Mediation and Moderation analyses

Maartje van Stralen - Mine Yildirim- Saskia te Velde

VU University medical center, EMGO Institute for Health and Care Research, Amsterdam

Today: morning session

- 9:00- 10:30 Lecture 1: Introduction to mediation analyses
  - Syllabus and Course goals
  - Definition Mediation analyses
  - Mediation analyses: single mediation analyses
  - Mediation analyses: multiple mediation analyses
- 10:30- 10:45 Coffee
- 10:45- 11:30: Practicum 1: Introduction to mediation analyses
  - Exercise 1.1: single mediation analyses Active plus project
  - Exercise 1.2: multiple mediation analyses Active plus project
  - Exercise 1.3: Bootstrapping
- 11:30- 12:00: Lecture 2: Introduction to moderation analyses
  - Definition Moderation analyses
  - How to conduct Moderation analyses
- 12:00- 12:30 Practicum 2: Introduction to moderation analyses
  - Exercise 2.1: moderation analyses DOiT project

Today: afternoon session

- 12:30-13:30 Lunch
- 13:30-15:15 Lecture 3: Advanced mediation and moderation models
  - Mediation with a categorical outcome variable
  - Moderation of a mediated effect
  - Longitudinal mediation analyses
- 15:15-15:30 Coffee
- 15:30-17:00 Practicum 3:
  - Practice with your own data, OR
  - Exercise 3.1: Longitudinal mediation analyses
  - Exercise 3.2: Moderated of a mediated effect
  - Exercise 3.3: Mediation analyses with a categorical outcome variable
- 17:00- ??:?? Drinks (at own costs)

Lecture 1: Introduction to mediation analyses

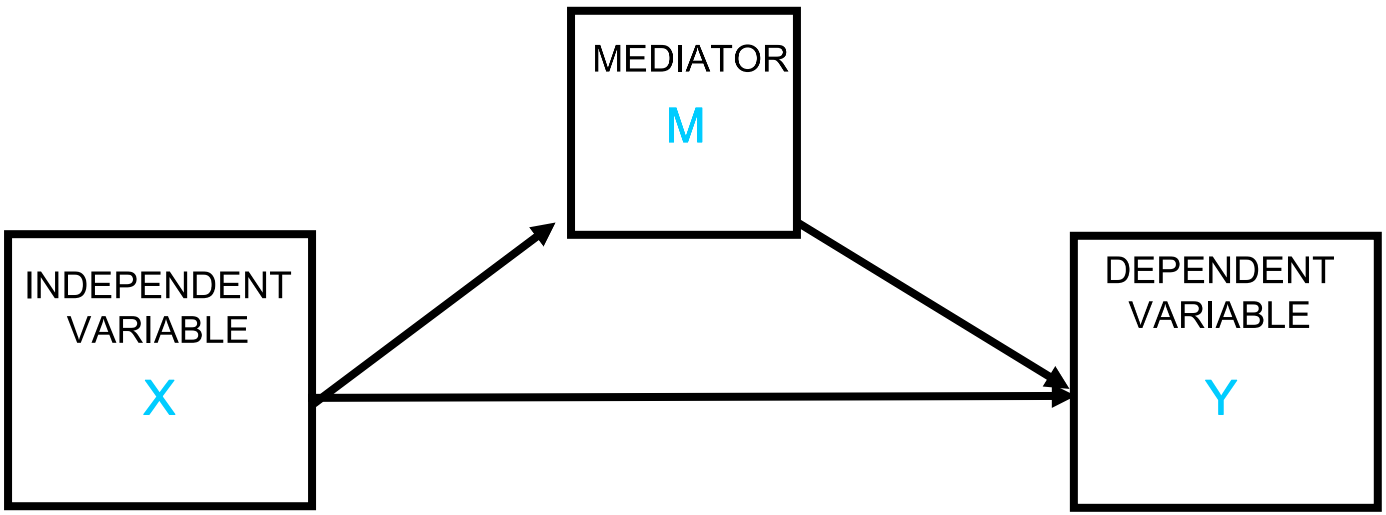
Mediation: syllabus and course goals

- Syllabus, 2 sections
  - Hand-outs
  - Exercises

- Course goals
  - Understand how to conduct ***single and multiple mediation*** models
  - Understand how we conduct ***moderation*** analyses
  - Understand how we conduct ***moderated mediation*** analyses
  - Understand how we conduct mediation analyses with ***categorical*** outcome variables
  - Understand how we conduct ***longitudinal*** mediator models

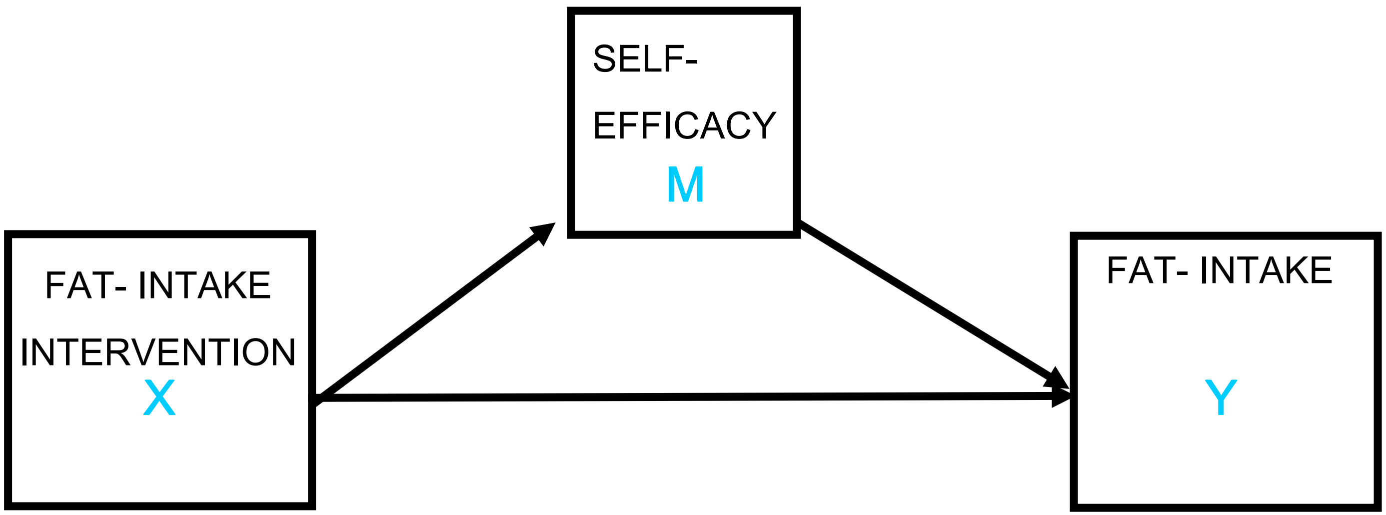
1.1 Single mediation analyses

Mediators: What?



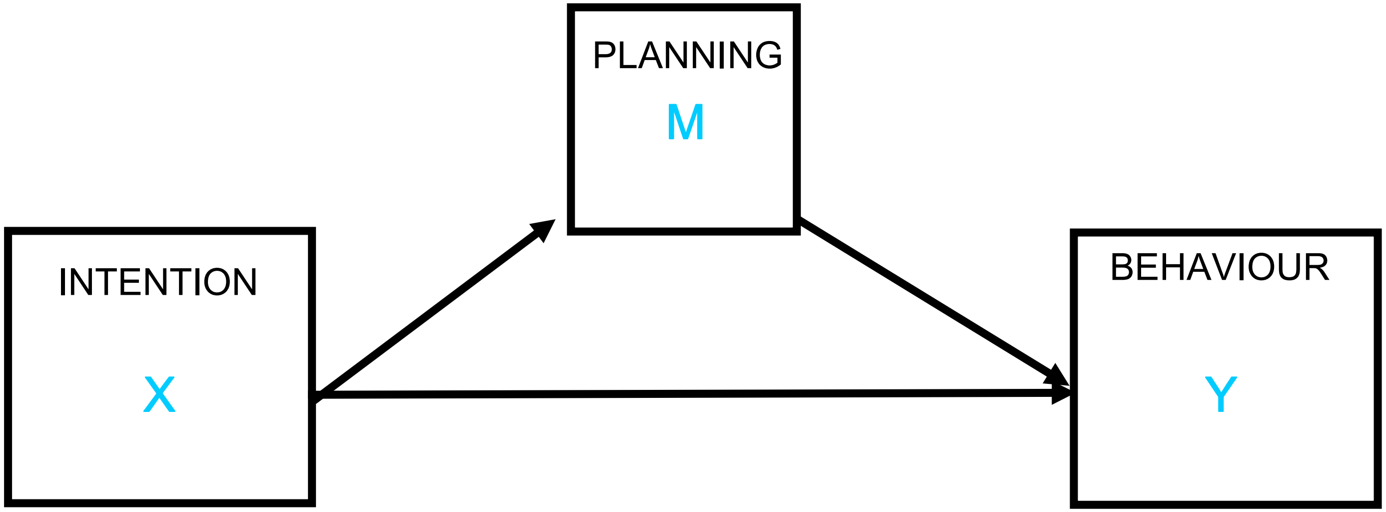
- A mediating variable is the intermediate in the causal sequence relating the independent to the dependent variable

Mediators: What?



- Working mechanisms intervention:
  - Mediator is variable through which an intervention exerts its effect on behavior

Mediation: What?



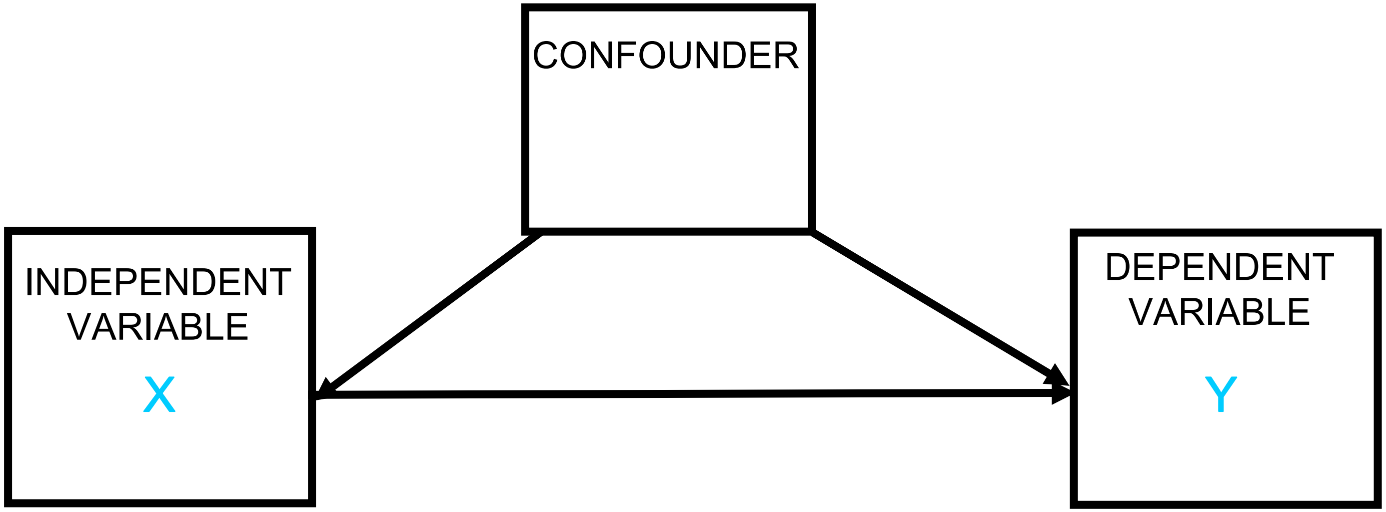
- Useful for theory testing/ theory augmentation

Mediation: What?

- Other third variables influencing associations between independent and dependent variables:
  - Confounder
  - Moderator
  - Covariate

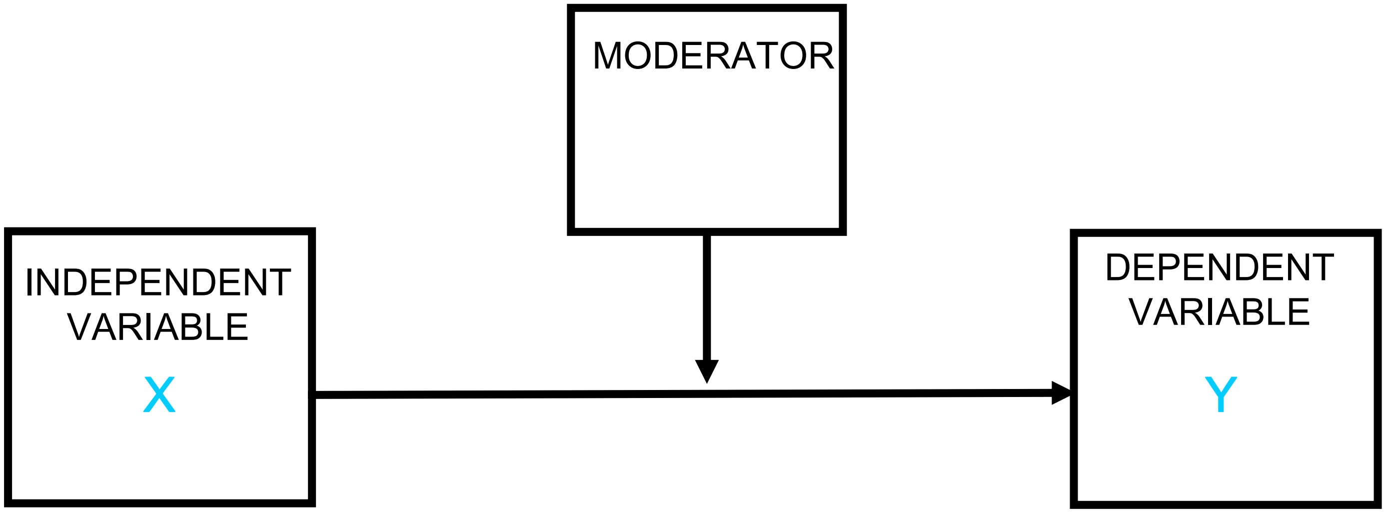
Mediator vs confounder

- Confounder is a variable related to two variables of interest that falsely obscures or accentuates the relation between them (Meinert & Tonascia, 1986)
- A confounder is, like a mediator, a variable that accounts for the relation between a predictor and an outcome (Baron & Kenny, 1986, p.1176). However it is not intermediate in a causal sequence.



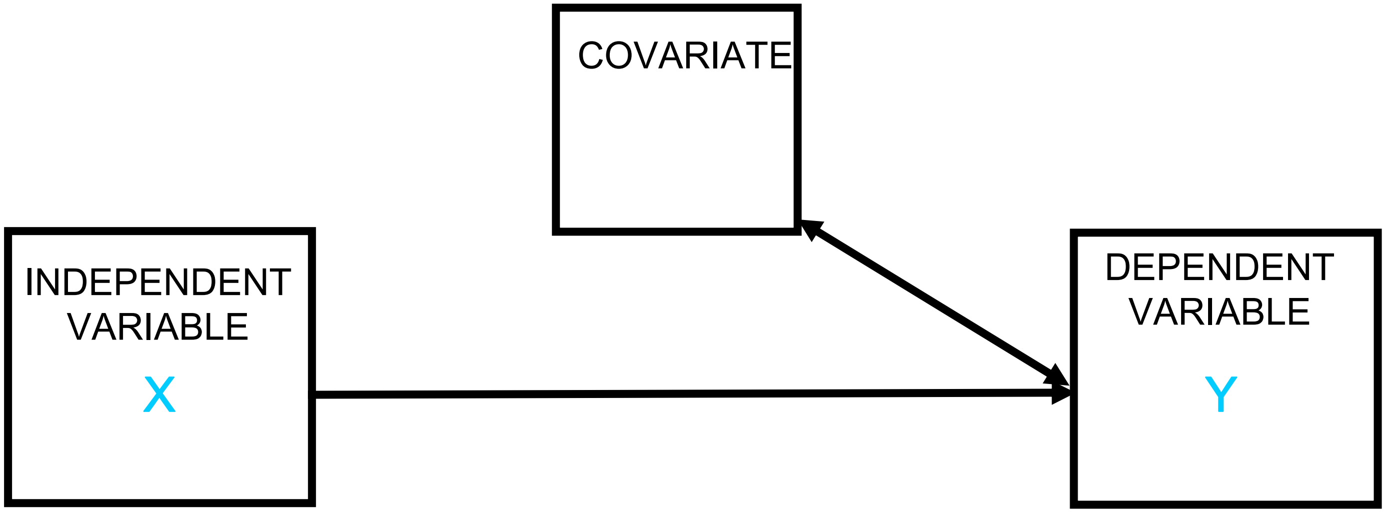
Mediator vs Moderator

- Moderator is a variable that affects the strength or direction of the relation between two variables. The variable is not intermediate in the causal sequence so it is not a mediator.
- Moderator is usually an interaction, the relation between X and Y depends on a third variable.

