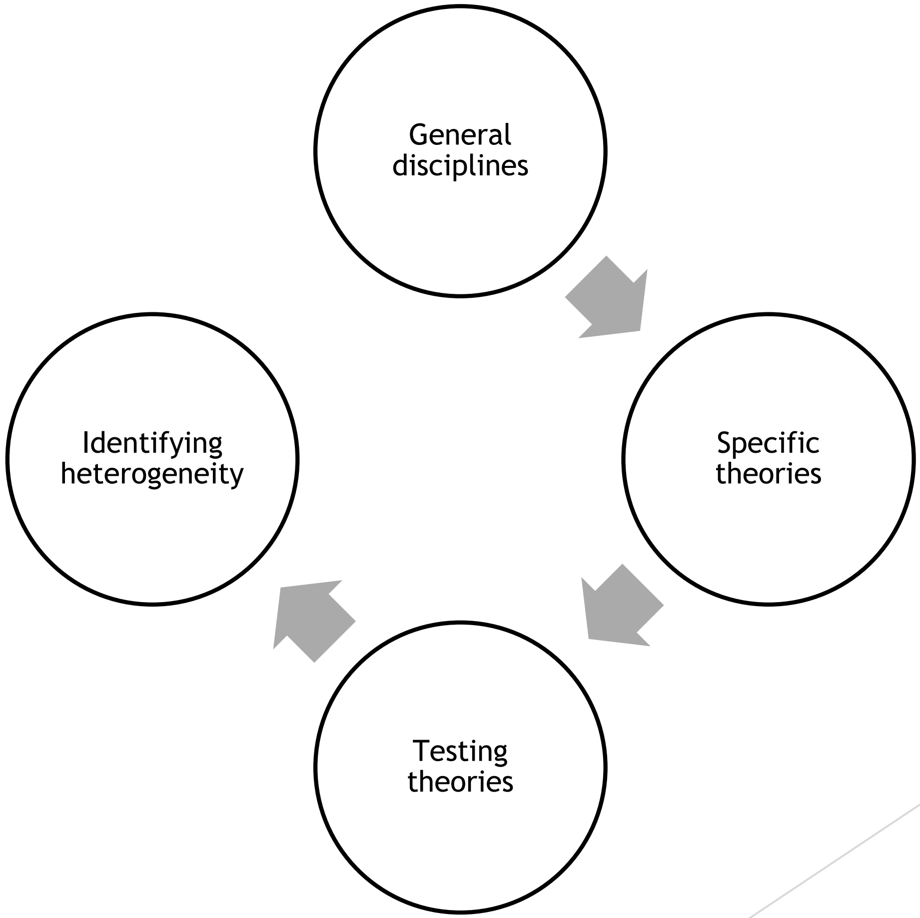
Advanced Theories of Consumer Research



A guide to structural equation modelling

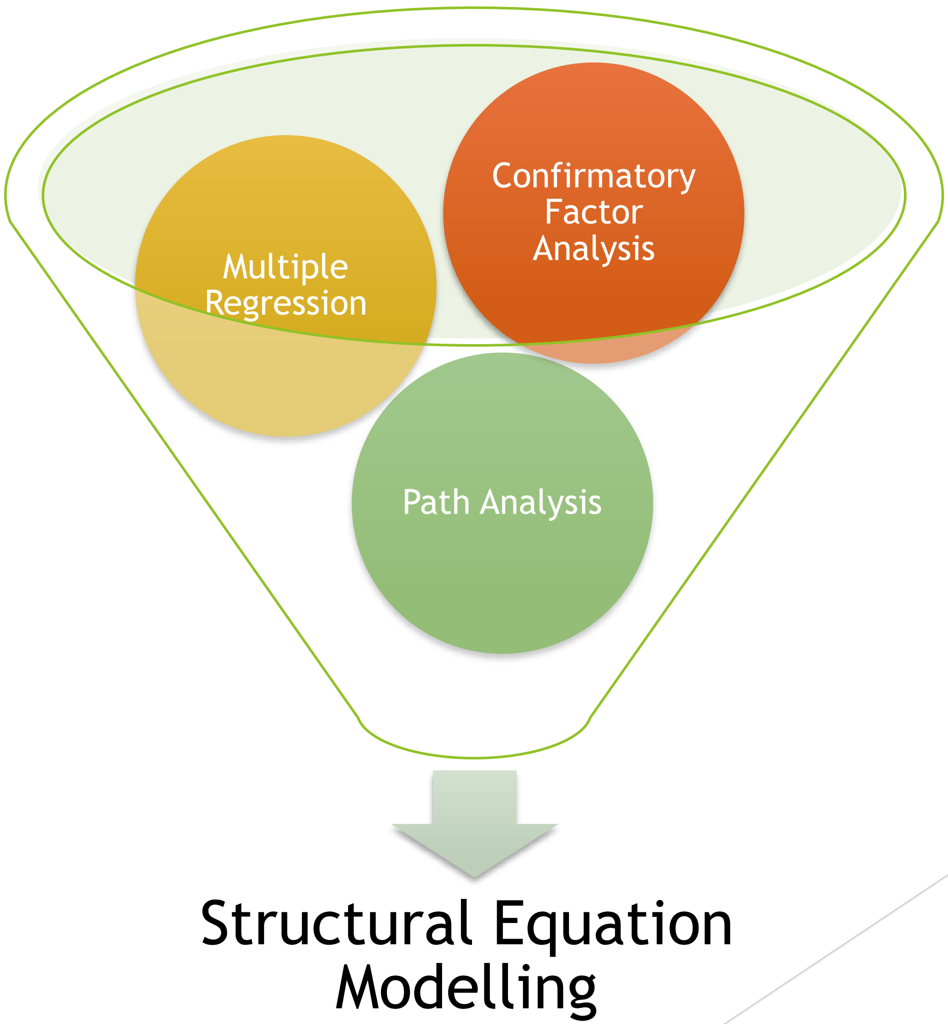
Outline:

- Morning: structural relationships and latent concepts/constructs
- Afternoon: SEM-Tutorials 1-6

Course material:

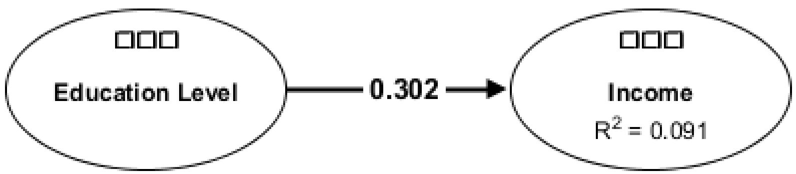
- Software: Adanco 2.2
- Course files https://github.com/dlemken/Structural-Equation-Modeling

Structural Relationships and Latent Concepts



Structural Relationships

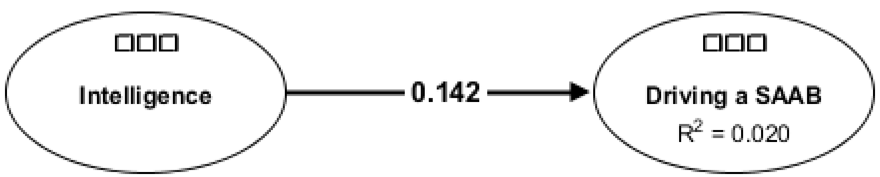
Research starts with an conscious or unconscious hypothesis:



Exogeneous

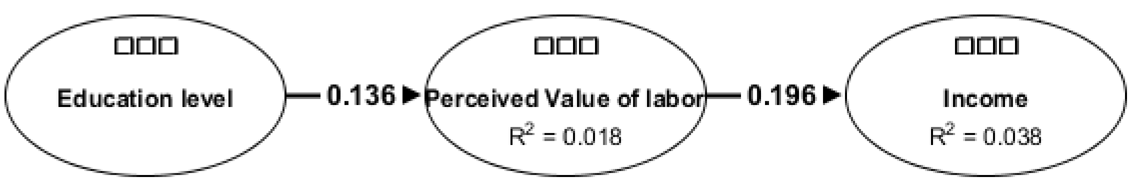
Endongeneous

Recursive?



Causality?

