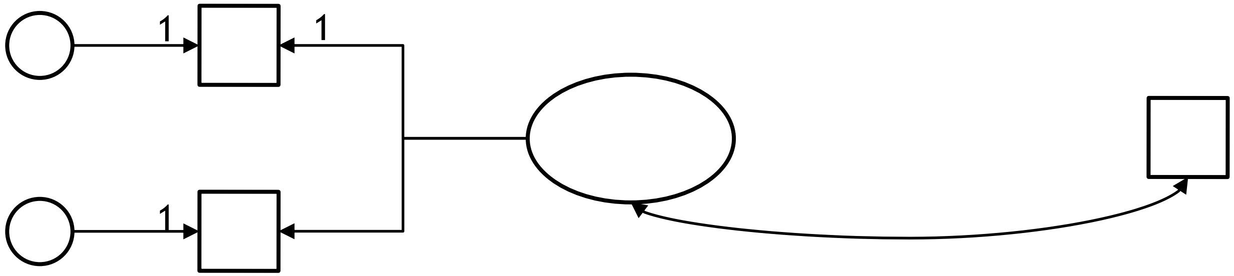
Determining the Relative Size of Complementary Mediating Effects

- Variance Accounted For (Shrout & Bolger, 2002):

$$\mathrm{VAF} = \frac{\text{Indirect effect}}{\text{total effect}} = \frac{a \cdot b}{a \cdot b + c}$$

- The VAF can become negative.
  - Interpretation as suppressor effects
  - Set to one.
- The VAF actually only makes sense for complementary mediation.

Rules for Latent Variables



<u>Rule 5:</u>* Latent variables with two indicators require a nomological net, i.e., they must have at least one non-zero path or covariance with some other variable in a structural equation model, or the analyst must add an additional constraint.

*assuming that there are no free covariances between measurement errors

PLS Path Modeling and Selected Extensions

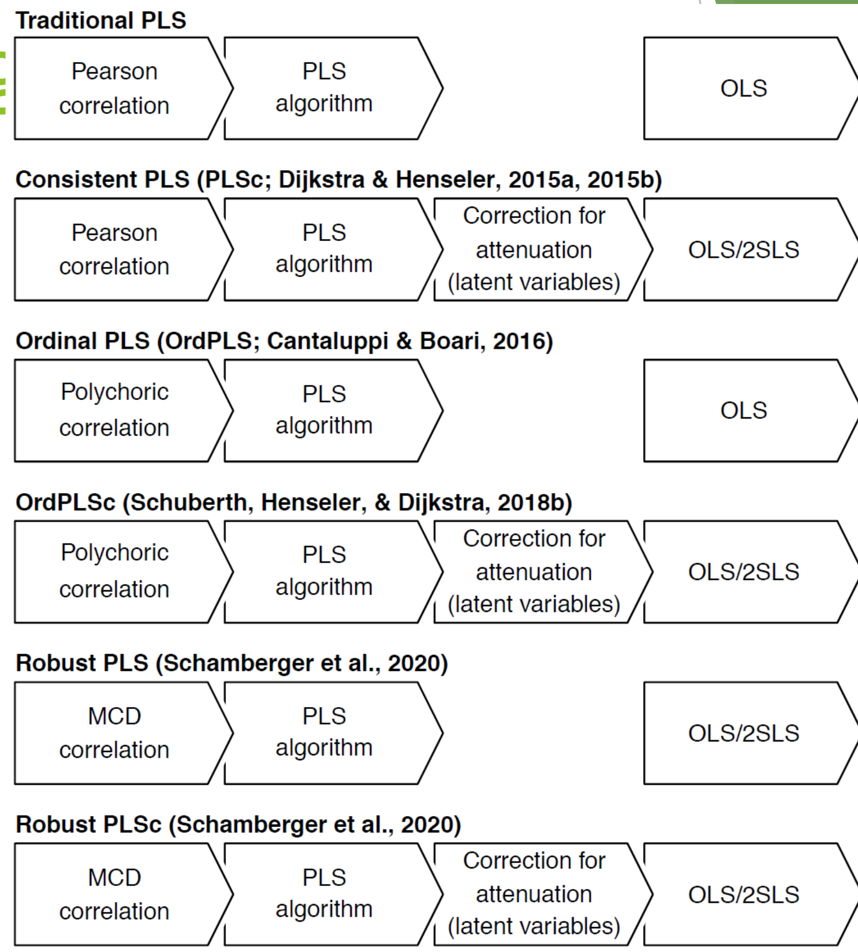


FIGURE 5.2. PLS path modeling and selected extensions.