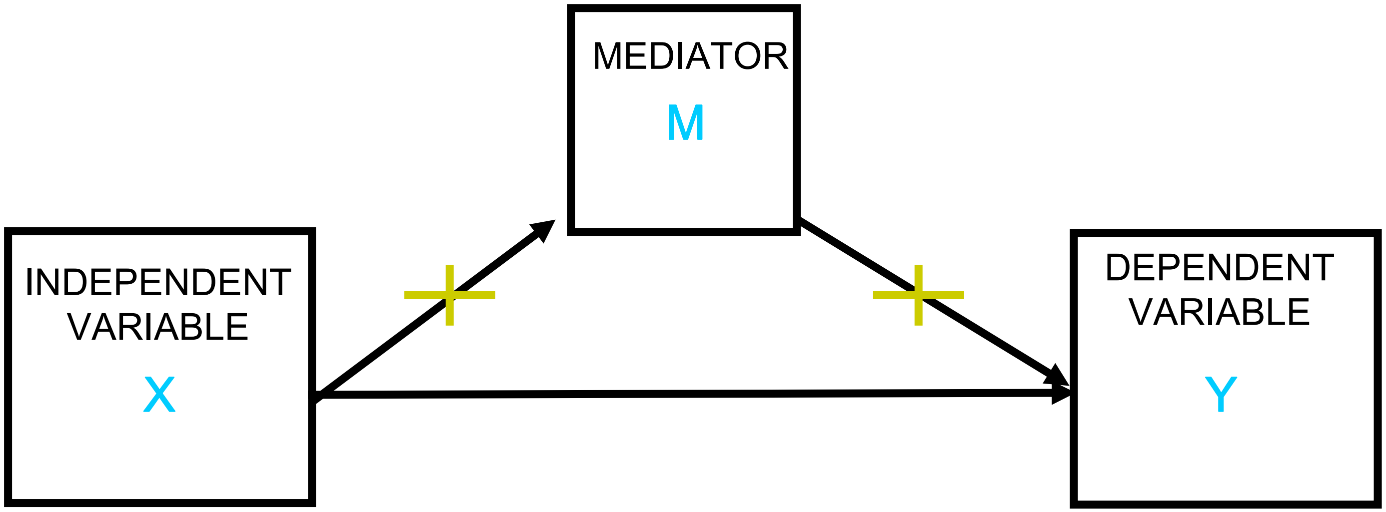


Mediator vs. covariate

- Covariate is a variable that is related to X or Y, or both X and Y, but is not in a causal sequence between X and Y, and does not change the relation between X and Y. Because it is related to the dependent variable it reduces unexplained variability in the dependent variable.
- A covariate is similar to a confounder but does not appreciably change the relation between X and Y so it is related to X and Y in a way that does not affect their relation with each other.

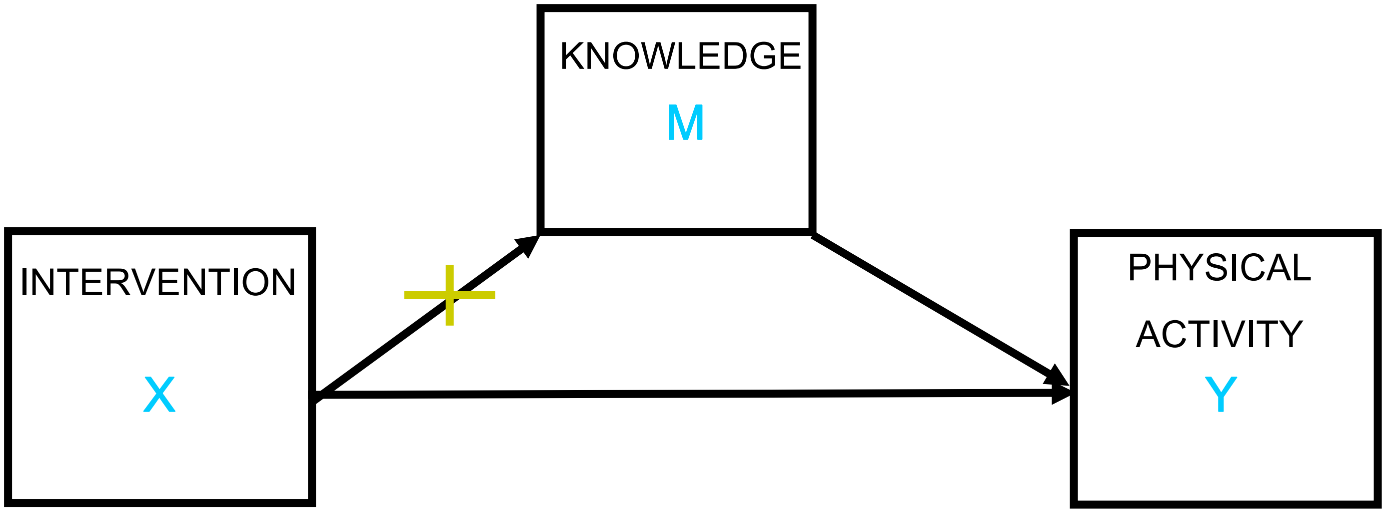


Mediators: Why?



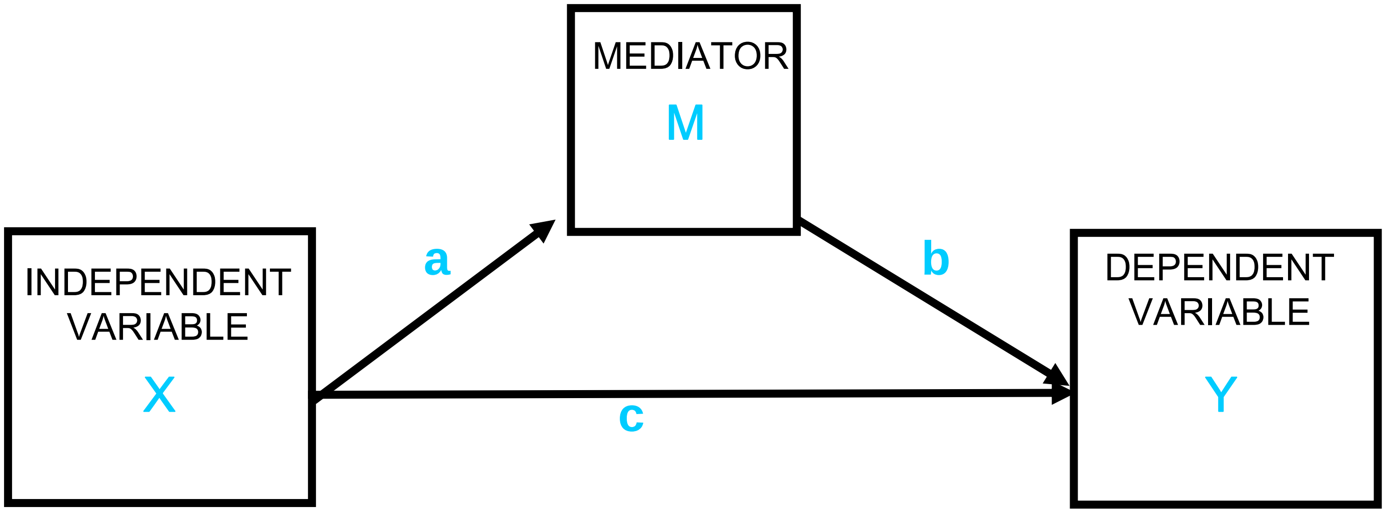
- Makes interventions more cost-effective
- More effective:
  - Mediator is important: to be included in future interventions
  - Intervention ineffective in changing mediator: Intervention needs adaptation
- Less costs:
  - Mediator not related to behavior: excluded from intervention

Mediators: Why? Example



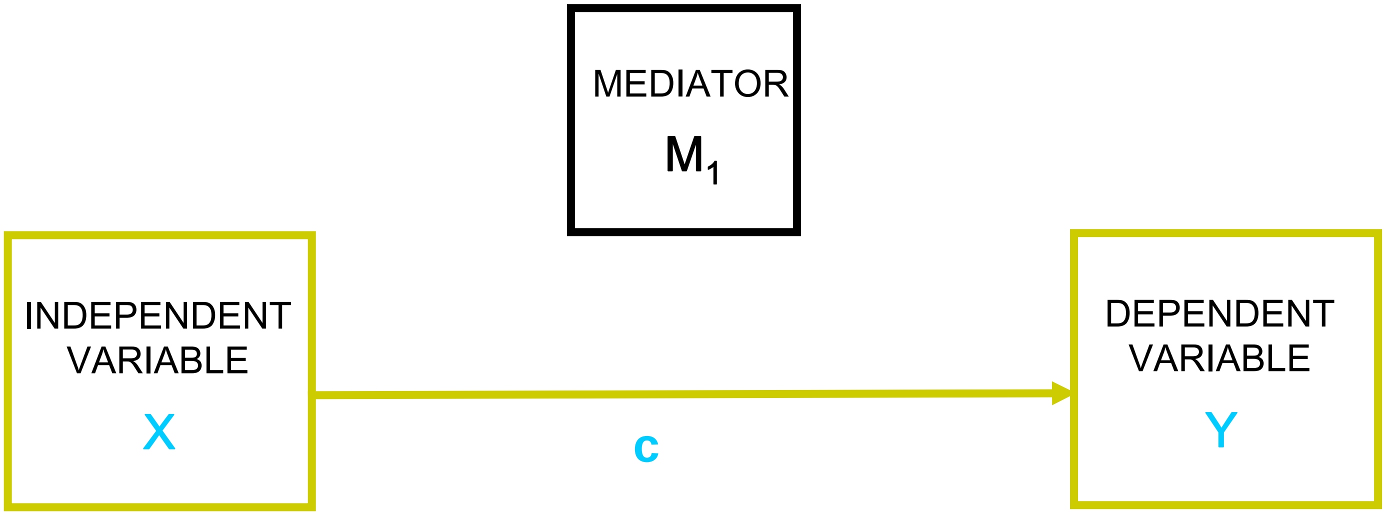
- Knowledge PA recommendation is not a mediator of the intervention effect on PA behaviour
- Knowledge is related to PA behaviour
- PA intervention did not affect knowledge of PA recommendation

Mediators: How?



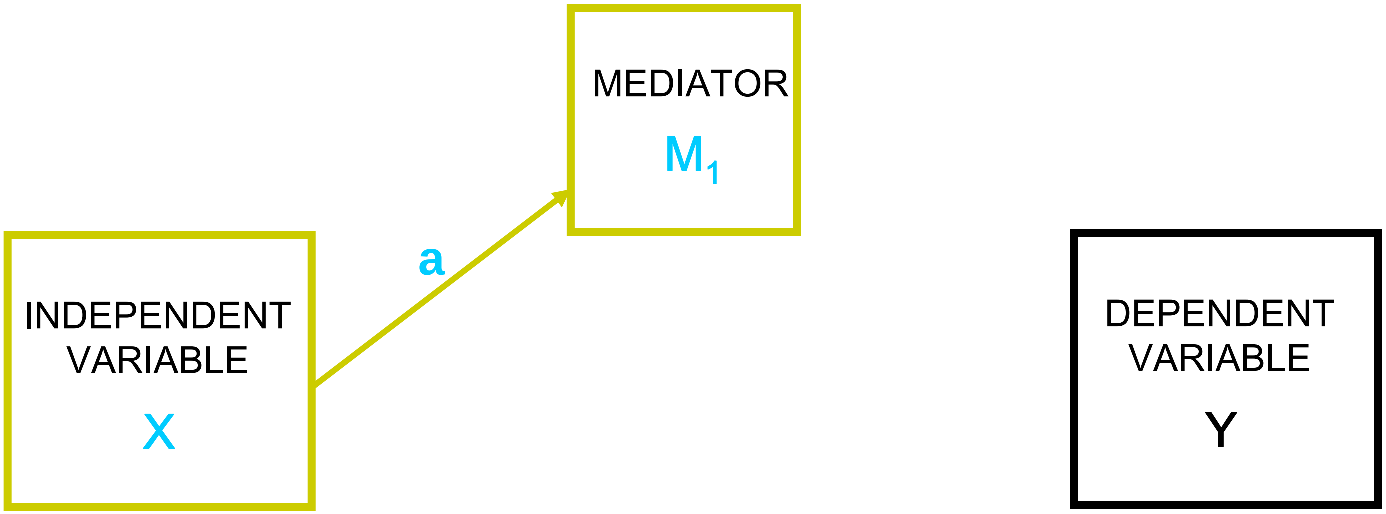
- Mediation analysis steps
  1. Main effect: association between predictor and outcome variable (c-path)
  2. Action theory (a-path)
  3. Conceptual theory (b-path)
  4. Test of mediation
     1. Joint significance test
     2. Difference in Coefficient test
     3. Product of Coefficient test

Step 1: association X and Y



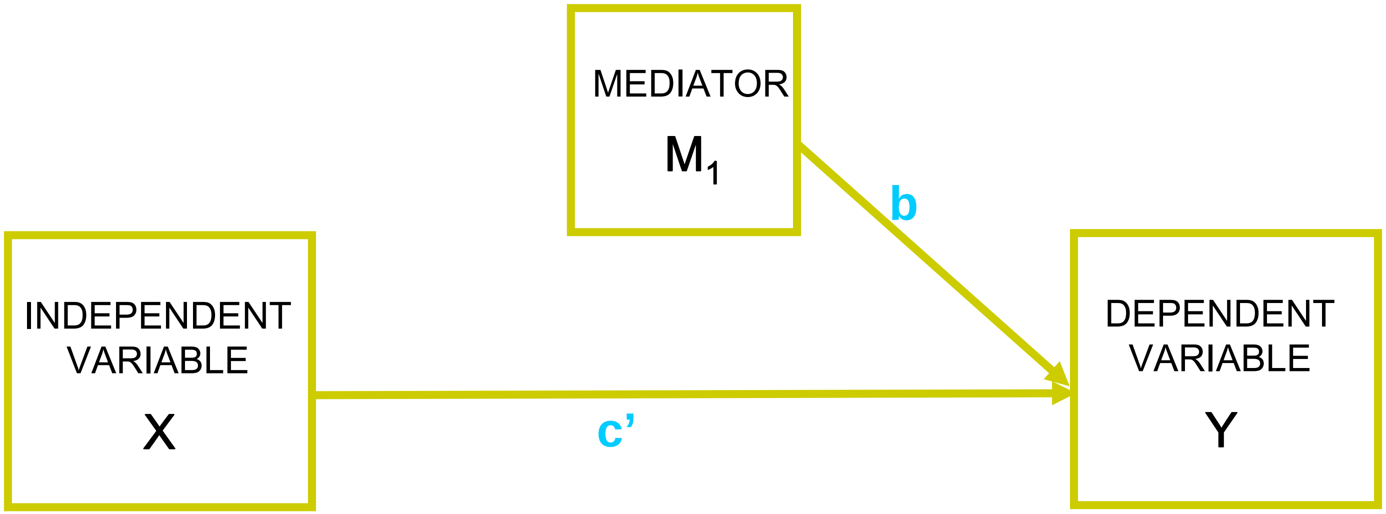
- Main effect: Assess association between predictor and outcome variable
- Y= c* X +ε
  - "c" is regression coefficient of the relationship between the independent variable (X) and the outcome variable (Y)