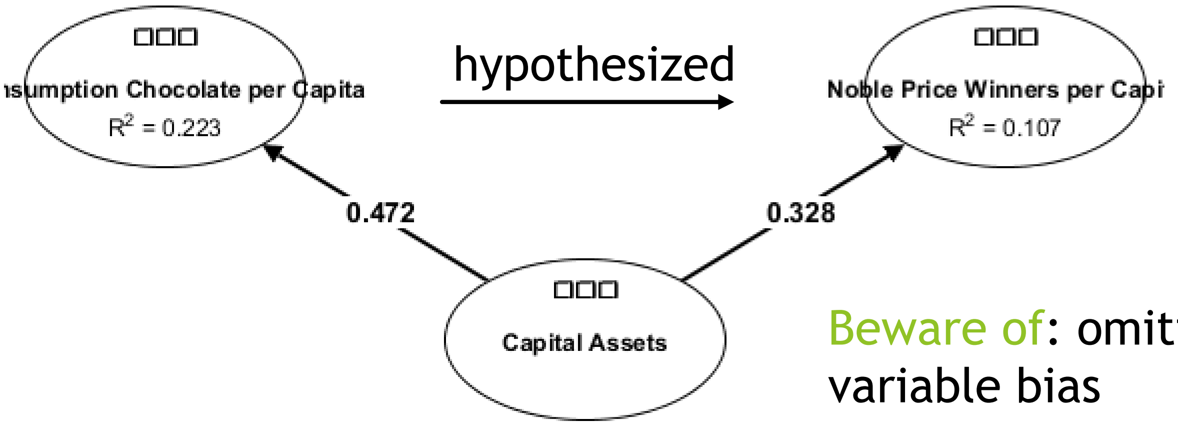
The effect mechanism



Structural Relationships

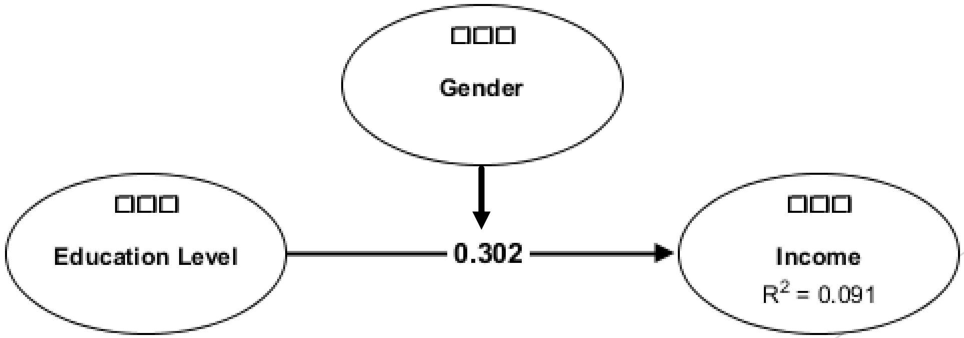
Path Analysis

Phantom Relationship: Revealing the need for a theoretical foundation



Beware of: omitted variable bias

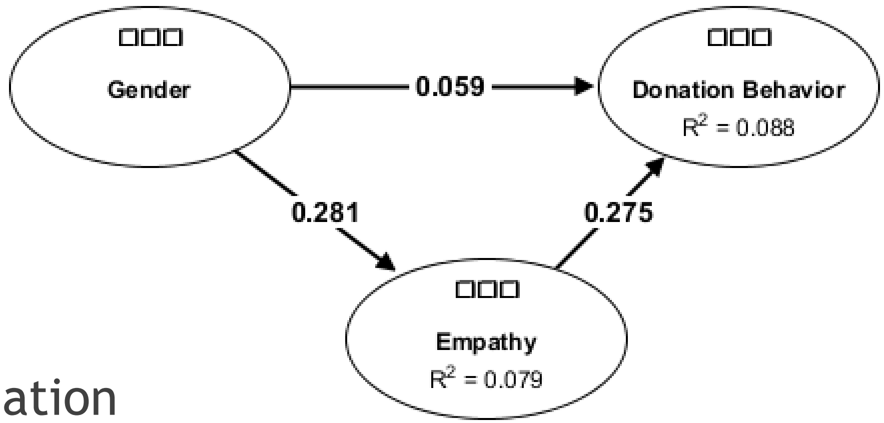
Moderated Relationship: Interaction Effects



Indirect effect: Triangle of Mediation

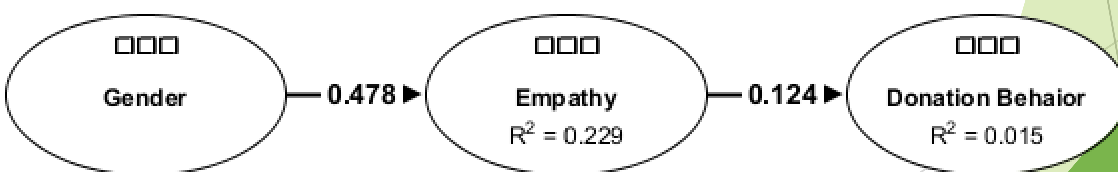
Path Analysis

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



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

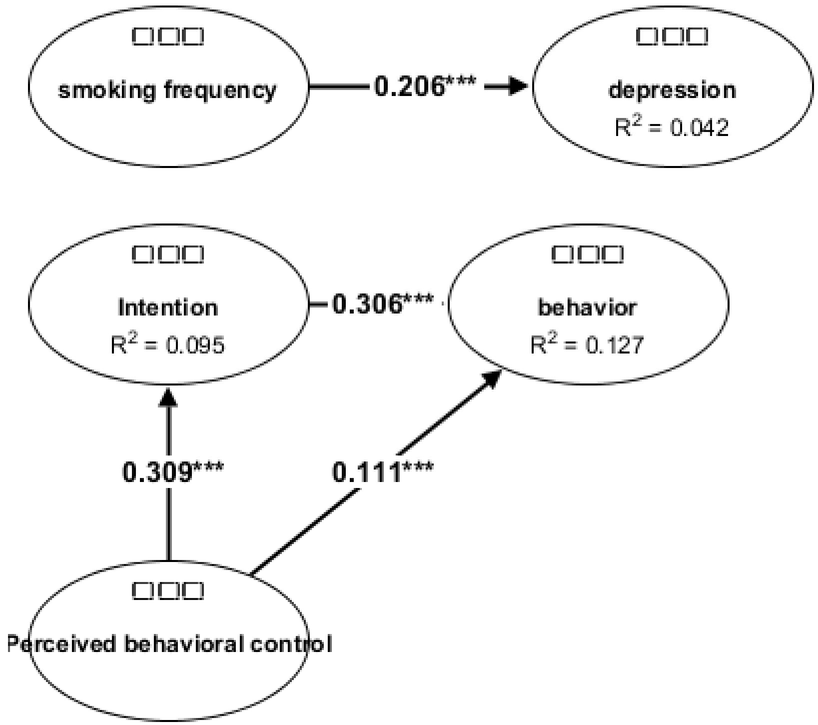
Willer, R. et al. (2015): What drives the gender gap in charitable giving? Lower empathy leads men to give less to poverty relief. In: *Social Science Research* 52, S. 83–98. DOI: 10.1016/j.ssresearch.2014.12.014



Structural relationships 2

Path Analysis

Quiz



Reverse Causality

Partial mediation

Quiz

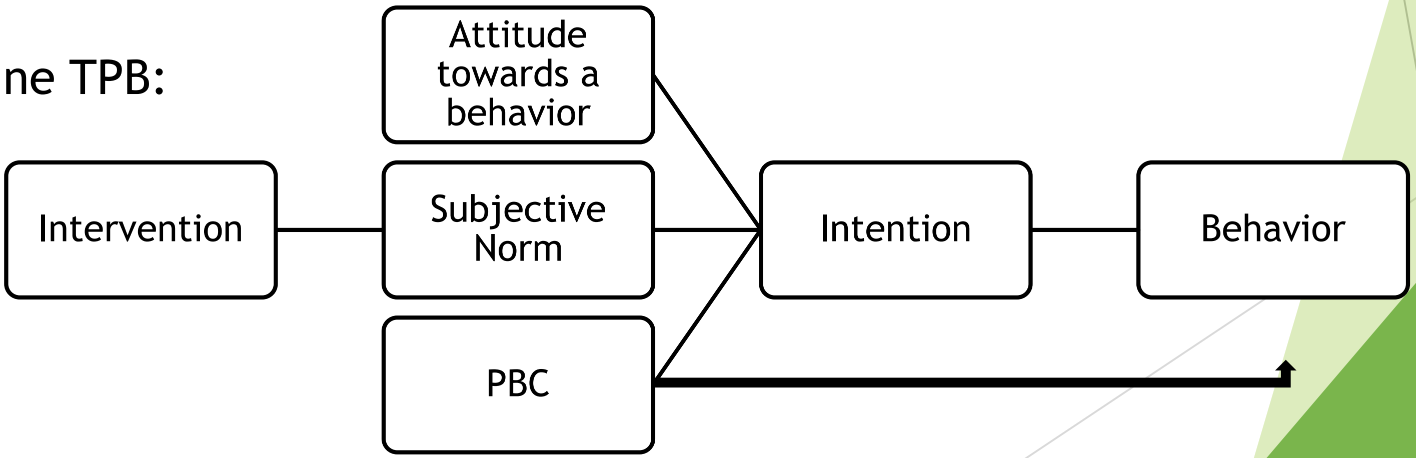
Can you name an example for
1. Mediation
2. Competitive mediation