Correction for Attenuation

The correlation between two latent variables equals the correlation between their construct scores divided by the geometric mean of their scores‘ reliabilities:

$$\operatorname{cor}\left(\eta_j, \eta_k\right) = \frac{\operatorname{cor}\left(\tilde{\eta}_j, \tilde{\eta}_k\right)}{\sqrt{\operatorname{reliability}\left(\tilde{\eta}_j\right) \cdot \operatorname{reliability}\left(\tilde{\eta}_k\right)}}$$

A Typology of Second-Order Constructs

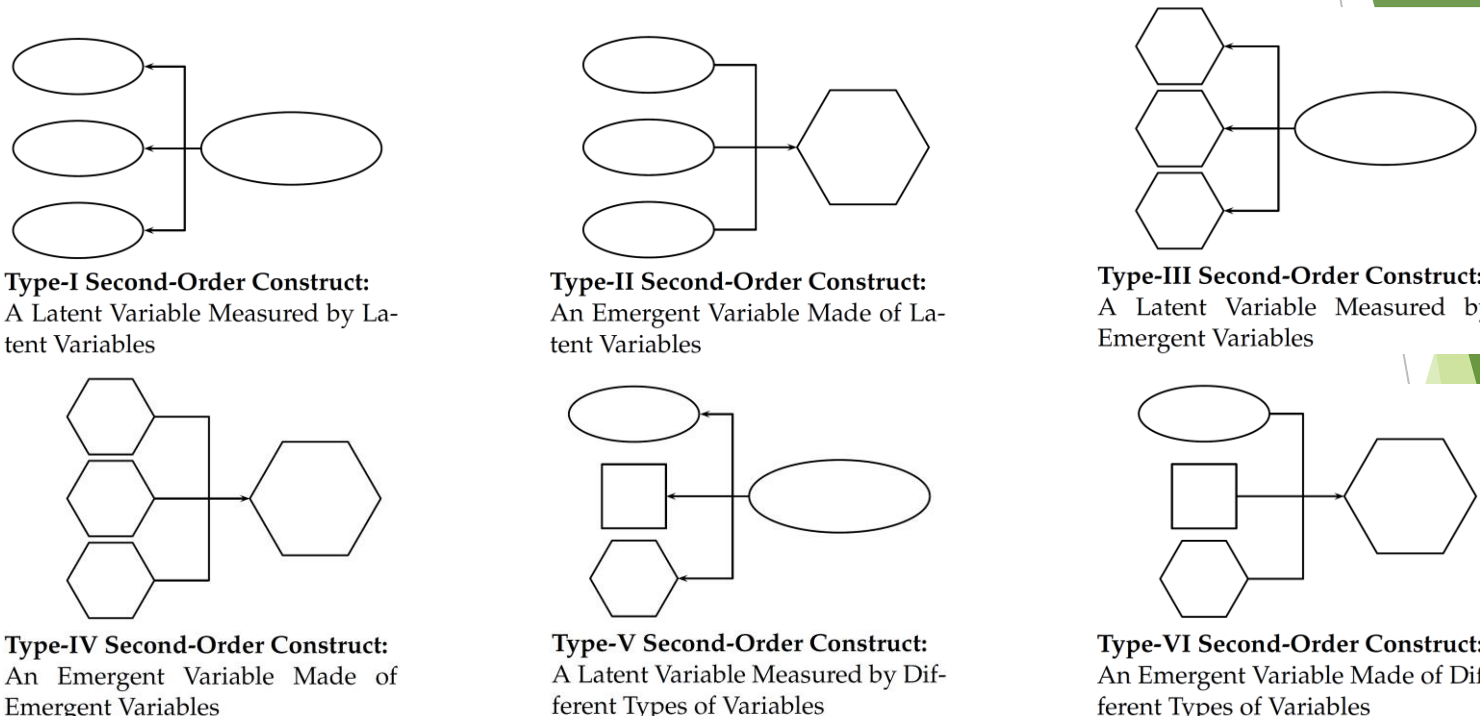


FIGURE 10.1. A typology of second-order constructs.