Step 2: Action theory



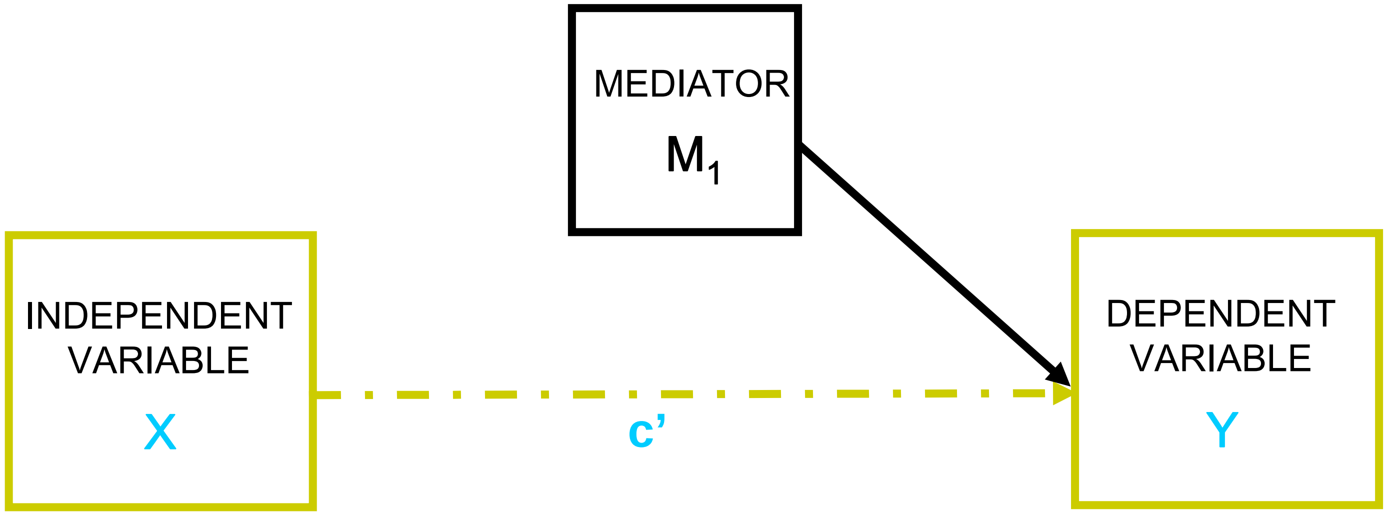
- Assess association between independent variable and mediator
- M= a * X + ε
  - "a" is regression coefficient of the relationship between X and M and needs to be significant

Step 3: Conceptual Theory



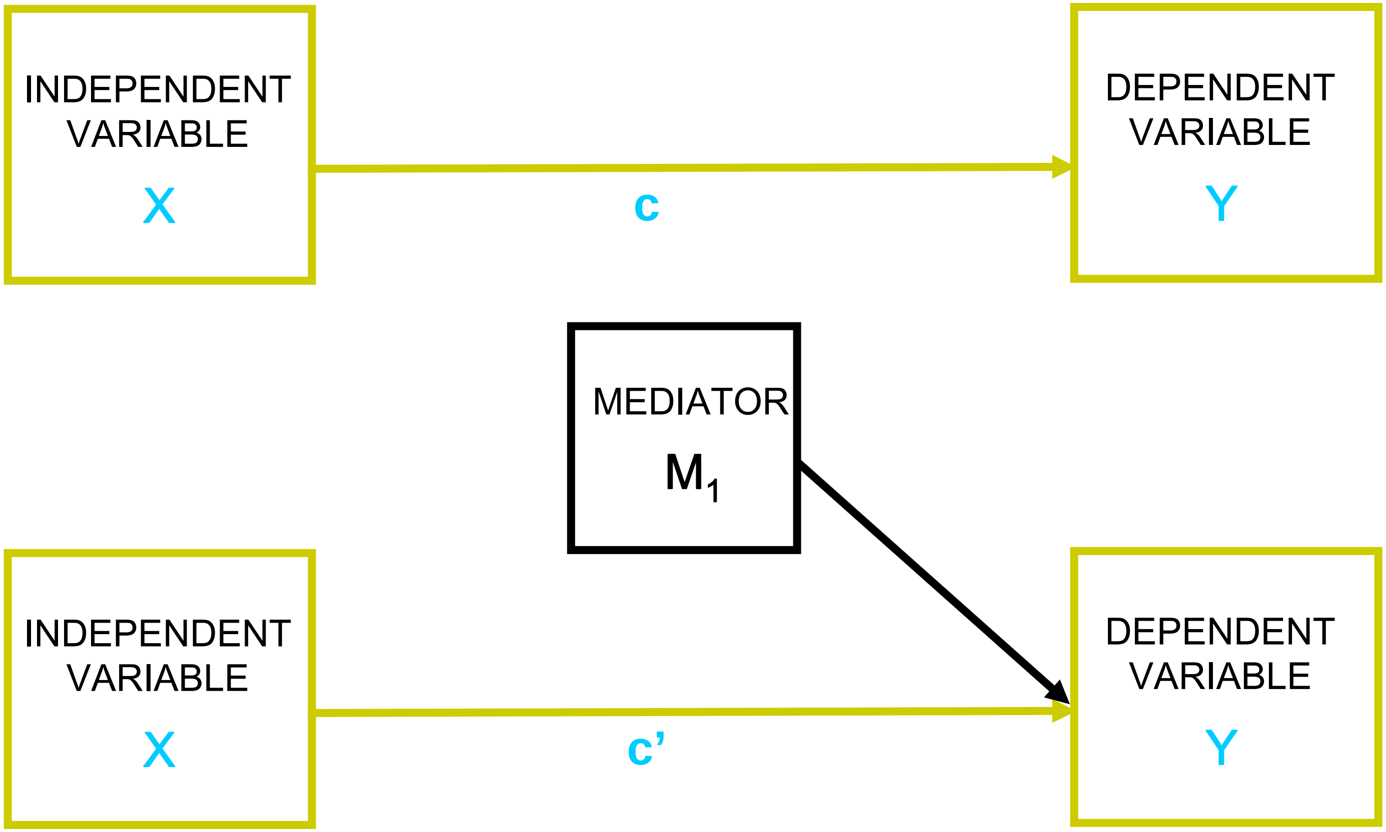
- Assess association between mediator and the outcome variable when controlled for the independent variable
- Y= b* M + c’* X +ε
  - “b” is the regression coefficient relating the mediator (M) and outcome variable (Y) controlled for the independent variable (X)
  - “c’” is the regression coefficient relating the independent variable (X) to the outcome variable (Y) controlled for the mediator (M)

Step 4: Test of mediation (1/3)



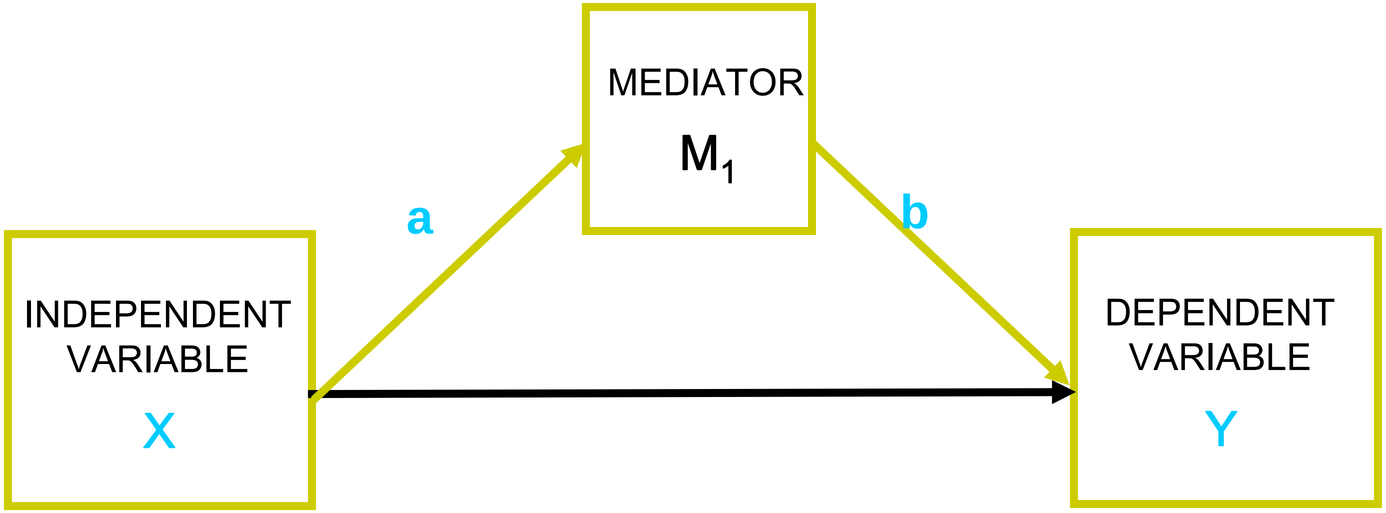
- Baron and Kenny (1986) causal steps
  - Full mediation: c'= 0
  - Partial mediation: c'< c

Step 4: Test of mediation (2/3)



- Difference in coefficient test (c-c')

Step 4: Test of mediation (3/3)



- Product of coefficient test (a*b)
- a*b= c-c′

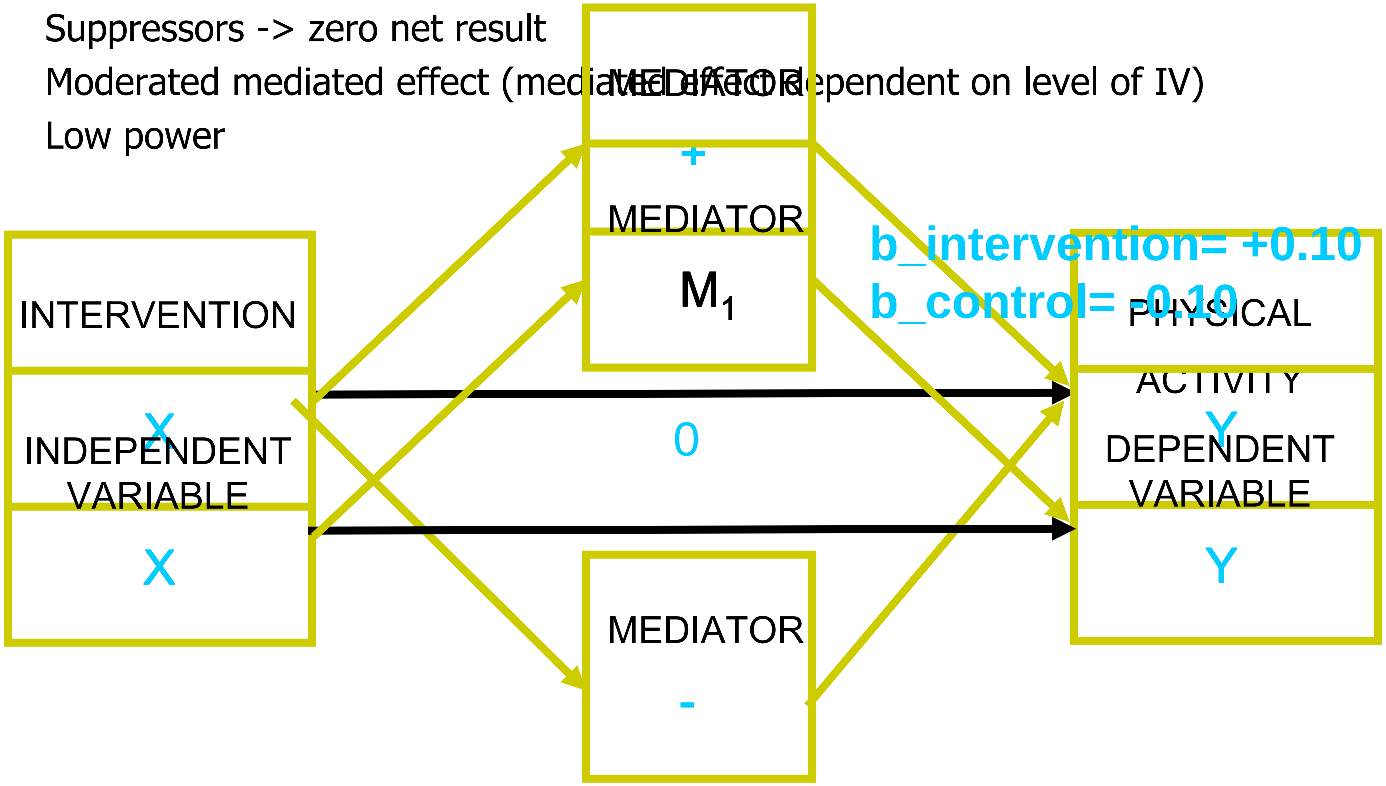
Different tests of mediation

- Baron and Kenny
  - Less powerful (small associations > 20,000 participants needed)
  - High risk of Type 1 Error (4 different analyses)
  - No test of significance of mediated effect

- Product of coefficient test/ Difference in coefficient test
  - More powerful than Baron and Kenny (no need to test main effect)
  - No need to have a **significant main effect**
  - Significance testing: **Sobel test**
  - However, Sobel test mainly useful in large samples (>400). Biased in smaller samples (assumption of normal distribution of a*b relationship is not tenable): **Bootstrapping**

Main effect: to be or not to be significant?

- Baron & Kenny: main effect needs to be significant
- Product of Coefficient test: no significant main effect necessary
- Reasons for a non-significant main effect are
  - Suppressors -> zero net result
  - Moderated mediated effect (mediated effect dependent on level of IV)
  - Low power



Significance test + % mediation

- Sobel test:
  1. calculate Standard Error

$$s_{First} = \sqrt{\hat{a}^2 s_{\hat{b}}^2 + \hat{b}^2 s_{\hat{a}}^2}$$

2. Wald statistics: divide estimate (a*b) by its standard error ($ab/\ s_{ab}$).

- If Z> 1.96 than mediated effect is significant

3. Calculate Confidence intervals

- Upper Confidence Interval (UCL) = $ab\ + Z_{.975}\ s_{ab}$
- Lower Confidence Interval (LCL) = $ab + Z_{.025}\ s_{ab}$
- The mediated effect is statistically significant if 0 is not in the interval.

Bootstrapping

- Multivariate normality:
  - Assumptions Sobel test: the paths that constitute the mediated effect, total effect and direct effect follow a multivariate normal distribution
  - In small samples this is normally not true
- Bootstrapping
  - Method to test significance and confidence interval around mediated effect
  - re-samples from the existing data many times (>1,000) to create an empirically-derived sampling distribution against which to test the sampled indirect effect (to get around the assumption of normality)
  - Because bootstrapping is based on random resampling of the data, bootstrap confidence intervals will differ slightly each time the macro is run as a result of the random sampling process. The more bootstrap samples that are requested, the less this variation between runs.
  - Allows you to create more accurate confidence intervals around mediated effect
  - Useful when sample size is smaller ($N < 400$)
  - Useful in multiple mediation models

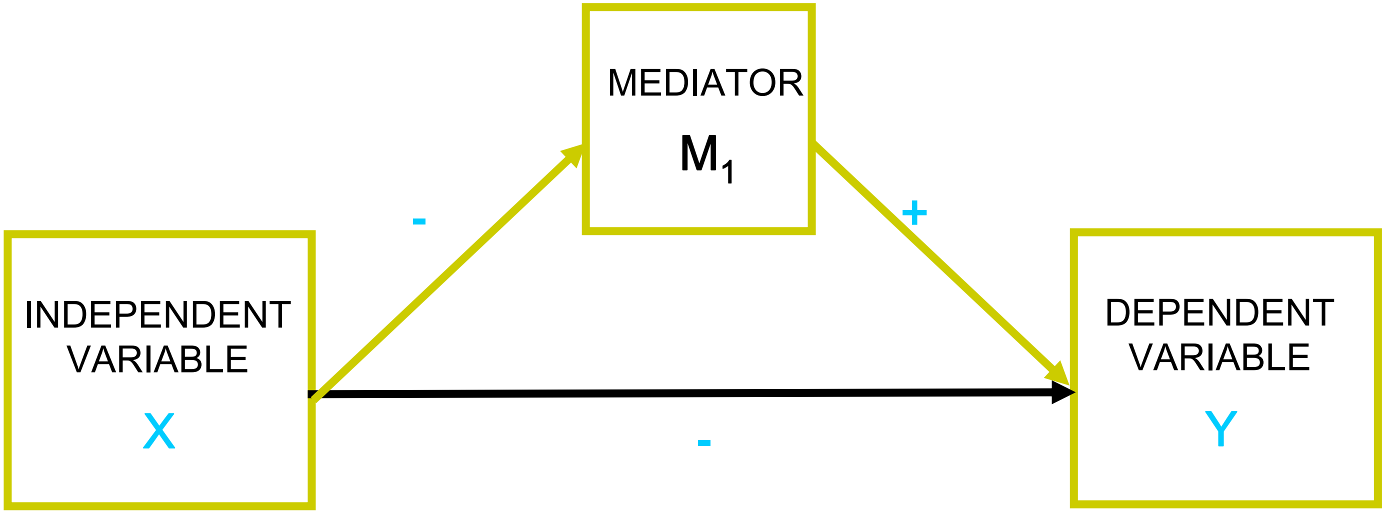
In addition….