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:



The concept of latent variables (LV)

Confirmatory Factor Analysis

- A variable is not directly observable:
  - Hidden variables, e.g. Racism, quality of life, intelligence,
  - Abstract concepts or mental states, e.g. confidence, extraversion, wisdom, Sustainable behavior in the food domain, welfare

  ⇒ LV reduces dimensionality

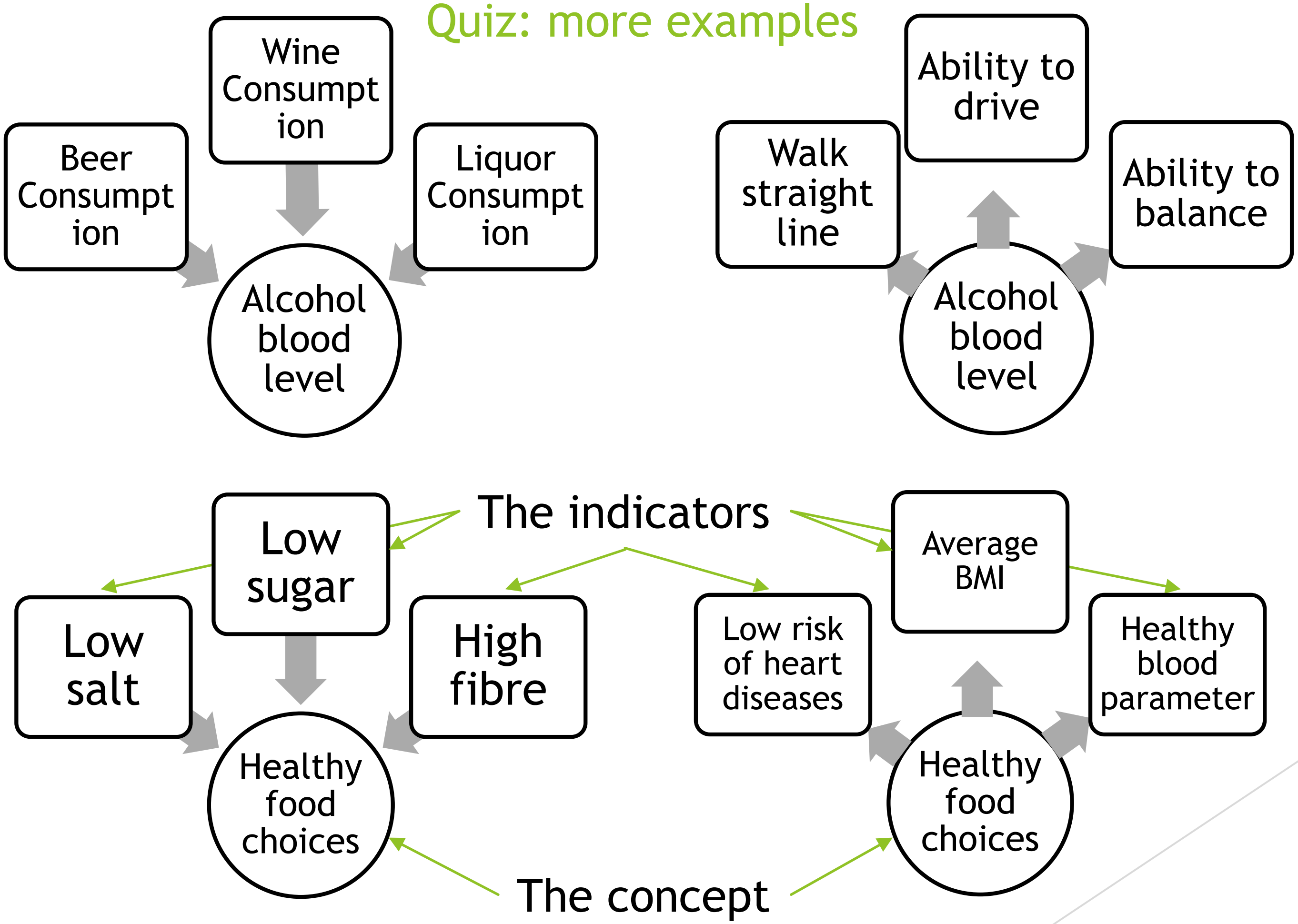
  ⇒ LV lead to improved concept measurements, i.e. less measurement error

  ⇒ LV leads to inter-individual, i.e. generalizable concepts

Confirmatory Factor Analysis

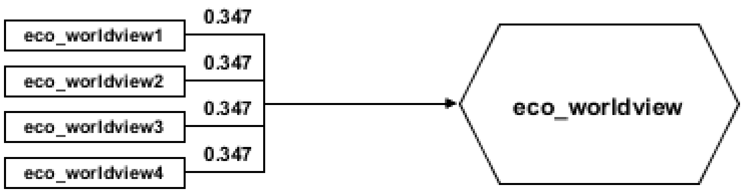
Latent Variables: formative vs. reflective

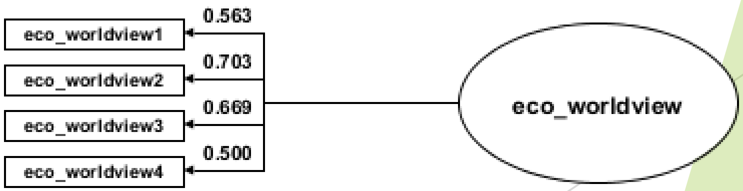
Quiz: more examples



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 should be correlated
Dropping an indicator from the measurement model increases error of LV measurement	Dropping an indicator may not induce error





Confirmatory Factor Analysis

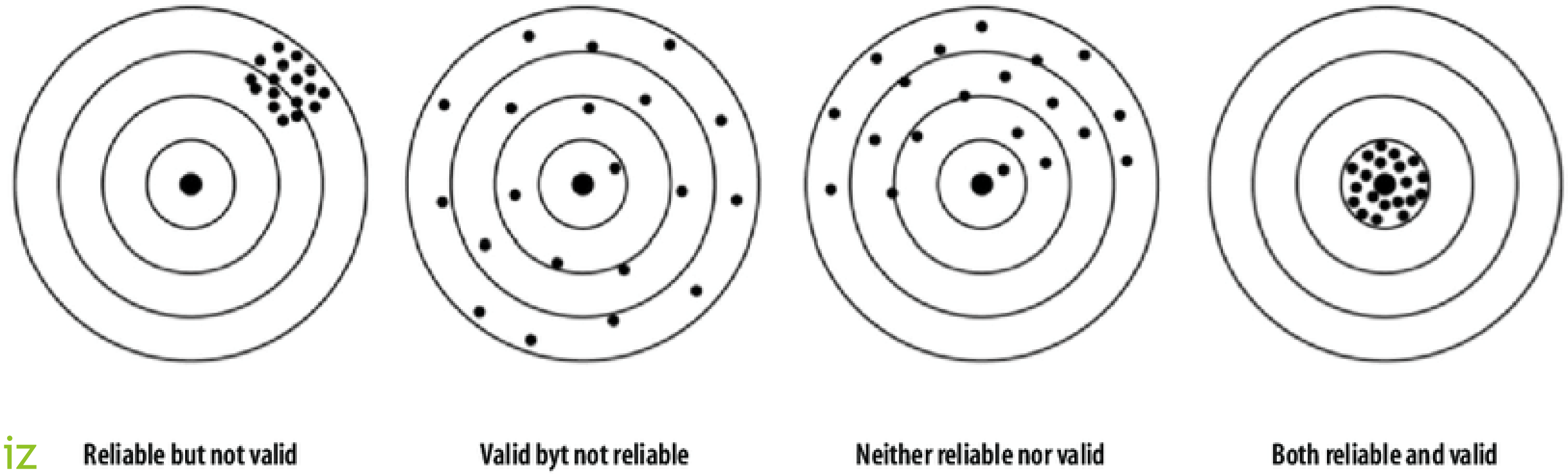
Latent Variables: formative vs. reflective

	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

Confirmatory Factor Analysis

Reliability and Validity (Measurement error):



Quiz

- Estimate Reliability and Validity

Assessing Reflective Measurement Models

Confirmatory Factor Analysis

Reliability and Validity – capture measurement error:

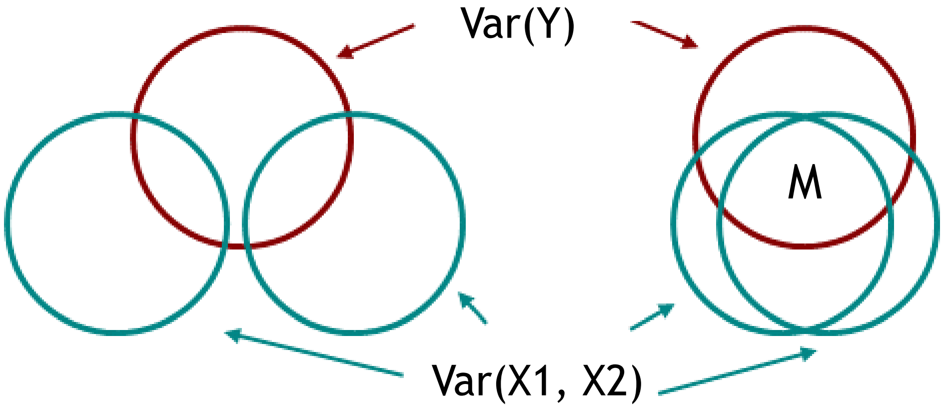
1. **Concept Reliability: Cronbach's alpha, Dijkstra-Henseler's rho**)
2. Concept Validity
   - **Convergence Validity:** Average Variance extracted, i.e. **AVE** (similar to explained variance concepts or LV explanatory power for indicator variance)
   - **Discriminant Validity** (Heterotrait-monotrait 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

Confirmatory Factor Analysis

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



- External validity: empirical confirmation

Value Belief Norm Theory (VBN)

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

Values:
Altruistic,
Egoistic,
Biospheric,
Schwartz values

Belief 1:
Ecological worldview (NEP-scale)

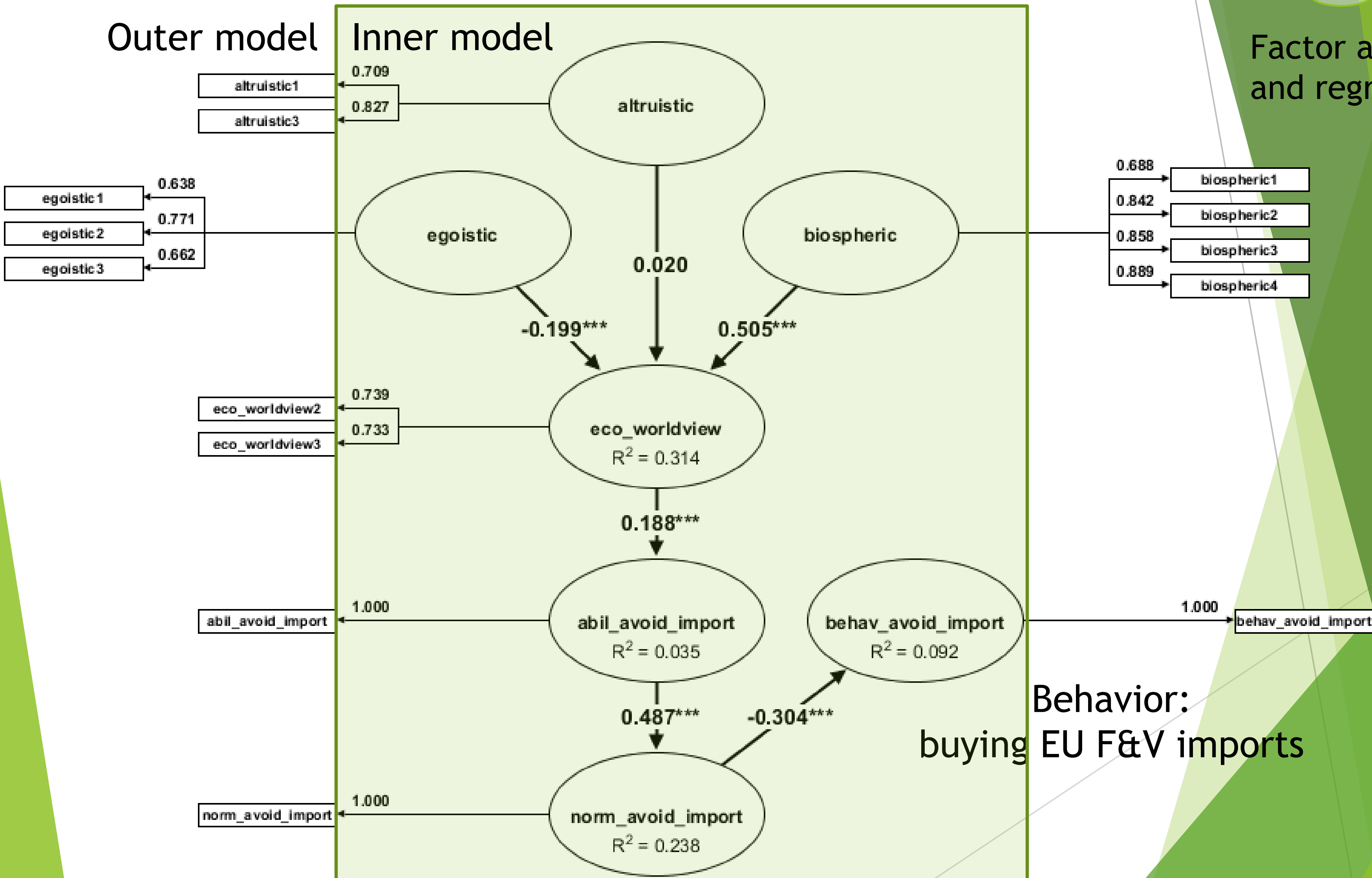
Belief 3:
Perceived ability to reduce threat

Norm:
Sense of obligation to take action

Sustainable behavior in the Food domain

Empricial Data: VBN-Theory (M1)

Factor analysis and regressions!



Behavior:
buying EU F&V imports

Assessing Reflective Measurement Models: Construct Reliability

Confirmatory Factor Analysis

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

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

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

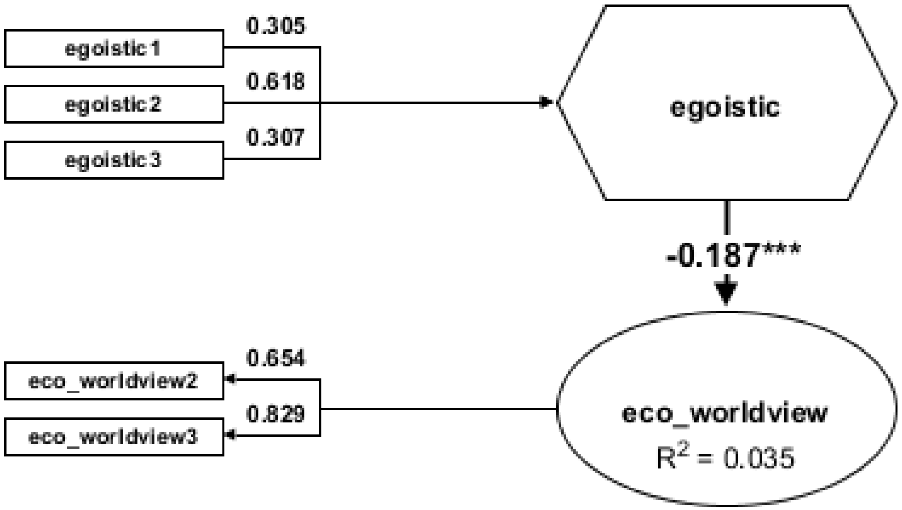
Forner Larcker Criterion or HTMT: An LV should explain variance of its own indicators better than indicators of other LVs

Solution: Merging or Dropping of constructs or indicators

Assessing formative Measurement models

Confirmatory Factor Analysis

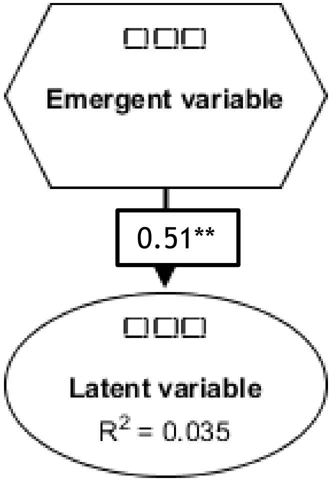
- You may argue Egoistic Values based on Schwartz are not reflected by importance of authority, wealth, social power and influence, but egoistic values consist of these 4 attributes:



Confirmatory Factor Analysis

Assessing formative Measurement models: External validity

- 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

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

Assessing formative Measurement models: Indicator relevance

Confirmatory Factor Analysis

- 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

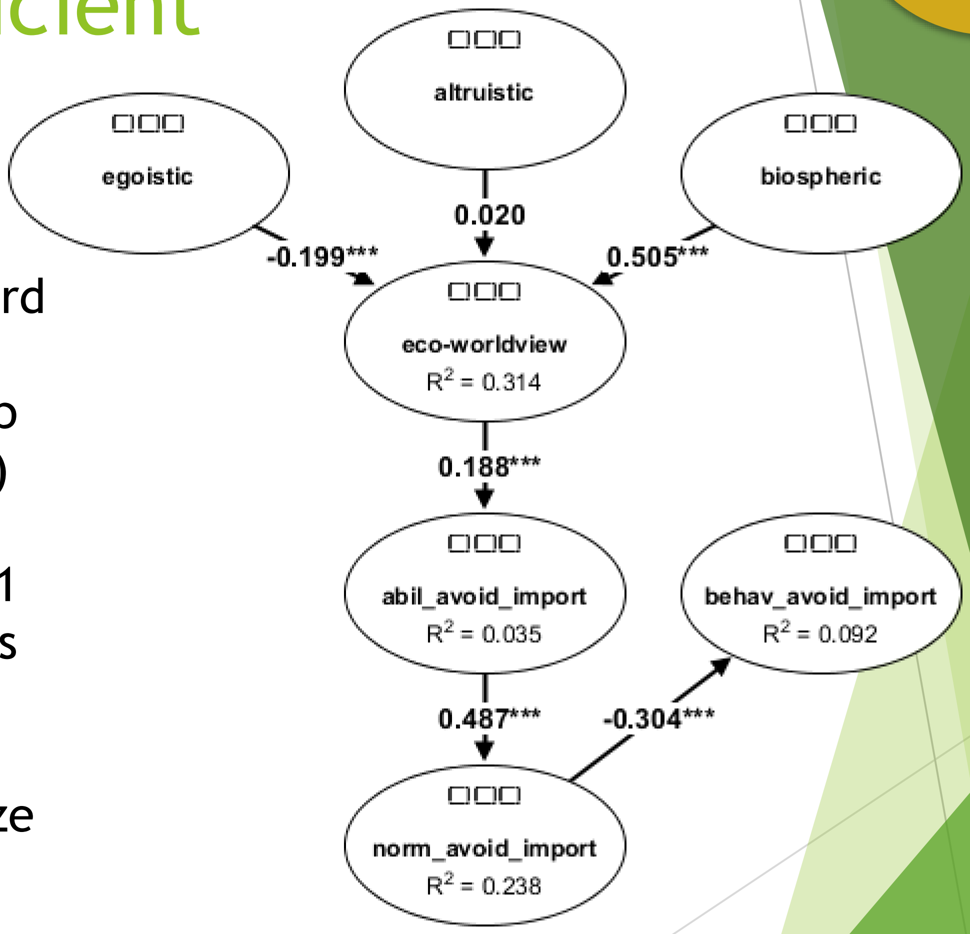
Assessing the Structural models: The path coefficient

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
4. b holds if everything else is constant

⇒ Interpret sign

⇒ Interpret relative effect size



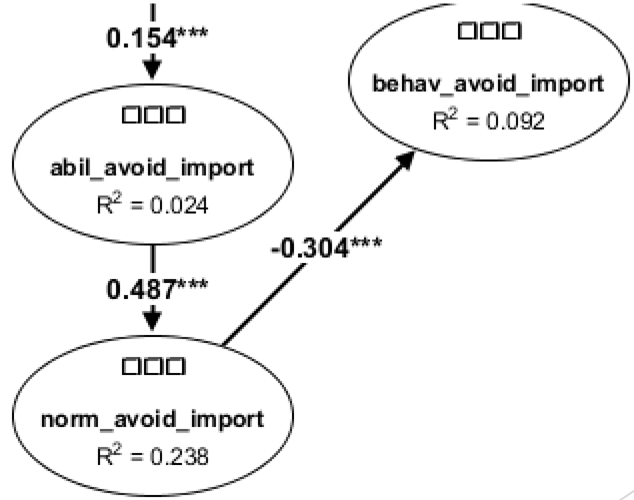
Quiz: What is the range of values b can obtain?

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

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



Quiz: Do I get a larger f^2 for „abil“ if I model „abil“ directly on the behavior?

Assessing the Structural models: Significance and Bootstrapping

Multiple Regression

Can 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

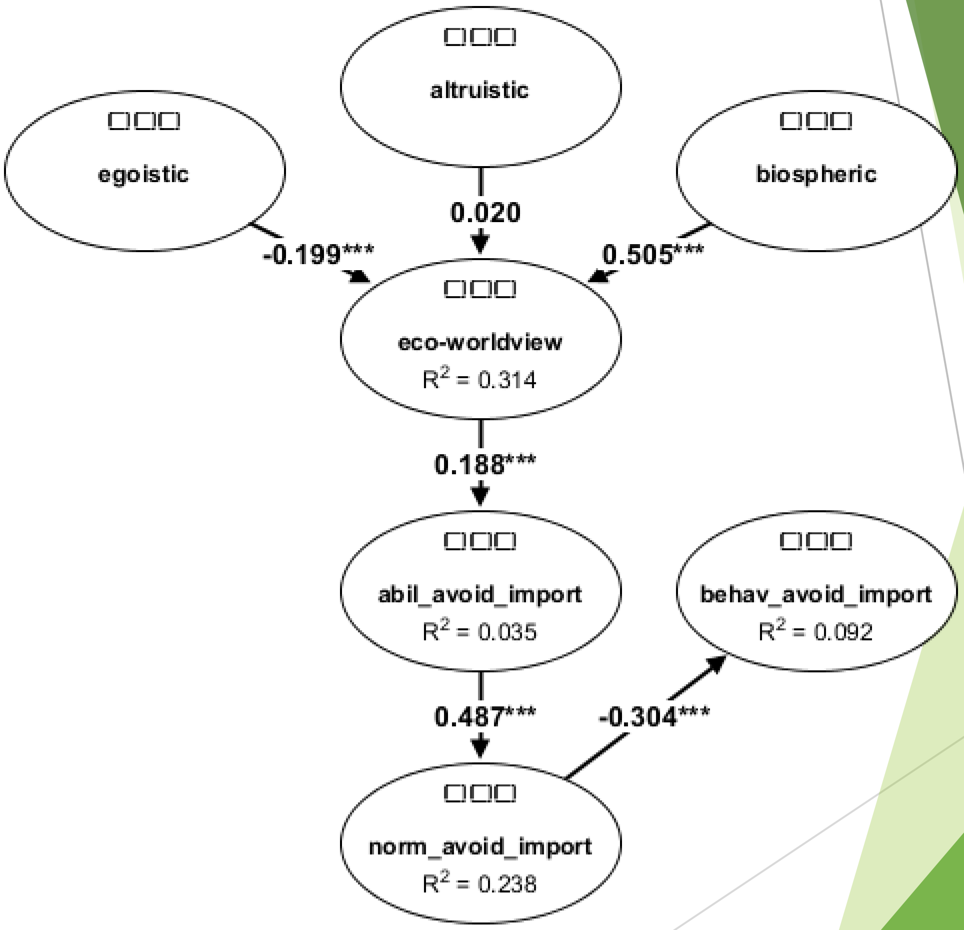
Example: A binary Adanco test

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

Assessing the Structural models: R^2

Multiple Regression

(Adj.) R^2: measures within sample prediction
⇒ Here R^2 is not global model fit criteria



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 better, SRMR<0,08 is considered a good model fit
- 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

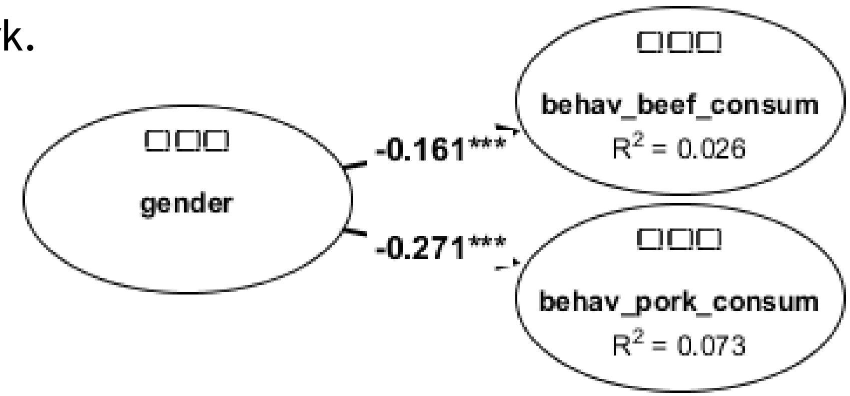
Now it's your turn: Tutorials 1 to 6

- Open a new project with Adanco
- Import data: NH_cleaned and open tutorials: https://github.com/dlemken/Structural-Equation-Modeling
  - 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