The Different Shapes of a Quadratic Effect

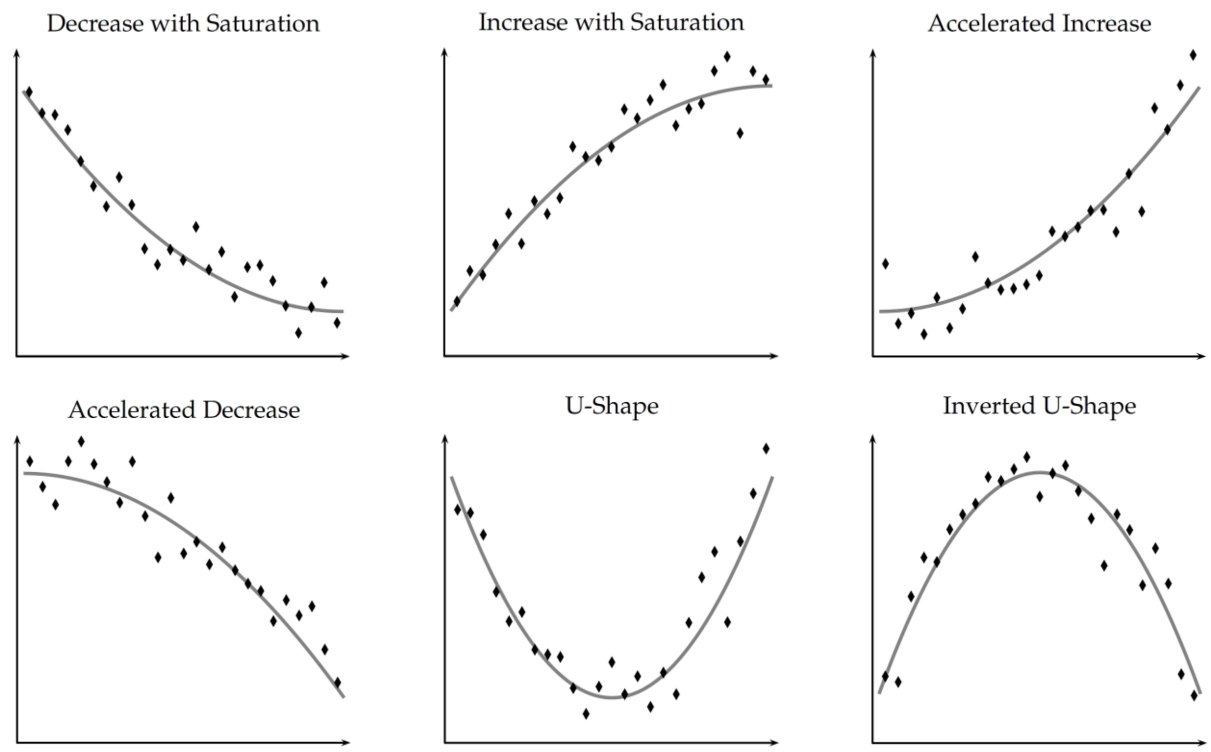


FIGURE 11.11. Different shapes of a quadratic effect.

Step 4: Parameter Estimation

Importance values:

- Total effects (unstandardized) on the final outcome variable
- If a variable is increased by one unit, to what extent does the final outcome variable change?

Performance values:

- Location parameters of the construct scores (typically: the arithmetic mean of the construct scores)
- What is the current level of the variable?