Proportion mediation

- Proportion mediation: the proportion of the total effect of the predictor on the outcome variable that can be explained by the indirect (cq. mediated) effect
  - Total effect= Indirect effect (*ab*) + Direct effect (c')
  - Proportion mediated= (Indirect effect/ Total effect) * 100%
  - Proportion mediated= (*ab*)/ (*ab + c'*)* 100%

Suppressor effects

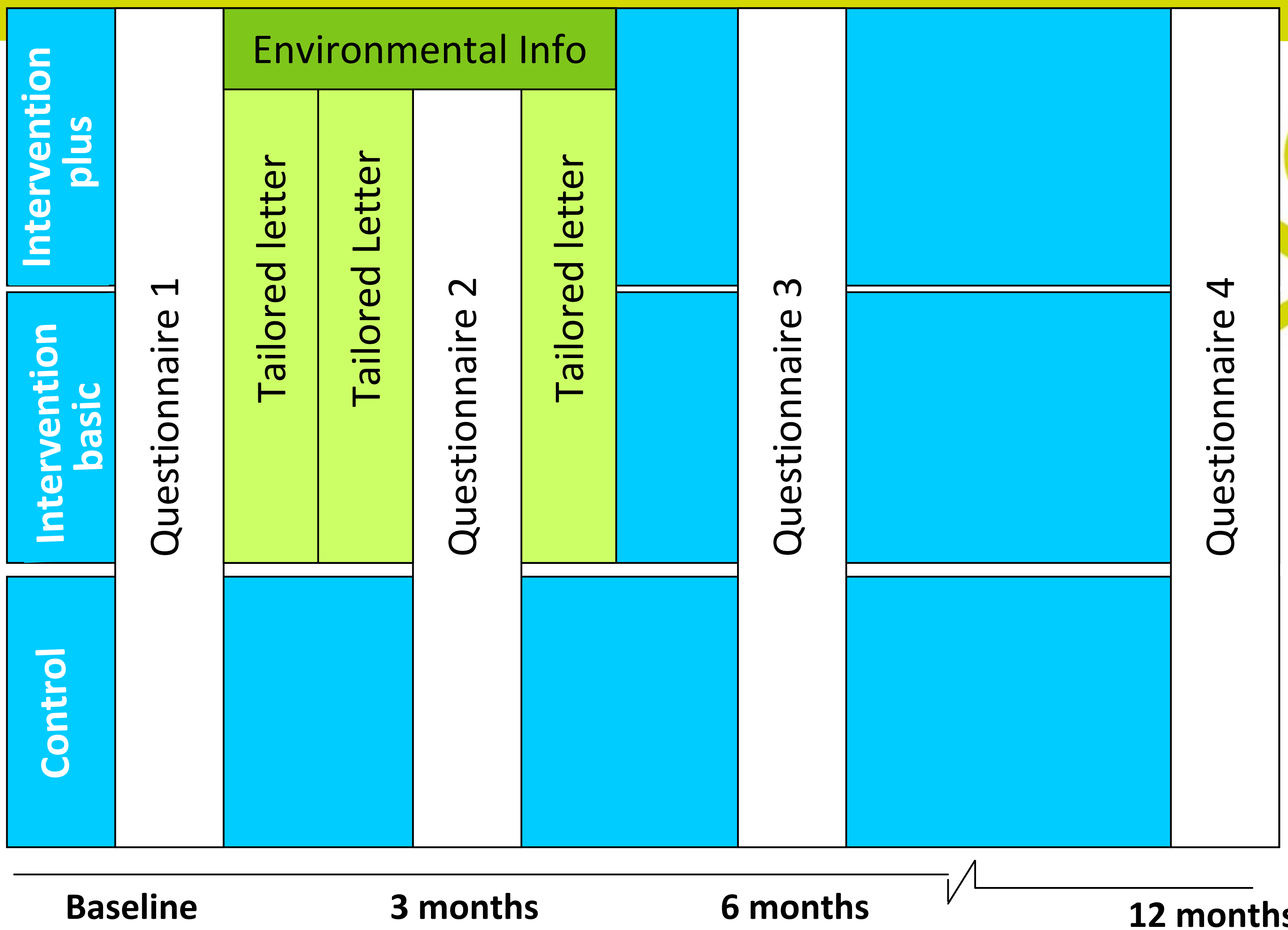
- In most situations, the relation between X and Y is reduced when the third-variable is included because it is a mediator or a confounder. There are cases where the X to Y relation gets bigger or reverses sign when a third variable is included.
- A suppressor variable is a variable that increases the magnitude or changes the sign of the relation between X and Y when it is included in the analysis.