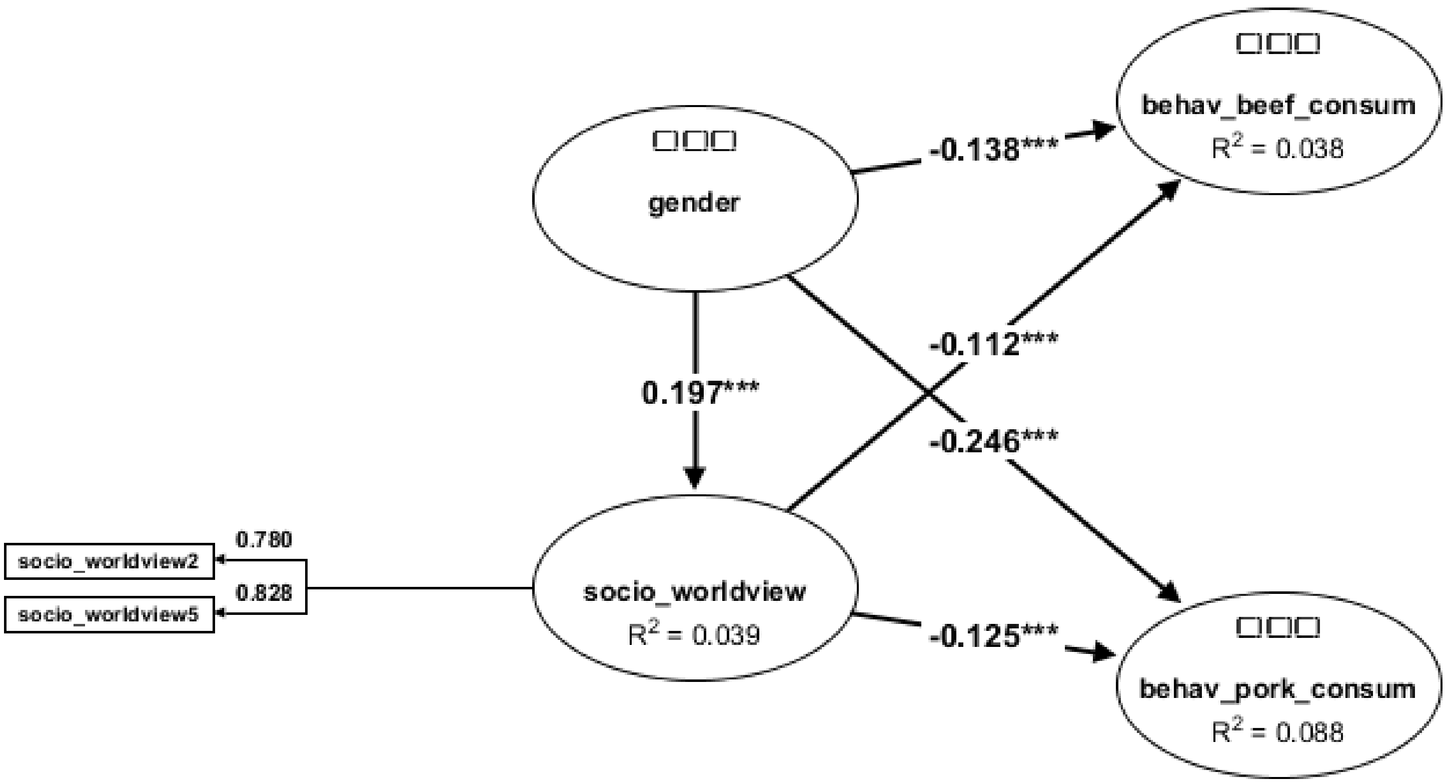
There is a gender gap in red meat consumption. Women consume less beef and particularly pork.



1. Can animal welfare related attitudes (here: socio_worldview2 and 5) explain the gender gap 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

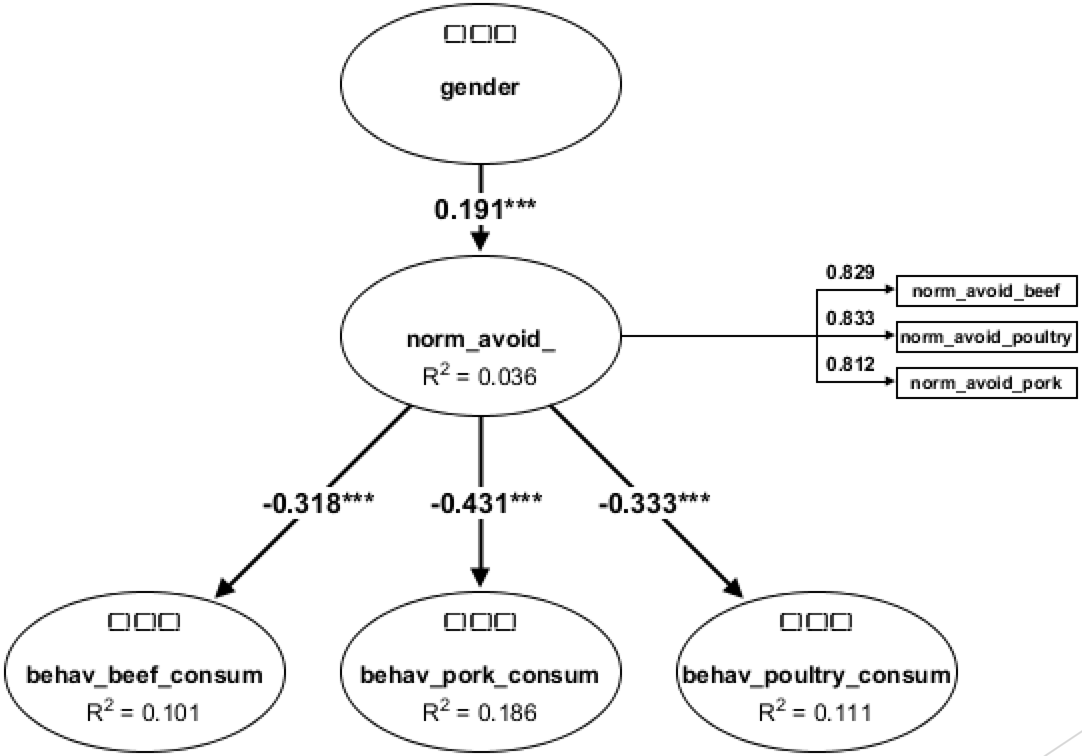
Solution 1



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

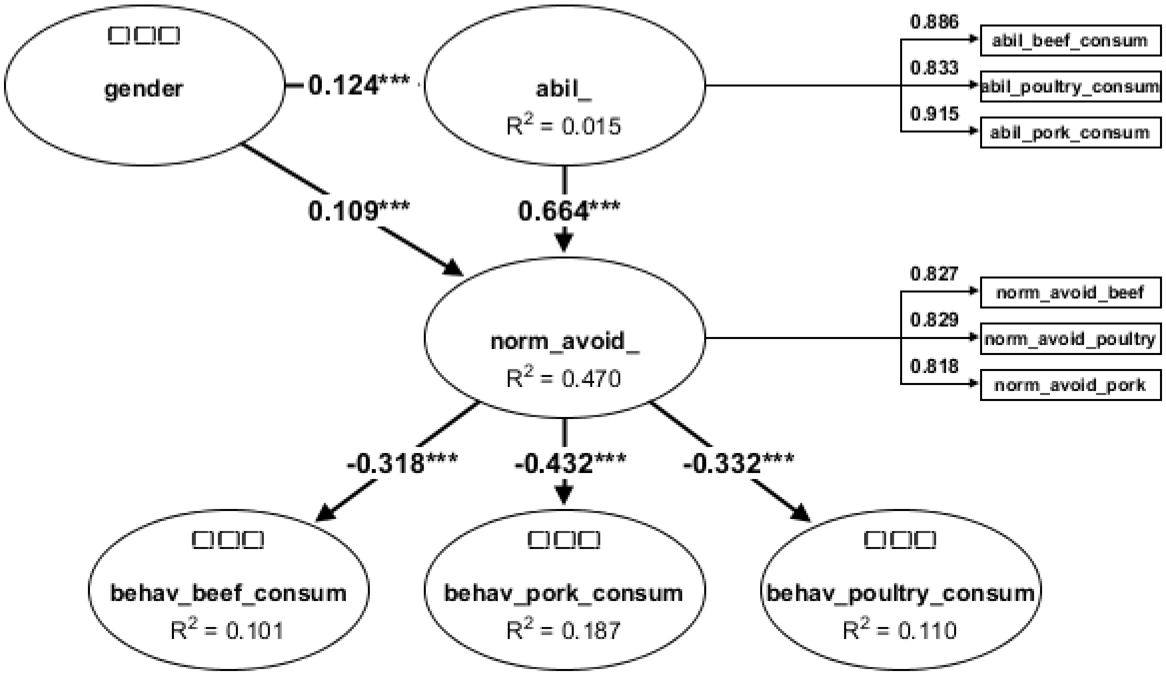
Tutorial 2

- One more try: We are interested in the felt norm to avoid meat consumption and why women experience the norm stronger than men. Can the perceived ability to avoid meat (abil_beef_consum, _pork, _poultry) explain the gender gap for the norm?