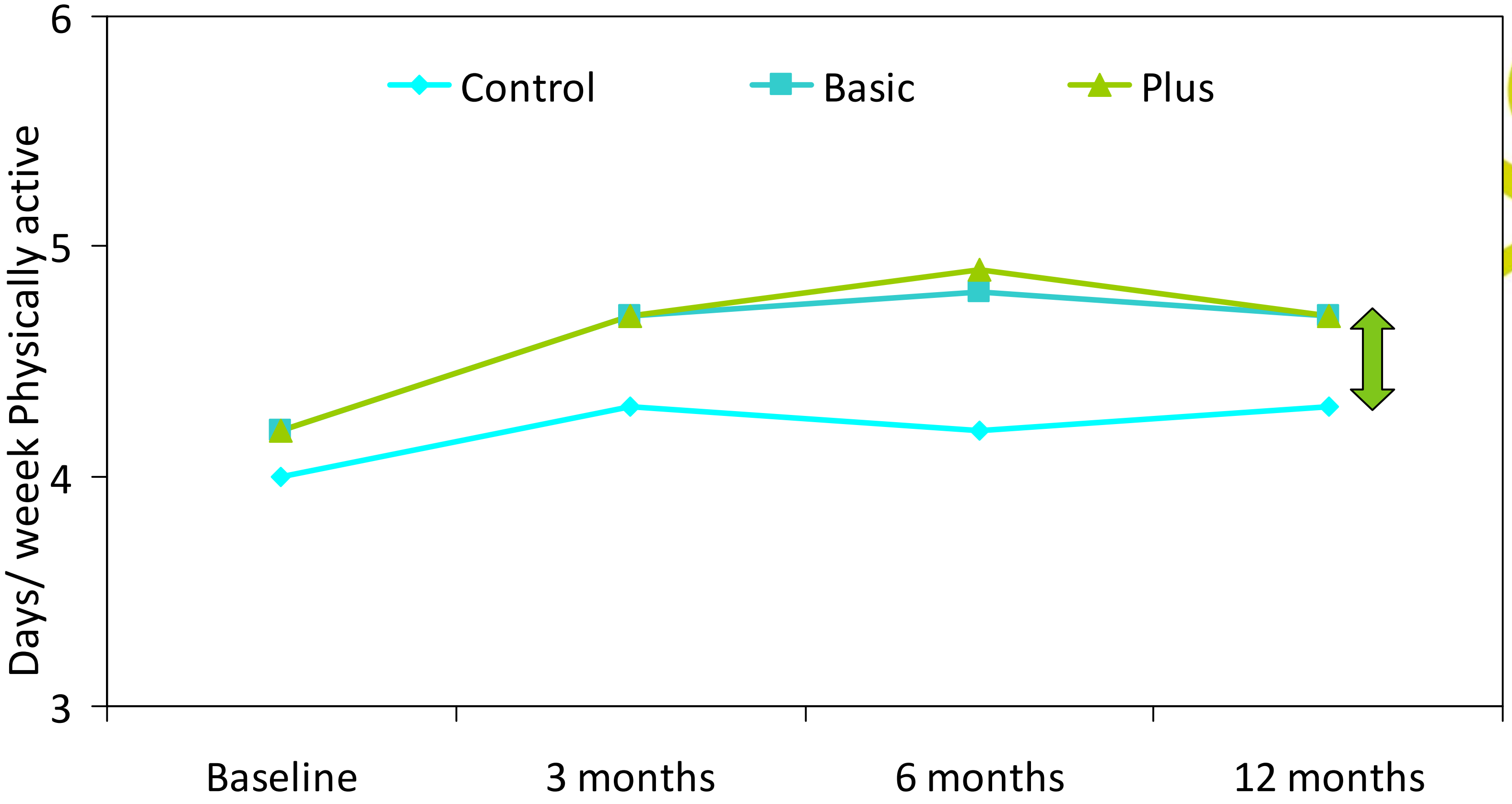
EXAMPLE: the Active *plus* project

Example: Active *plus* intervention

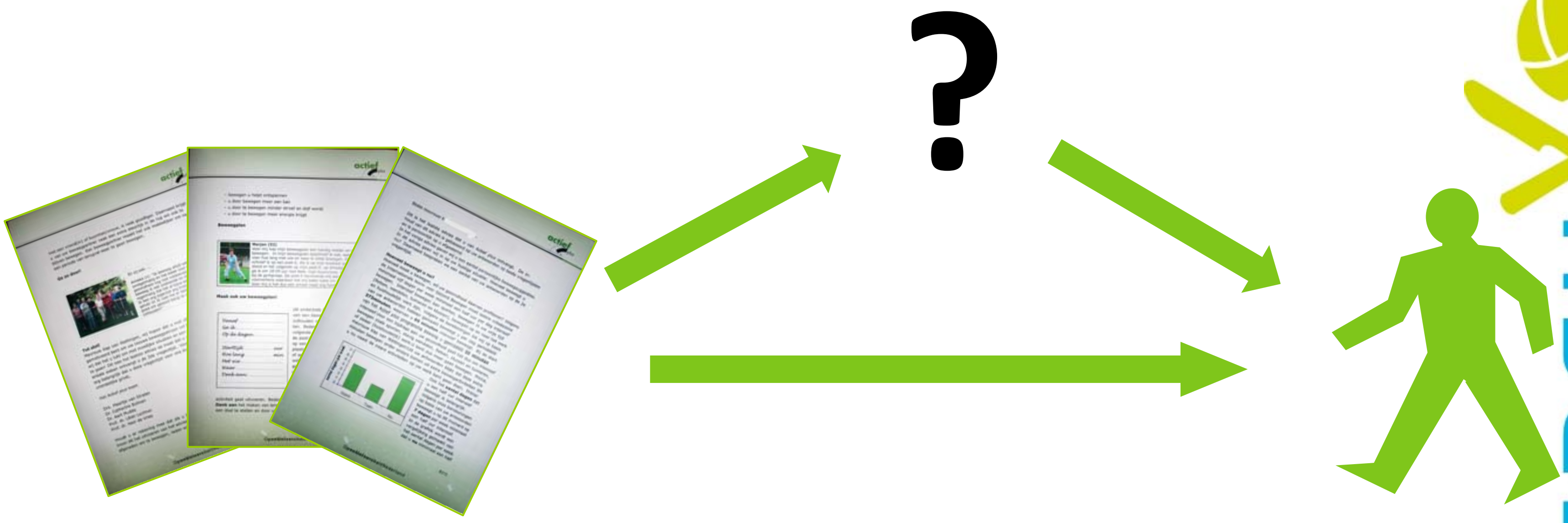
Design Active plus intervention



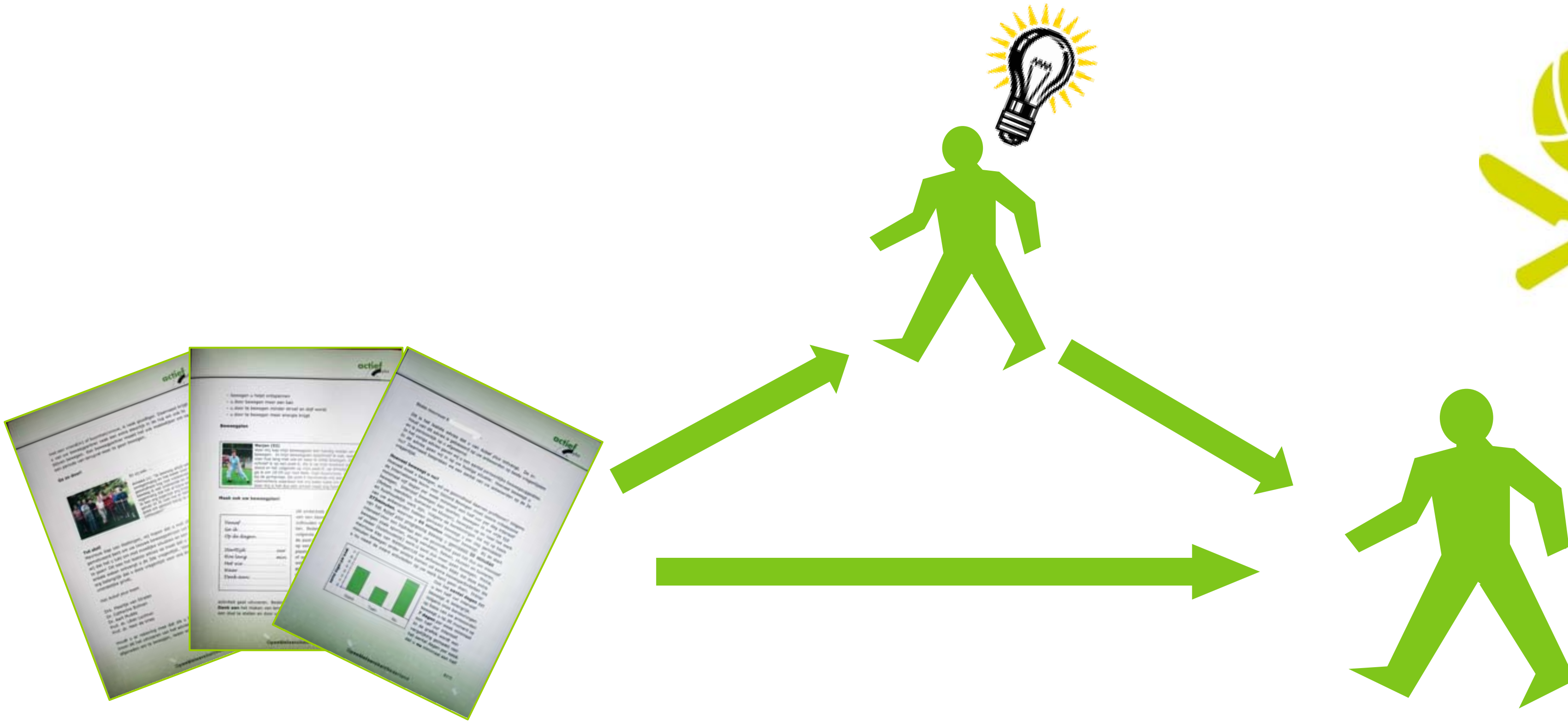
Efficacy Active plus intervention



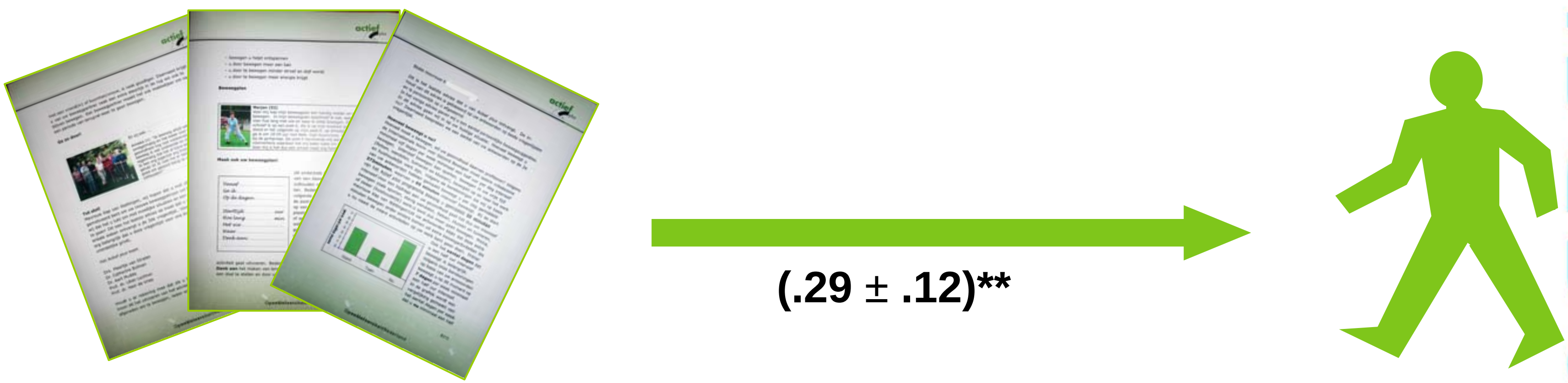
Working mechanisms intervention



Working mechanisms intervention



Step 1: Intervention effect

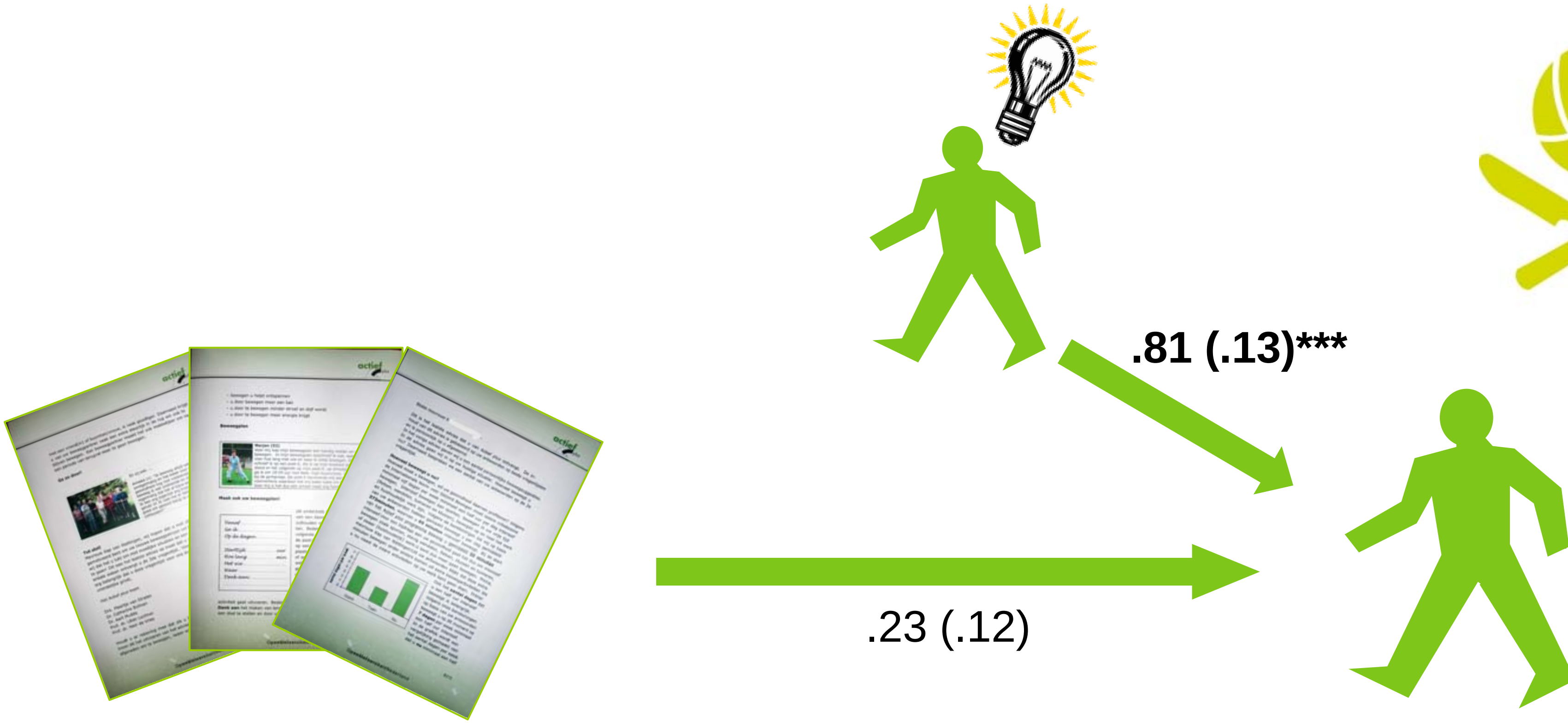


** p< 0.01

Step 2: Action theory test

** p<0.01

Step 3: Conceptual theory test



*** p<0.001

Step 4: Test of mediation

- Product of coefficient test (a*b)
  - a= .07 ± .03*
  - b= .81 ± .13***
  - Mediated effect= a*b= 0.08 *.82= 0.06

- Difference in coefficient test (c-c′)
  - c= 0.29
  - c′= 0.23
  - c-c′= 0.29- 0.23= 0.06

Step 5: test of significance + CI

- Test of significance

$$s_{First} = \sqrt{\hat{a}^2 s_{\hat{b}}^2 + \hat{b}^2 s_{\hat{a}}^2}$$

- $s_{ab}= \sqrt{(0.07^2*0.13^2+ 0.81^2*0.03^2)}$
- $s_{ab}= 0.03$
- Wald statistics: $ab/ s_{ab}= 0.06/ 0.03= 2.14$= significant

- Confidence intervals
  - Upper Confidence Interval (UCL) = $ab + Z_{.975}\, s_{ab} = 0.07 + 1.96 * 0.03= 0.114$
  - Lower Confidence Interval (LCL) = $ab + Z_{.025}\, s_{ab} = 0.07 – 1.96 * 0.03= 0.005$
  - 0 is not in the interval: mediated effect is significant

Step 6: proportion mediated

- Total effect= Indirect effect (*ab*) + Direct effect (c′)
- Proportion mediated= (Indirect effect/ Total effect) * 100%
- Proportion mediated= (*ab*)/ (*ab + c′*)* 100%

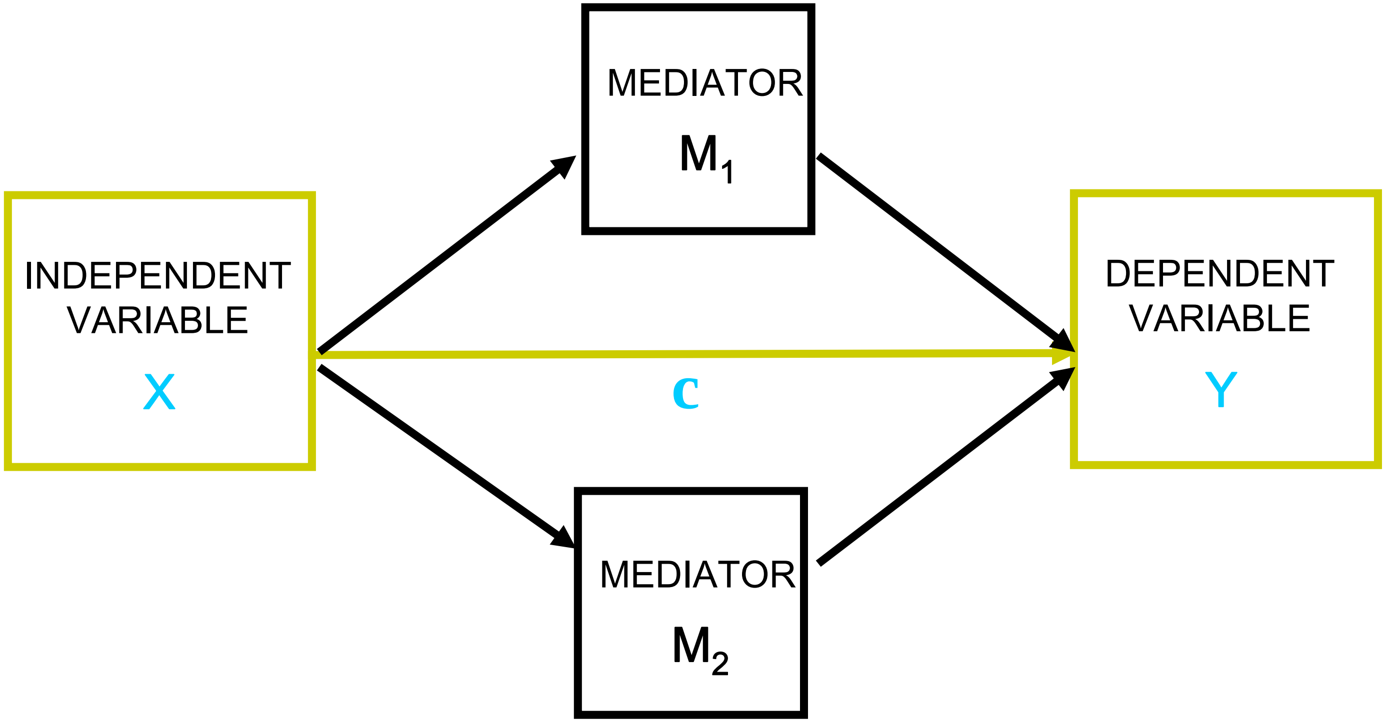
- Total effect= 0.06 + 0.23= 0.29
- Proportion mediated= (0.06/ 0.29) * 100%= 20.8%

- 20.8% of the intervention effect on PA behavior could be explained by changes in awareness induced by the intervention

1.2 Multiple mediation analyses

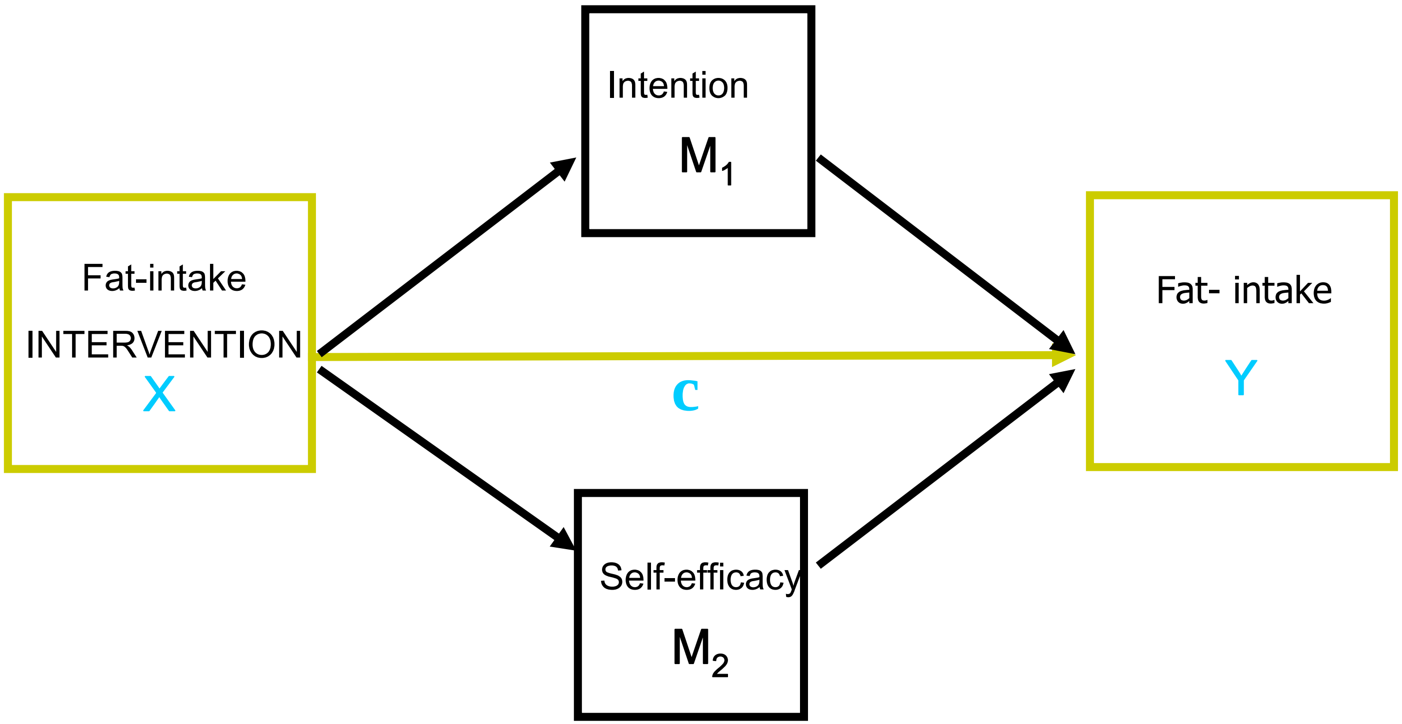
Multiple mediation

- Most behaviors are affected by multiple variables to it makes sense that there are multiple mediators.



Multiple mediation

- Most behaviors are affected by multiple variables to it makes sense that there are multiple mediators.



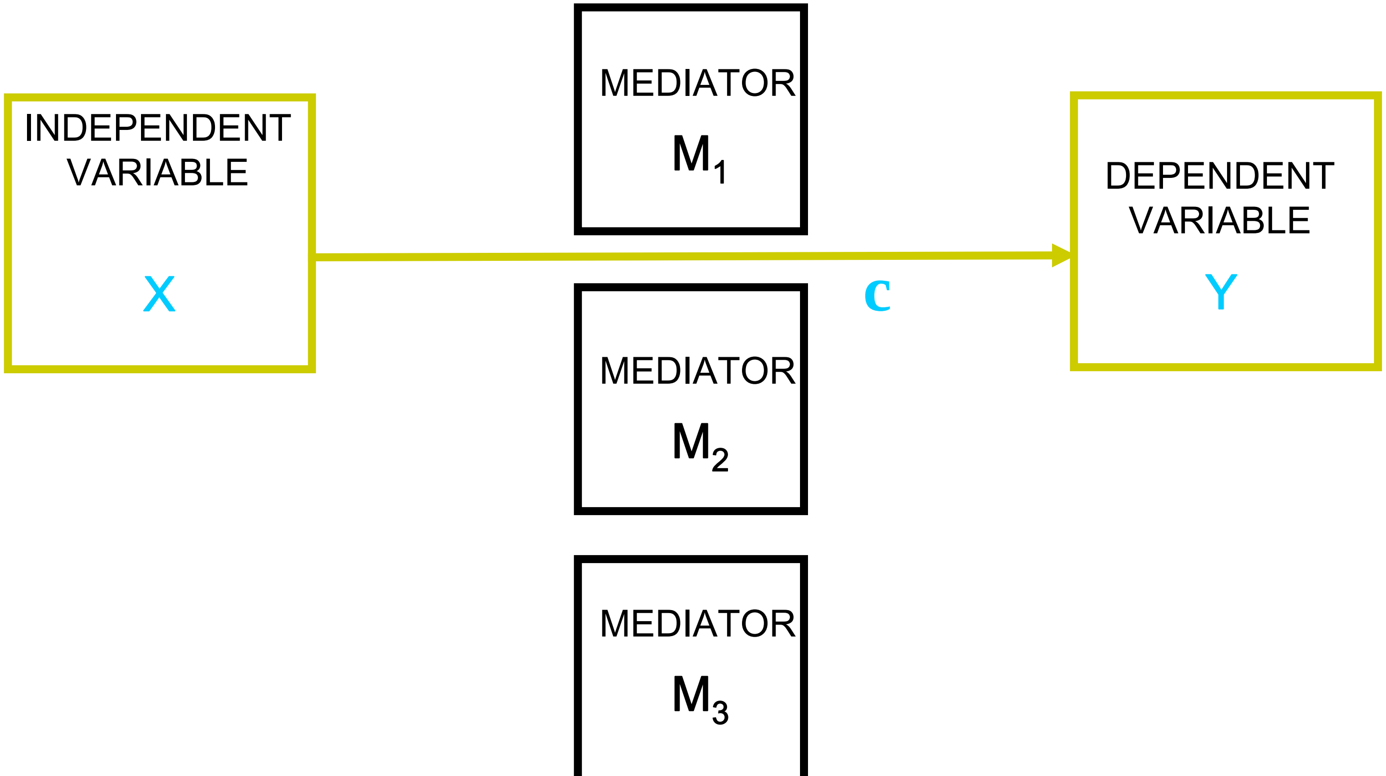
Multiple mediation

- Straightforward extension of the single mediator case but interpretation can be more difficult especially when considering all possible relations among variables.

- The product of coefficients methods is the best way to evaluate models with multiple mediators.

- Four steps:
  - Step 1: Main effect
  - Step 2: Action theory test
  - Step 3: Conceptual theory test
  - Step 4: Mediation effects

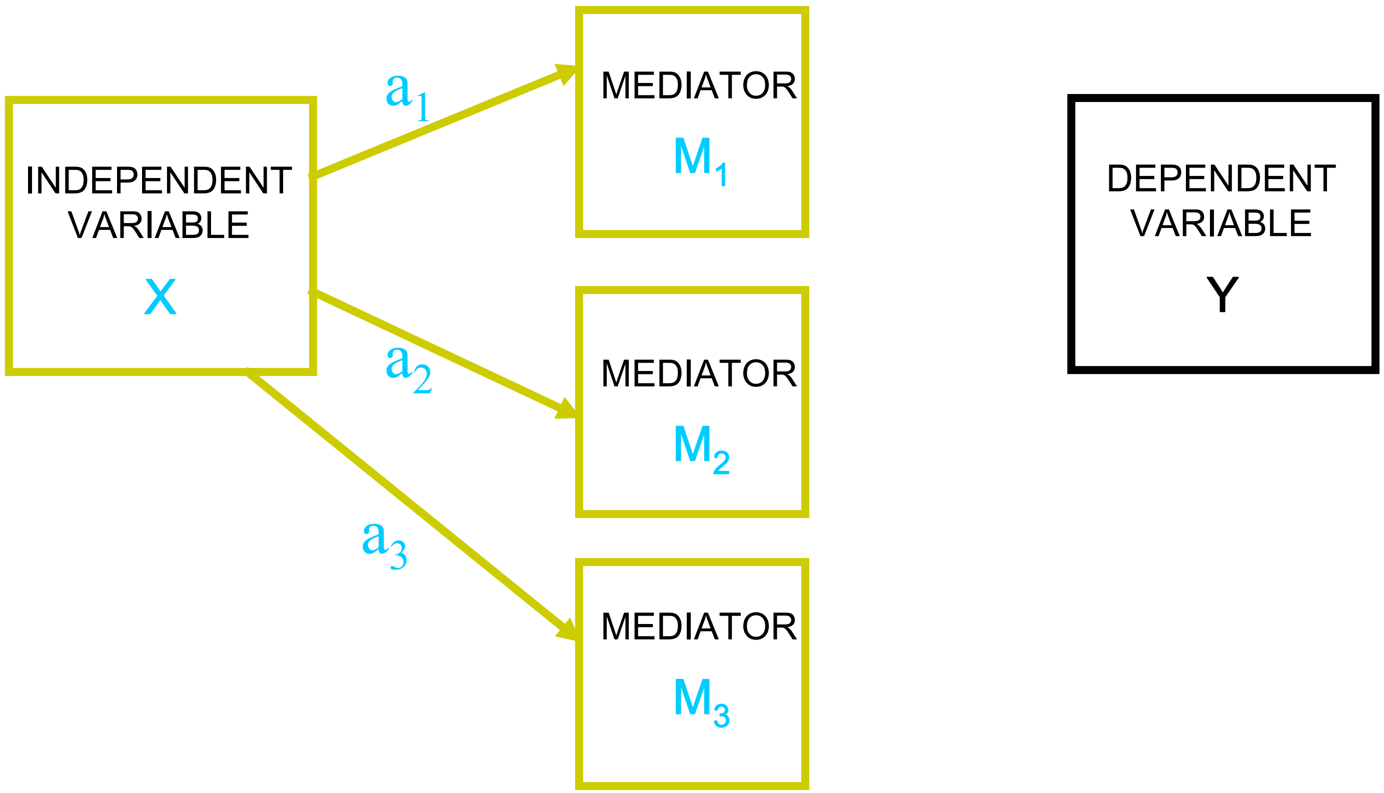
Step 1: Main effect



1. The independent variable causes the dependent variable:

$$\hat{Y} = \hat{c}X + \varepsilon_1$$

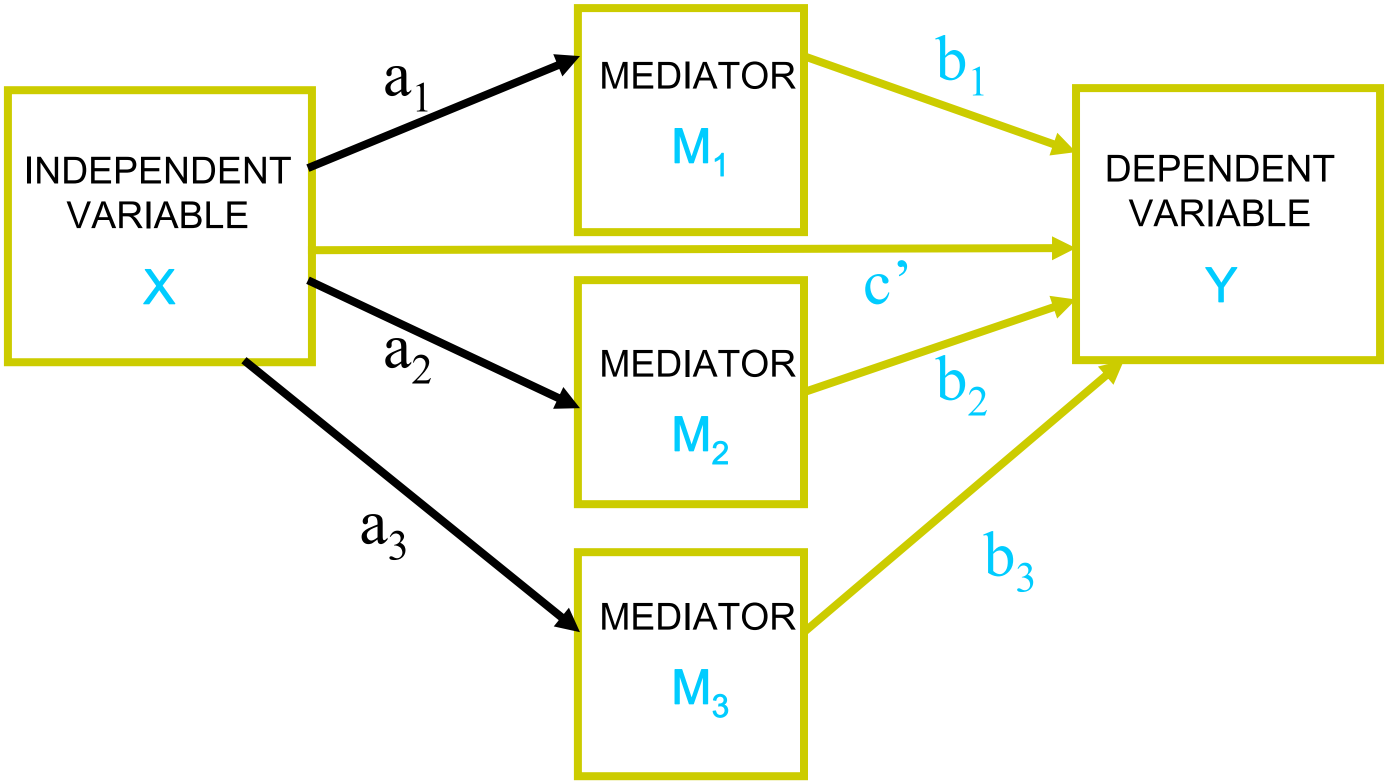
Step 2: Action theory test



2. The independent variable causes the potential mediators:

$$\hat{M}_1 = \hat{a}_1 X + \varepsilon_2,\ \hat{M}_2 = \hat{a}_2 X + \varepsilon_3,\ \hat{M}_3 = \hat{a}_3 X + \varepsilon_4$$

Step 3: Conceptual theory test



3. The mediators cause the dependent variable controlling for exposure to the independent variable:

$$\hat{Y} = \hat{c}'X + \hat{b}_1 M_1 + \hat{b}_2 M_2 + \hat{b}_3 M_3 + \varepsilon_6$$

Step 4: Mediation Effects

Mediated effects **=** a_1b_1, a_2b_2, a_3b_3

Standard error **=** $\sqrt{(s^2_{a1}*b_1^2 + s^2_{b1}*a_1^2 + s^2_{a2}*b_2^2 + s^2_{b2}*a_2^2 + s^2_{a3}*b_3^2 + s^2_{b3}*a_3^2 + \; + 2*a_1*a_2*COVb_1b_2 + 2*a_1*a_3*COVb_1b_3 + 2*a_2*a_3*COVb_2b_3\ )}$

COV= covariance between the b_i regression estimates

Total mediated effect **=** a_1b_1 **+** a_2b_2 **+** a_3b_3 **=** c - c′

Direct effect **=** c′ Total effect **=** a_1b_1 **+** a_2b_2 **+** a_3b_3 **+** c′ **=** c

Test for significant mediation= $z' = \Sigma a_ib_i / se_{aibi}$

Proportion Mediated = $a_ib_i/(c' + a_ib_i) = a_ib_i/c$

To summarize…

- A mediating variable
  - is the intermediate in the causal sequence relating the independent to the dependent variable
- Mediation analyses
  - Provides an explicit check on an intervention's theoretical underpinnings and whether the proposed process was achieved
  - Can make interventions more cost-effective
  - Can test theories and determine the associations between theoretical constructs
  - Consists basically of 4 steps
    1. Test of main effect
    2. Action theory test
    3. Conceptual theory test
    4. Test of mediated effect
  - Test significance by Sobel test or Bootstrapping
- Multiple mediation is a straightforward extension of the single mediator

Moderation Analysis

Yıldırım M., Van Stralen M.M. and Te Velde S.J.

20-05-2010

Outline of the presentation

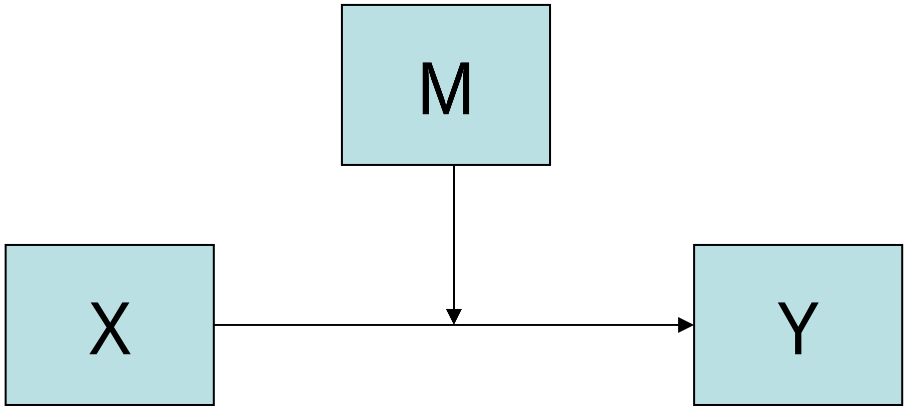
- Introduction to moderating effect
- How to do the analysis
- Plotting the interaction
- Problems with moderation analysis
- Resources

Introduction

- Interventions may not be equally effective across subgroups
- To plan intervention strategies that deals with the needs of subgroups
- To explore → Moderation analysis

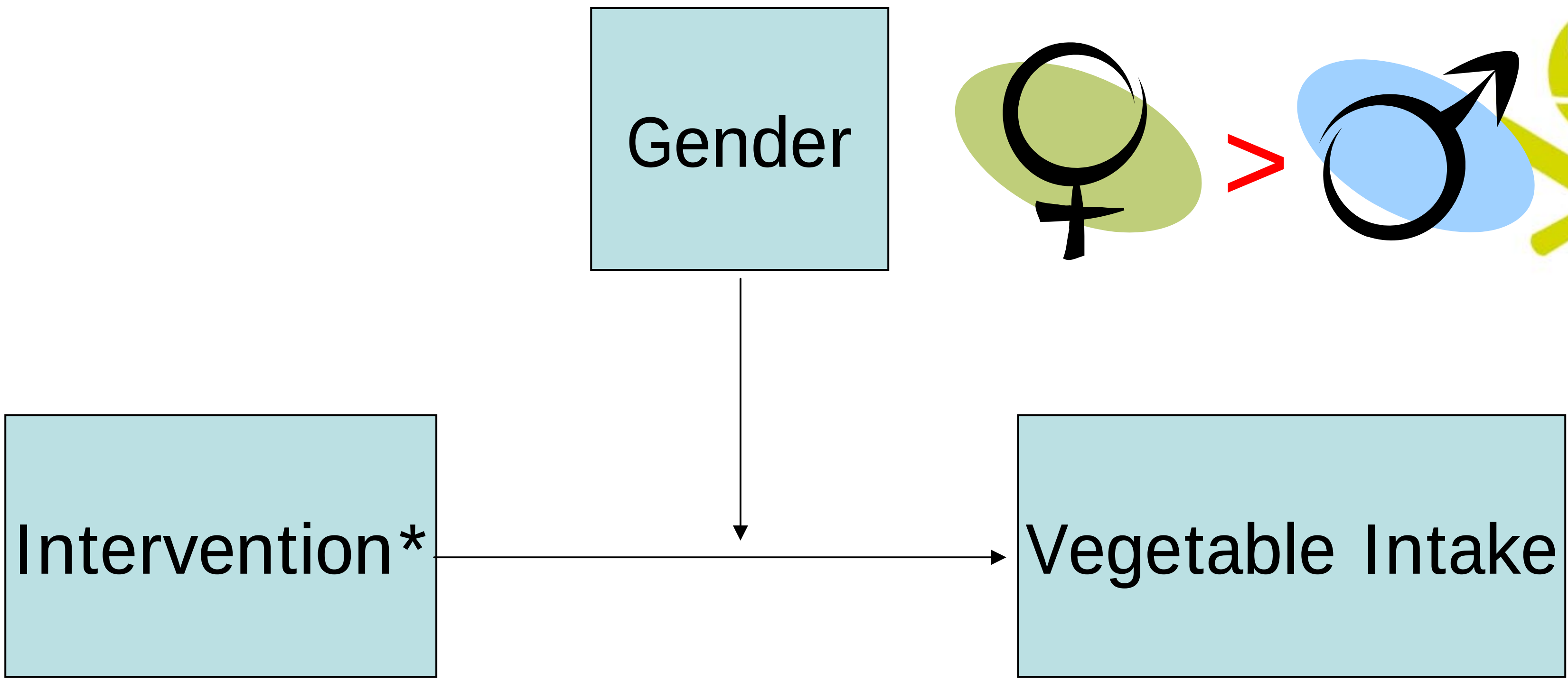
Moderators

Moderators (effect modifiers) are the variables that affect the direction and/or strenght of the relationship between variables



M as **MODERATOR**

Moderating Effect



* Tak N., Te Velde S. and Brug J. Pub Health Nutr 2007; 10: 1497-1507.

Moderation analysis

- Most common; Moderated multiple regression

$$\hat{Y} = b_0 + b_1 X$$

Y: Independent variable

X: Dependent variable

Moderation analysis

- Including an interaction term

is created by multiplying the moderator and independent variable

$$\hat{Y} = b_0 + b_1 X + b_2 M + b_3 X \cdot M$$

M: Moderator

Coding of categorical variable

- Dummy coding;

  - Most frequently used
  - Easy interpretation
  - It uses only ones and zeros
  - E.g. Gender; female-0, male-1

Test of significance

- t-statistics, H_0: $b_3 = 0$

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	9.777	1.232		7.934	.000
	group	.893	.367	.056	2.435	.015
	gender.dummy	-1.682	.395	-.108	-4.262	.000
	waist circum T0	.928	.018	.860	52.436	.000
	gen.dummy.mod	-1.440	.515	-.085	-2.799	.005

a. Dependent Variable: waist circum T3

- F-test; H_0: R^2 s should be equal *

** Aguinis H, Gottfredson RK. J Organiz Beh (2010)*

Moderation Analysis

Effect Sizes

- $R_2^2 - R_1^2$

$$\hat{Y} = b_0 + b_1 X + b_2 M \qquad R_1^2$$

$$\hat{Y} = b_0 + b_1 X + b_2 M + b_3 X \cdot M \qquad R_2^2$$

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.232[a]	.054	.029	1.348
2	.331[b]	.110	.074	1.317

a. Predictors: (Constant), agec, heightc

b. Predictors: (Constant), agec, heightc, heightc.agec

=0.110-0.054

=0.056 → 5.6%

Moderation Analysis

Effect Sizes

- In case of violation of homogeneity of error variances

$$f^2 = \frac{R_2^2 - R_1^2}{1 - R_2^2}$$

- Cohen (1988)

  f^2 = .02: small effect

  f^2 = .15: medium effect

  f^2 = .26: large effect

What next after the significant interaction?