Solution 2

Complementary Partial mediation: Explains about half of the gender gap (M0: b=0.191, M1: b=0.109)

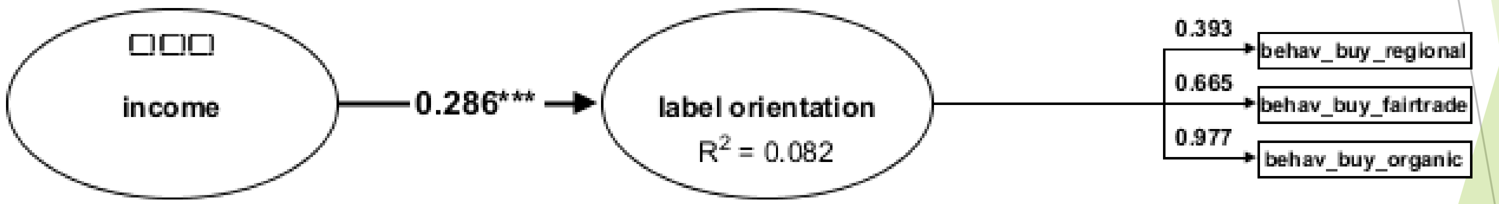


Note: Theoretically, We can test for significant differences of coefficients between models

Tutorial 3

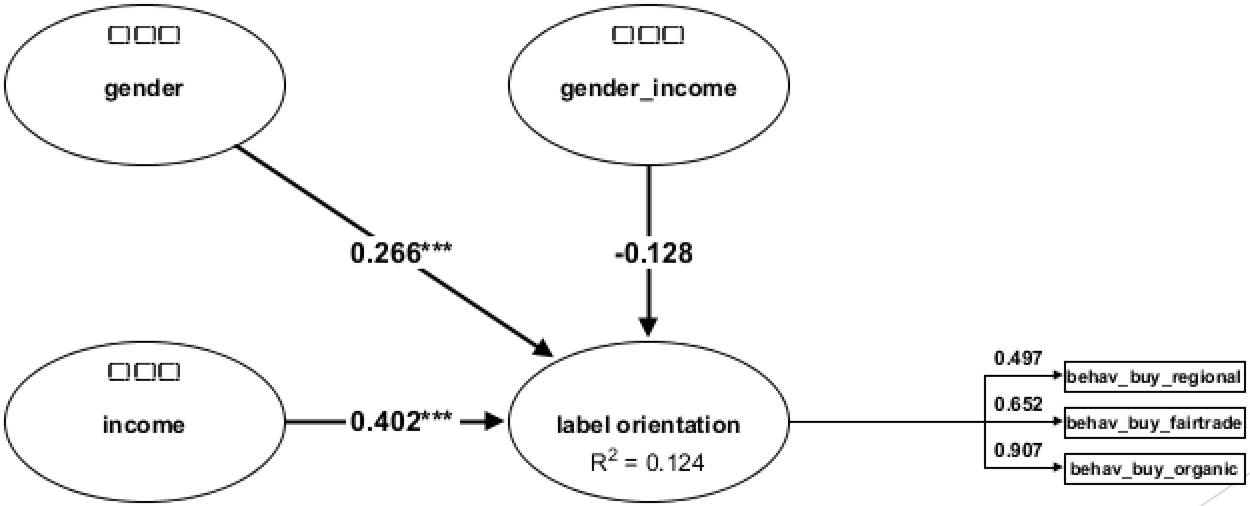
Consumers with a higher income consume more labelled food products (organic, regional, fairtrade). Is the effect of income moderated through gender differences?

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 orientation than for women. However, the variance is high. A significant relationship cannot be concluded. **We cannot confirm a moderator relationship**



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

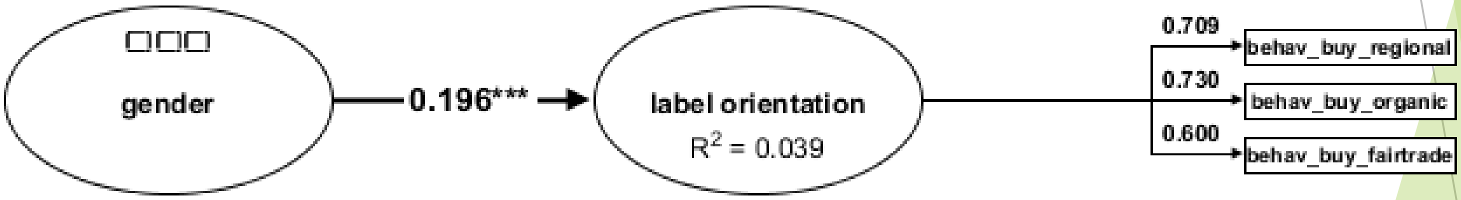
- Robust to small sample sizes (distribution assumption)
- **Robust to non-normal data**
- **Allows for formative measures**
- Prediction value
- Large number of indicators
- Theory development

Conclusions

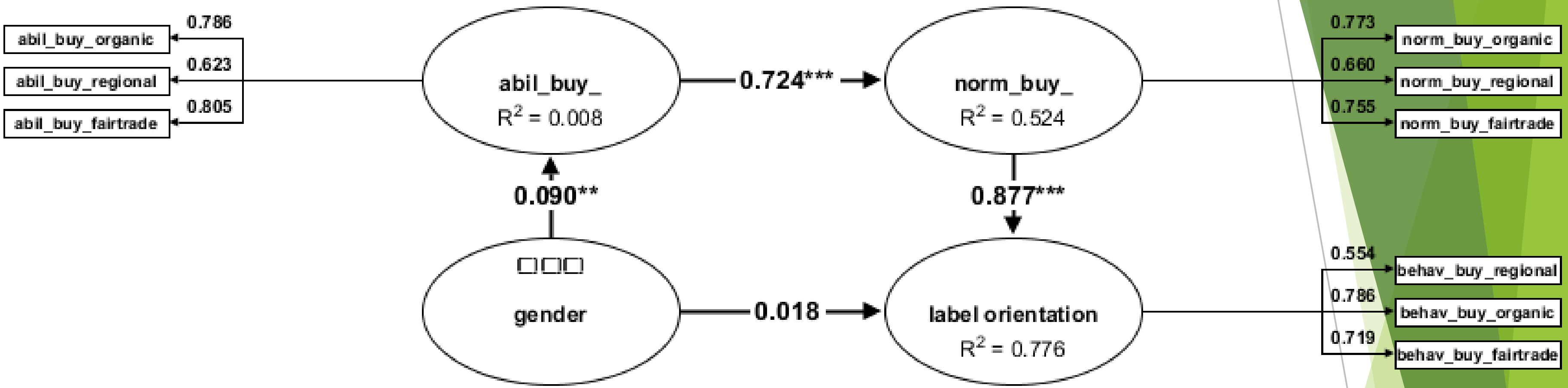
- Understand the structural relationships you hypothesize
- Measure your concepts as good as possible
- Provide evidence for both with PLS

Tutorial 4

1. What type of measurement problem do we create, if we model the mediation of ability and norm to buy labelled products (abil_, norm_: regional, organic, fairtrade) to predict gender differences in label orientation?
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 concept of label orientation better than the indicators themselves

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

2. Full mediation

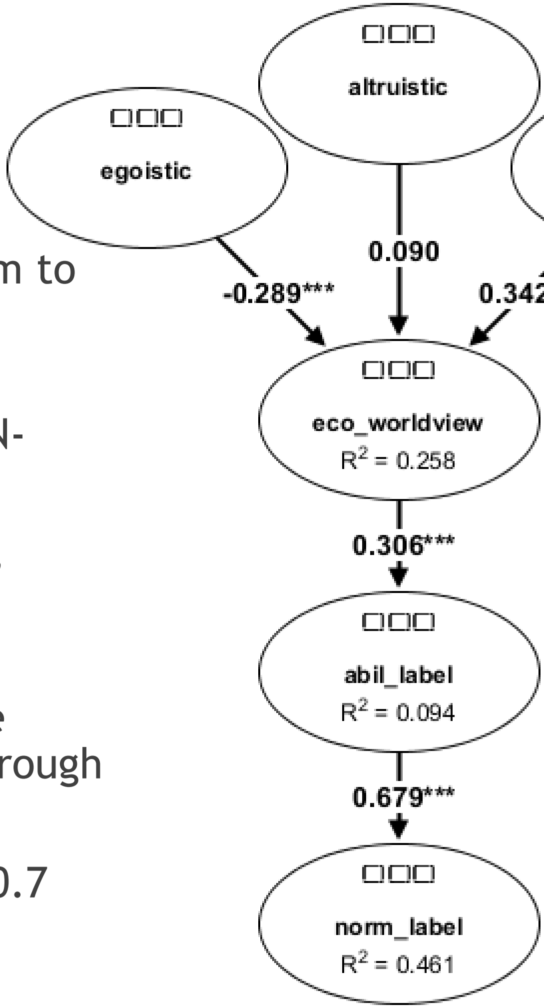
Tutorial 5

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.