- Stratified analyses

Table 3 Baseline and follow-up median (25th and 75th percentiles) and mean scores on the outcome measures for the multicomponent programme. Regression coefficients (β) from the multilevel regression analyses with intake at follow-up as dependent variable and group as independent variable

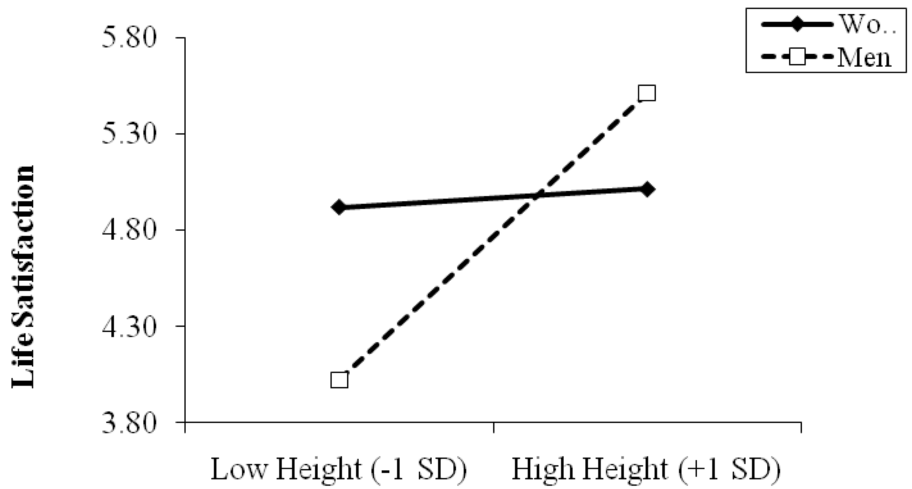
		Baseline median (P25, P75); mean (SD)		Follow-up median (P25, P75); mean (SD)			
		MC	C	MC	C	β	Net effect
Fruit (portions per day)		n = 209; 1.1 (0.7, 1.7); 1.3 (0.8)	n = 439; 1.1 (0.7, 1.7); 1.2 (0.7)	n = 206; 1.4 (0.9, 2.1); 1.5 (0.8)	n = 439; 1.1 (0.7, 1.7); 1.2 (0.7)	0.11***	0.2
Vegetables grams per day		n = 196; 53 (38,68); 58 (32)	n = 421; 58 (41, 75); 61 (29)	n = 190; 52 (35, 74); 56 (30)	n = 406; 55 (40, 74); 60 (29)	−0.03	−1
Vegetable snack (times per day)	Age group I	n = 49; 0.1 (0.1,0.4); 0.2 (0.2)	n = 76; 0.3 (0.0, 0.4); 0.3 (0.3)	n = 53; 0.1 (0.1, 0.6); 0.3 (0.3)	n = 77; 0.3 (0.1, 0.6); 0.3 (0.3)	−0.01	0.1
	Age group II	n = 87; 0.1 (0.0, 0.3); 0.2 (0.2)	n = 164; 0.3 (0.0, 0.4); 0.3 (0.2)	n = 84; 0.3 (0.1,0.4); 0.3 (0.3)	n = 162; 0.3 (0.1, 0.4); 0.3 (0.3)	0.06*	0.1
	Age group III	n = 66; 0.1 (0.0,0.3); 0.2 (0.2)	n = 197; 0.1 (0.0, 0.4); 0.2 (0.2)	n = 64; 0.2 (0.0, 0.4); 0.3 (0.3)	n = 192; 0.1 (0.1, 0.4); 0.2 (0.2)	0.03	0.1
	Boys	n = 90; 0.1 (0.0, 0.3); 0.2 (0.2)	n = 208; 0.1 (0.0, 0.4); 0.2 (0.3)	n = 91; 0.1 (0.0, 0.3); 0.2 (0.3)	n = 207; 0.3 (0.0, 0.4); 0.3 (0.2)	0.00	−0.1
	Girls	n = 112; 0.1 (0.1, 0.3); 0.2 (0.2)	n = 226; 0.3 (0.1, 0.4); 0.3 (0.2)	n = 110; 0.3 (0.1, 0.6); 0.3 (0.3)	n = 221; 0.3 (0.1, 0.4); 0.3 (0.3)	0.07**	0.1
24 h FJV (times per day)	Native	n = 168; 2.0 (2.0,3.0); 2.5 (1.2)	n = 332; 2.0 (2.0, 3.0); 2.4 (1.1)	n = 160; 3.0 (2.0, 4.0); 3.0 (1.2)	n = 323; 3.0 (2.0, 3.0); 2.7 (1.2)	0.24*	0.2
	Non-native	n = 21; 3.0 (2.0, 4.0); 2.7 (1.3)	n = 43; 4.0 (3.0, 5.0); 3.8 (1.5)	n = 19; 4.0 (3.0, 5.0); 3.8 (1.4)	n = 41; 3.0 (2.0, 5.0); 3.3 (1.8)	1.30**	1.6

P25 = 25th percentile; P75 = 75th percentile; SD = standard deviation; MC = multicomponent programme; C = control group; FJV = fruit, vegetables and fruit juice.
* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

Reinaerts E., et al. Pub Health Nutr 2007; 10 (9):939-947.

Plotting the interaction

- Available programs for graphical display of moderation analysis (ModGraph, Preacher)
- (Continous IV)
- -1SD, mean, +1SD



Moderation analysis

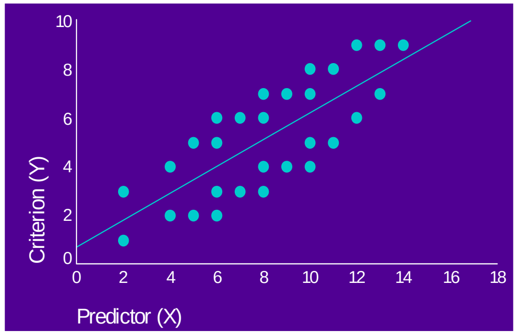
Several conditions are required, e.g.;

- Rationale (theory or evidence base)
- Homogeneity of residual variances (Applies to categorical moderators)

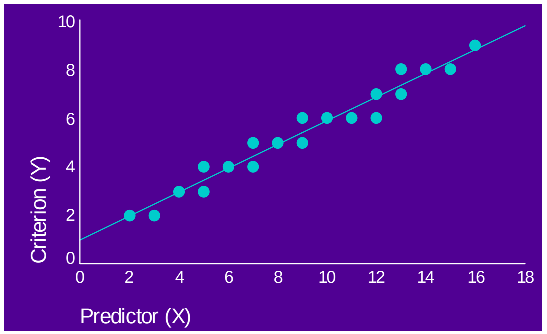
Checklist *Frazier et al. J Couns Psychol 51: 115-134, 2004.*

Homogeneity of residual variances

- Visually



Women



Men

Homogeneity of residual variances

- Regression analysis- Aguinis program:ALTMMR

ALTMMR Results

Sub-Group	Sub-Group Size(n)	Predictor Std.deviation (S_x)	Criterion Std.deviation (S_y)	Correlation for X and Y(r_{xy})
1	411	0.49	7.17	-0.053
2	427	0.49	7.27	-0.153

Error Variance Results

DeShon & Alexander's rule of thumb for homogenity is met. (The Error Variance ratio is 1:1.01)

Bartlett's Test indicates homogeneous error variance (M = 0.0049, p = 0.9444).

Alternative Differential Slopes Statistics

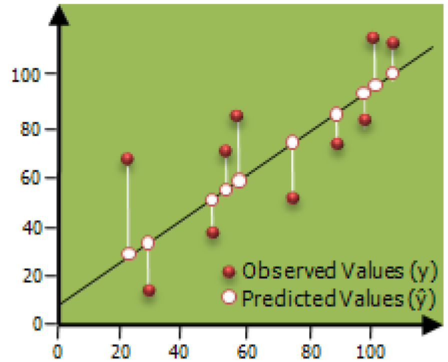
James's Test indicates no differential slopes. (p > 0.05) U = 2.173, and U(critical) = 3.8521

Alexander's Test indicates no evidence of differential slopes! (A = 2.1675, p = 0.141)

Violation of hom.

- Weighted Least Square

Instead of OLS



- Structural Equation Modelling

Moderation analysis

- Non-linear relationships
  - Cubic (adding third root)
  - Quadratic (adding square root)
  - E.g. quadratic M relationship and its interaction;

$$\hat{Y} = b_0 + b_1 X + b_2 M + b_3 M \cdot M + b_4 X \cdot M$$
$$+ b_5 X \cdot M \cdot M$$

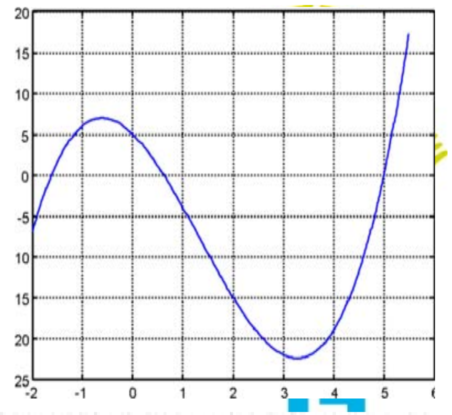
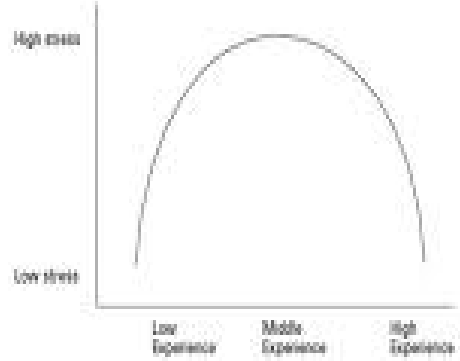


Figure 2 Schematic representation of quadratic relationship between stress and experience extrapolated from data in Table I



Moderation analysis

- Higher-order interactions

- Extending the moderation >2 predictors

$$\hat{Y} = b_0 + b_1 X + b_2 M + b_3 Z + b_4 X \cdot M + b_5 X \cdot Z + b_6 M \cdot Z + b_7 X \cdot M \cdot Z$$

Moderation analysis

- Centering-Mean centering

  Observed value – Sample mean value
- Not a requirement
- Only for interpretation of the regression coefficients

Problems with moderation analysis

- Low statistical power
  - Sample size
  - The equality if the subgroups sizes
  - Measurement errors in the variables
  - Categorization of continous variables
  - Small effect sizes

Recommendations

1) A large enough sample size (determined a priori by a power calculation), balanced subgroups, reliable measures and continuous scales (rather than artificially dichotomized)
2) Reporting moderation analysis in the abstract
3) Restricting the investigations to the specific rationale
4) In case of violation of homogeneity of (error) variances assumption, use alternative tests

Resources

- Checking the homogeneity of error variances;
- Aguinis's program (ALTMMR): http://mypage.iu.edu/~haguinis/mmr/index.html

- Power calculation programs;
- Gpower
- Aguinis's program: http://mypage.iu.edu/~haguinis/mmr/index.html

- Plotting the interaction;
- ModGraph: http://www.victoria.ac.nz/psyc/paul-jose-files/modgraph/modgraph.php
- Preacher: http://www.people.ku.edu/~preacher/interact/index.html

THANK YOU!

Mine Yildirim

m.yildirim@vumc.nl

Lecture 3: Advanced mediation and moderation analyses

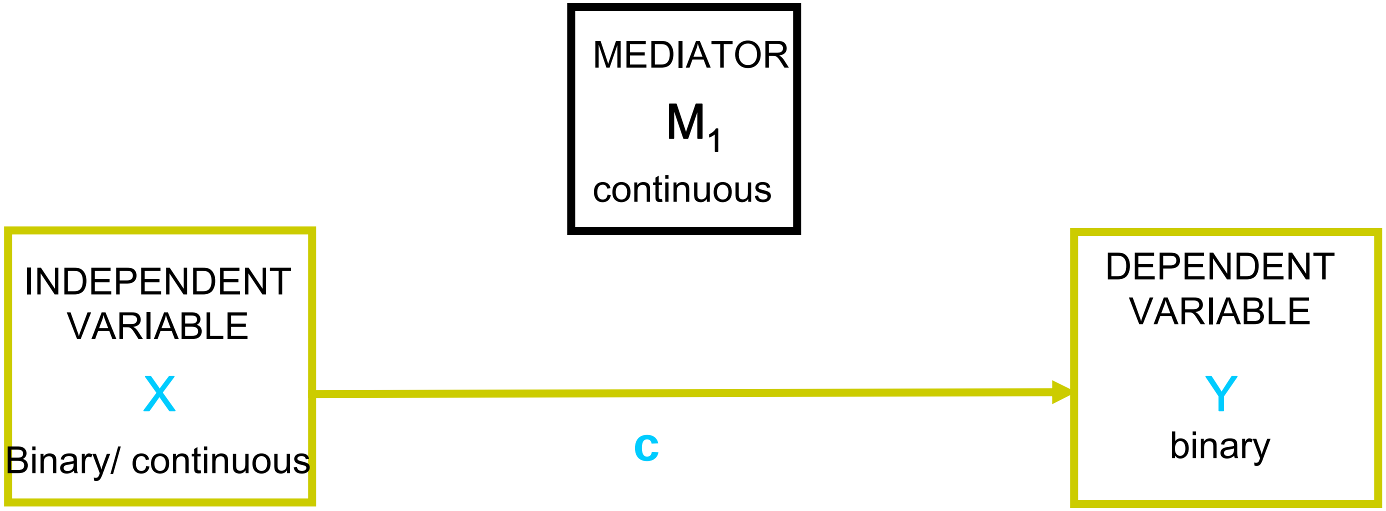
Today: afternoon session

- 13:30-15:15 Lecture 3: Advanced mediation and moderation
  - Mediation with a categorical outcome variable
  - Combining moderation and mediation analyses
  - Longitudinal mediation analyses
- 15:15-15:30 Coffee
- 15:30-17:00 Practicum 3:
  - Practice with your own data, OR
  - Extra exercises:
    - Exercise 3.1: longitudinal mediation analyses
    - Exercise 3.2: moderated mediation analyses
    - Exercise 3.3: mediation analyses with a categorical outcome
- 17:00- ??:?? Drinks

3.1. Categorical outcome variables

Categorical outcome variable

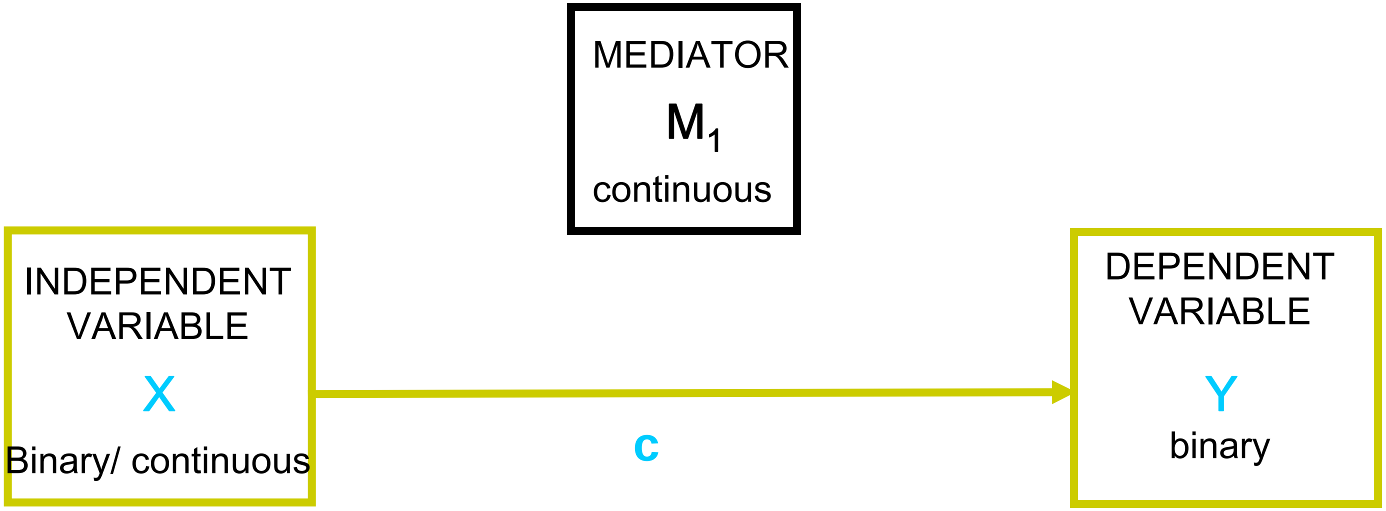
- A dependent variable is often binary such as whether a person complies with guidelines, used a condom or not, is dead or alive.
- In this case, logistic or probit regression is the method of choice.
- Estimates of the mediated effect using logistic and probit regression can be distorted using conventional procedures.