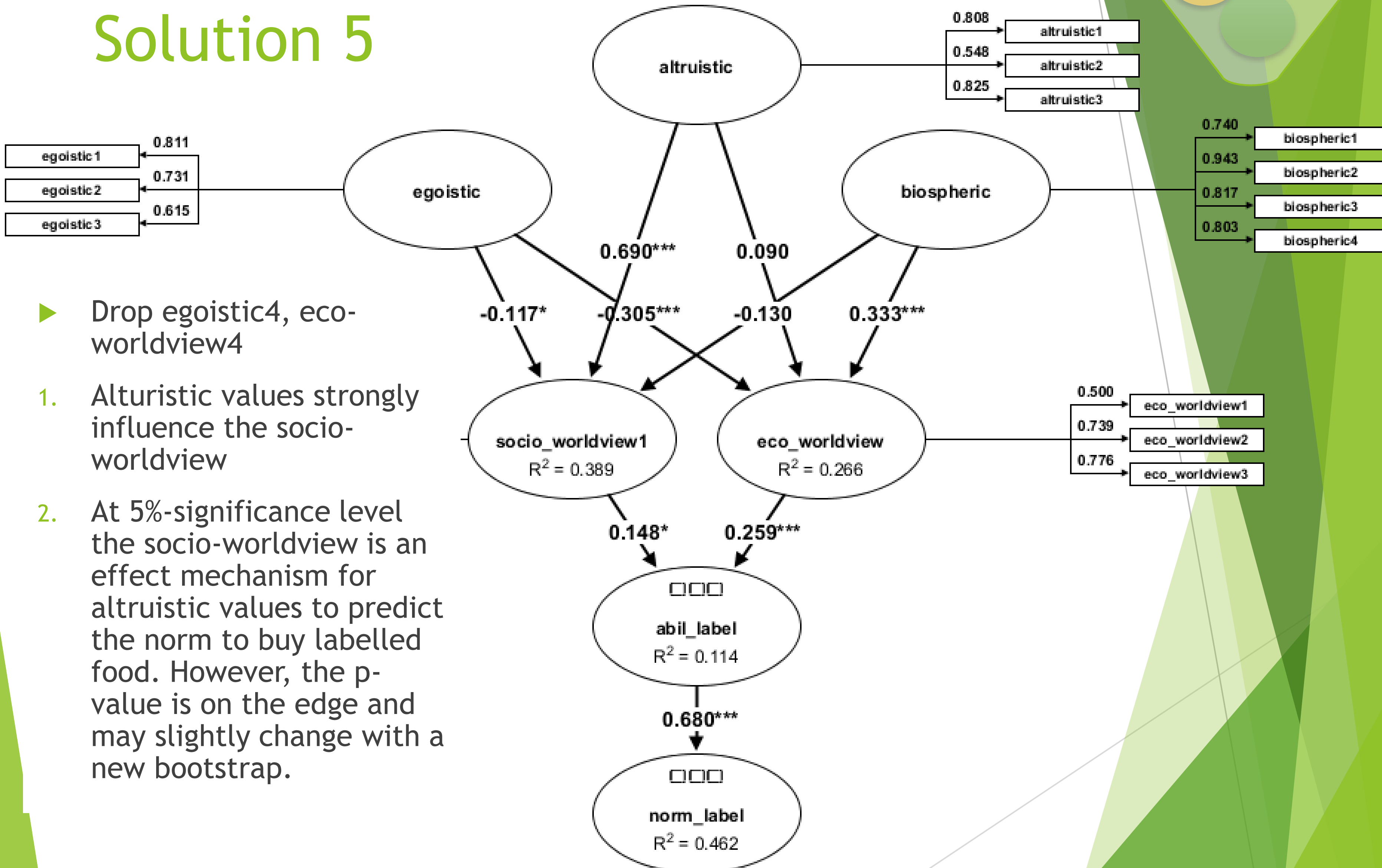
1. Which values (egoistic, altruistic, biospheric) influence the socio-worldview?
2. Do altruistic values matter to the norm to buy labelled products through the socio-worldview?

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



Item name	Wording (Likert-Scale)
Socio_worldview1	Es sollte der Anspruch der Menschheit sein, dass niemand Hunger leiden muss
Socio_worldview3	Alle Menschen müssen die Möglichkeit haben ihre Grundbedürfnisse zu decken

Solution 5



- Drop egoistic4, eco-worldview4

1. Alturistic values strongly influence the socio-worldview
2. At 5%-significance level the socio-worldview is an effect mechanism for altruistic values to predict the norm to buy labelled food. However, the p-value is on the edge and may slightly change with a new bootstrap.

Tutorial 6

Build a comparative VBN model to predict the Norm to

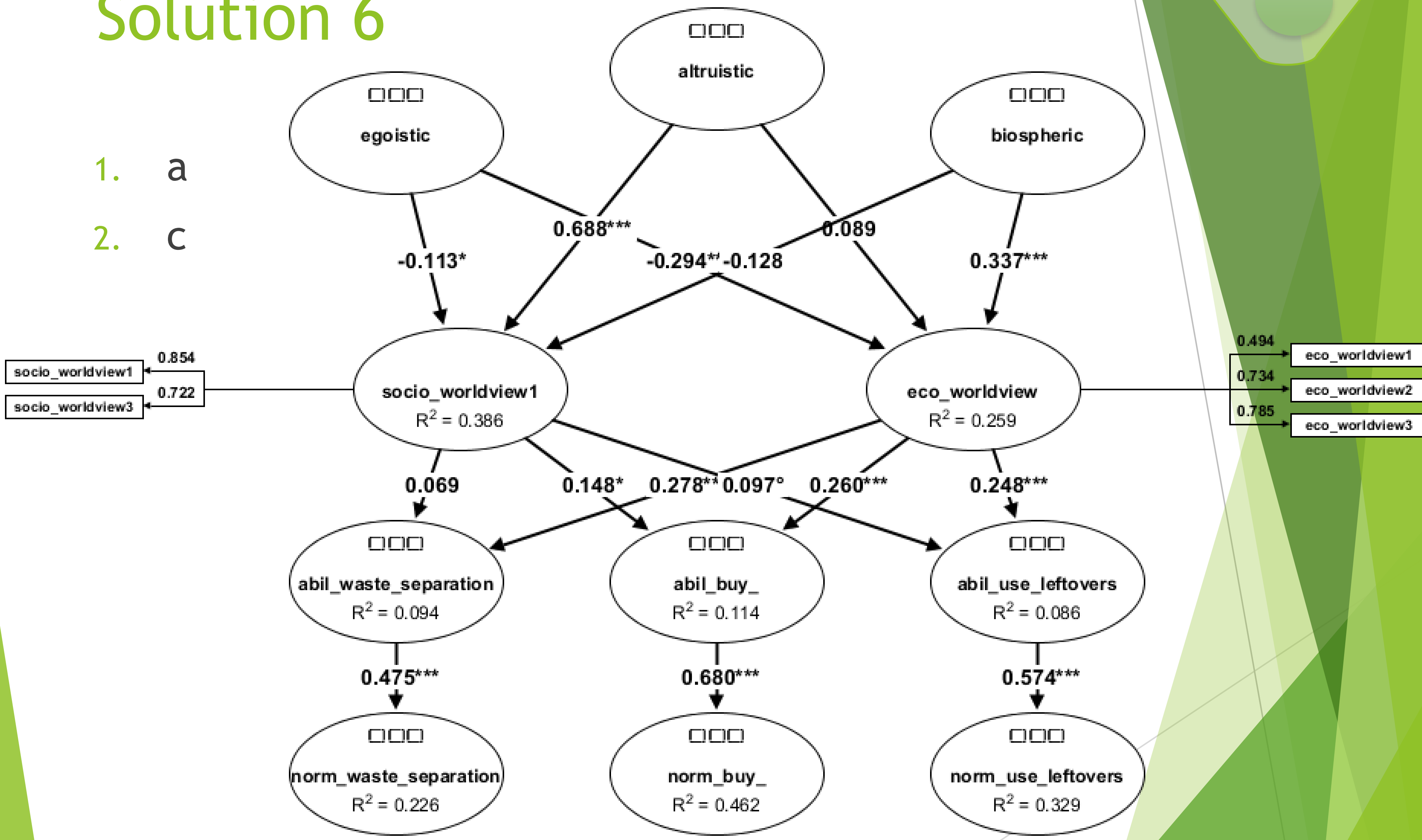
a. buy labelled products
b. to use leftovers
c. seperate waste

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
2. c



Outlook

- Causality: Group Comparisons allow tests for unobserved heterogeneity
- 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
- Books are available: Handbook of Partial Least Square: Concepts, Methods and Applications (Vinzi, Chin, Henseler 2010)
- Several options for model comparison and inter model testing (e.g. path coefficients between models

Good luck with your research!