Step 1: main effect

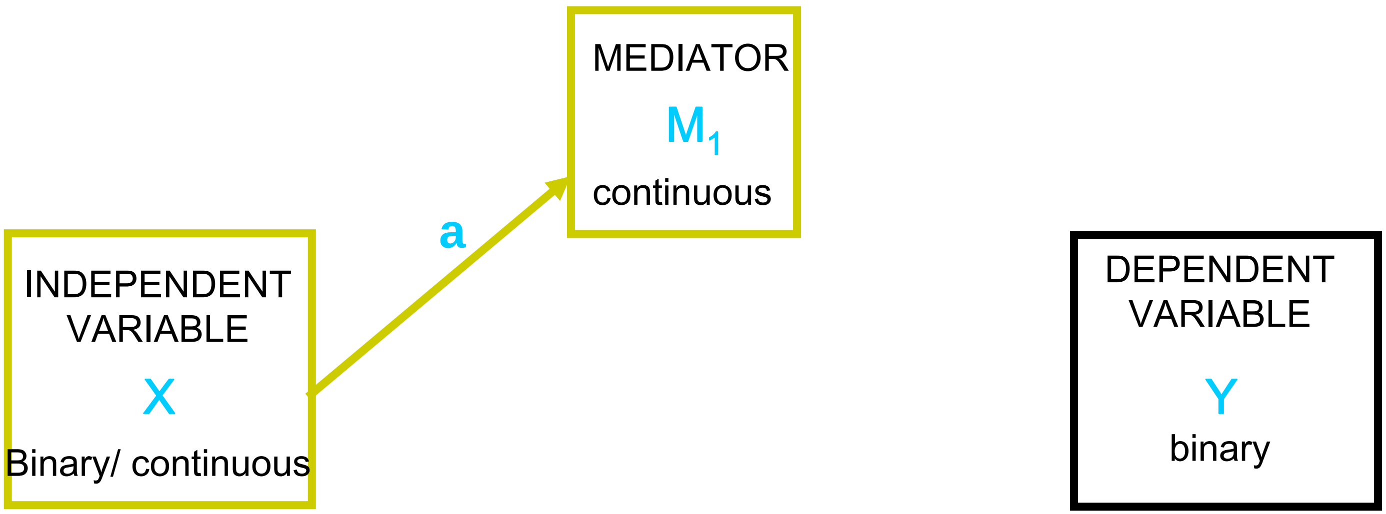
- Standard logistic regression model, where Y depends on X, i_1 is the intercept and c codes the relation between X and Y.
- $\text{logit } \Pr\{Y=1|X\} = i_1 + cX$

c is the regression coefficient *not* the OR



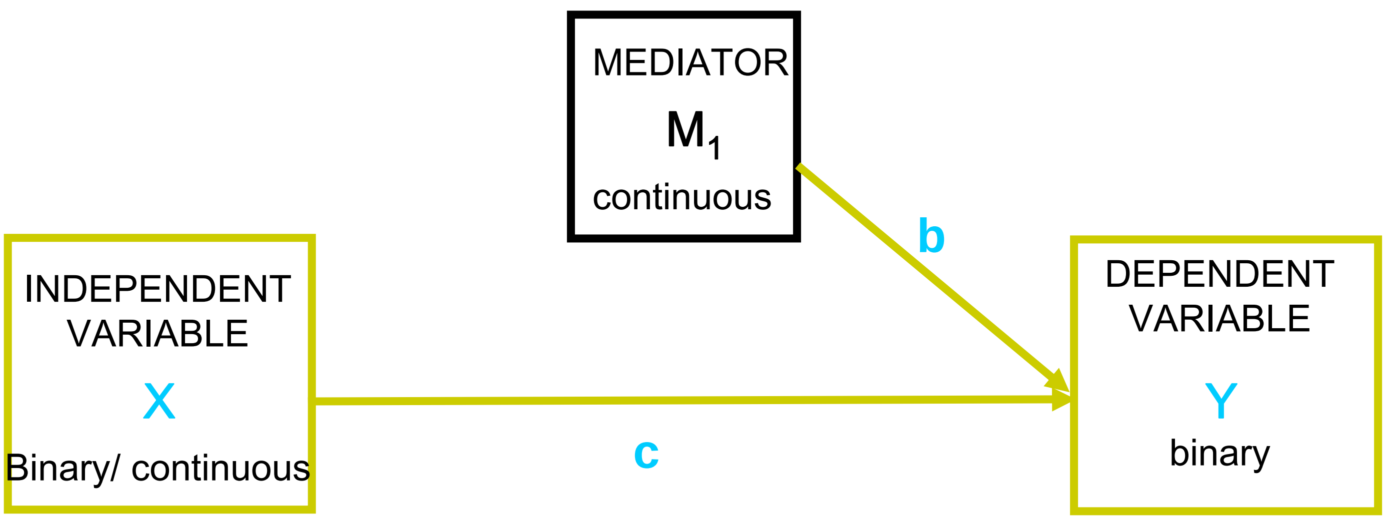
Step 2: Action theory test

- M is a continuous variable so ordinary least squares regression is used to estimate this model where *i* is the intercept, *a* represents the relation between X and Y, and ε is residual variability.
- $M = i + aX + \varepsilon$



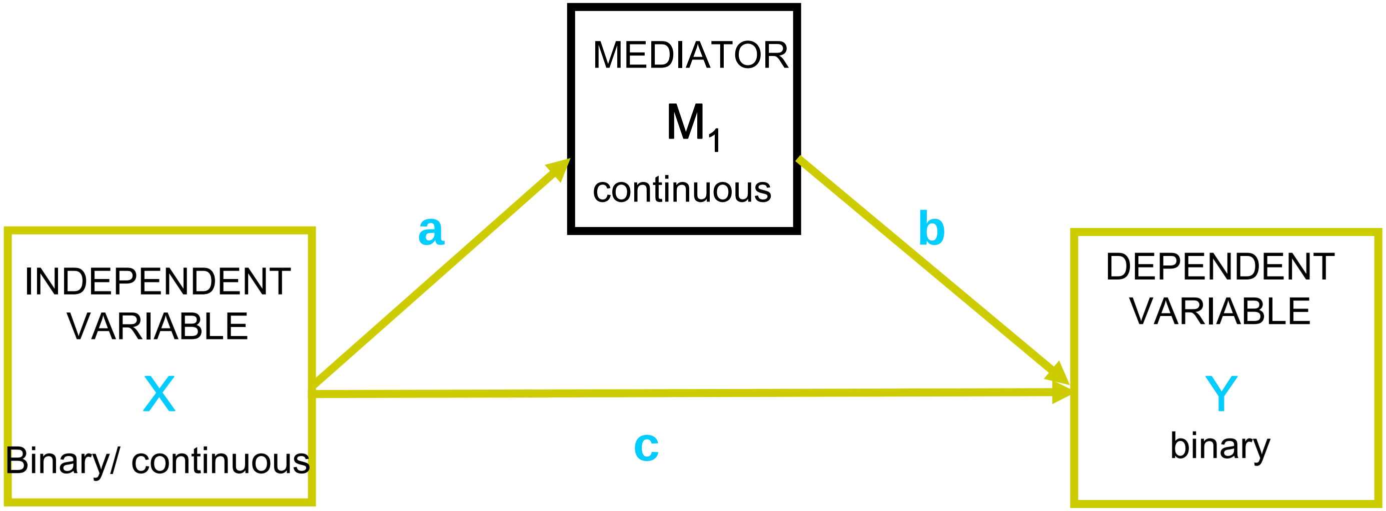
Step 3: Conceptual theory

- Standard logistic regression model, where Y depends on X and M, i_2 is the intercept, c' codes the relation between X and Y adjusted for M and b codes the relation between M and Y, adjusted for X.
- logit Pr{Y=1|X,M} = $i_2 + c'X + bM$ (2)



Step 4: Mediated effect

- *c* - *c′* Difference in coefficients. The coefficients are from separate logistic regression equations.
- *ab* Product of coefficients. The *b* coefficient is from a logistic regression model and *a* is from an ordinary least squares regression model.
- The difference in coefficients method can give distorted values for the mediated effect because of differences in the scale of separate logistic regression models.



Step 5: significance testing

- The standard error of a*b can be calculated using the earlier standard error equations

$$s_{First} = \sqrt{\hat{a}^2 s_{\hat{b}}^2 + \hat{b}^2 s_{\hat{a}}^2}$$

- Standard error of c-c′ is more complicated

- The two estimators, *ab* and *c-c′* are not identical in logistic or probit regression.
- There are solutions:
  - Standardize the values of the coefficients c and c′
  - Use a computer program such as Mplus that appropriately handles categorical variables in covariance structure models.

Multiple mediator models

- Multiple mediator models for continuous variables outlined in the morning, can also be evaluated when the dependent variables is categorical, using logistic regression analyses
  - $Y^* = i + cX + e$
  - $Y^* = i + c'X + b_1M_1 + b_2M_2 + b_3M_3 + e$
  - $M_1 = i + a_1X + e$
  - $M_2 = i + a_2X + e$
  - $M_3 = i + a_3X + e$
  - Total mediated effect= $a_1b_1 + a_2b_2 + a_3b_3 \neq c - c'$
- Can be conducted in SPSS, SAS and Mplus
- However, with more than two categorical dependent variables, or multiple mediator models with combinations of categorical and continuous variables, more complicated iterative approaches such as Mplus covariance structure analyses are needed.

3.2. Combining mediation and moderation

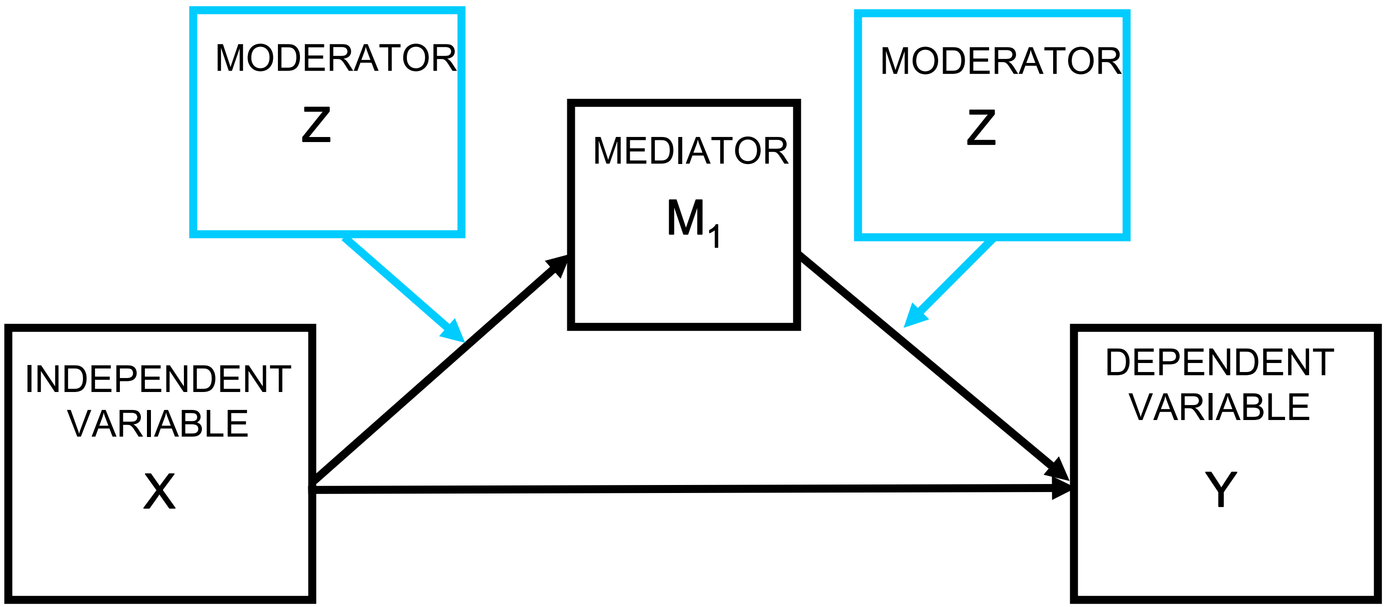
Moderation of a mediated effects

- Both moderated and mediated effects are important in research
  - Understand how manipulations achieve effects **and** identify characteristics of participants and/or environment that moderate effectiveness of a manipulation.
- Improve intervention strategies by understanding for whom and/or under what conditions they operate
- Better target subgroups by understanding how they differentially respond to intervention strategies
  - Does a program differentially impact participants based on level of risk

- Mediation of a moderator effect
  - Exploring mediating mechanisms to explain an overall interaction effect
- Moderation of a mediated effect
  - Investigating whether a mediated relation holds across levels of a fourth moderating variable

Moderation of a mediated effect

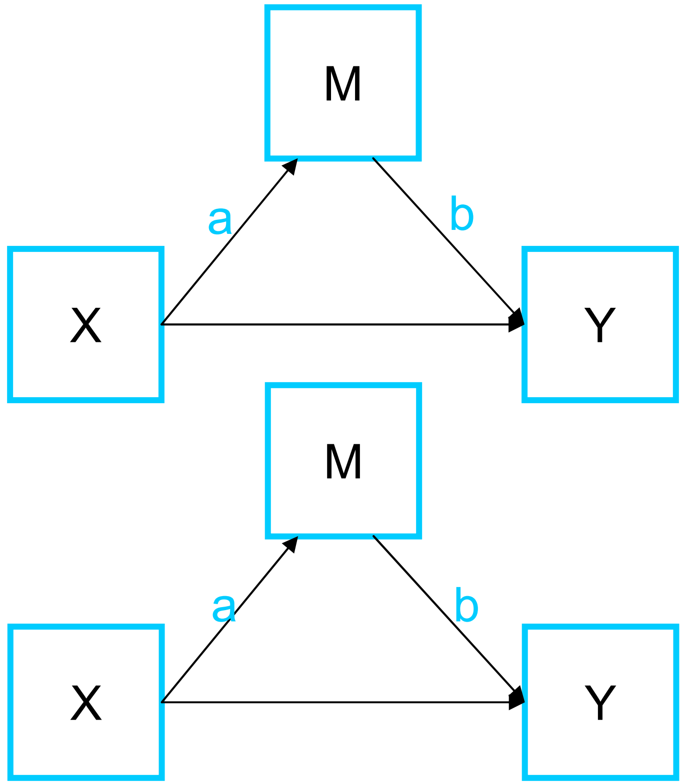
- Mediated relation depends on the level of a moderator variable
- Criteria for the moderation of a mediated effect:
  - a path is moderated
  - b path is moderated
  - or both a and b paths are moderated.



Dichotomous moderating variable

- **Separate mediation models**: estimate mediation models for each moderator-based subgroup and compare the equivalence of the ab point estimates across the subgroups.

  - Z= 0
  - Z=1



Always code Moderator (Z) as a 0 or 1

Dichotomous moderating variable

- H0: abgroup2 - abgroup1= 0
- H1: abgroup2 - abgroup1≠ 0

- If the point estimates (ab) in each subgroup are statistically significant from one another, there is **significant moderation of the indirect effect** (= heterogeneity in the ab product).

- To test the point estimates (ab) for statistical significance the difference is divided by a standard error for the estimate to form a z-statistics.
- Standard error of the difference between the two coefficients is:

  $SE_{pooled} = \sqrt{}(s^2_{ab_{group1}} + s^2_{ab_{group2}})/ 2$

- If the obtained Z value (=difference in point estimates/ spooled) is greater than 1.96 there is a significant moderation in the indirect effect such that *ab* is heterogeneous across moderator-based subgroups

3.3. Longitudinal mediation models

Longitudinal data vs. cross-sectional

- Longitudinal data provides more information regarding the temporal precedence of X, M, and Y (time ordering among variables based on theory or evidence).
  - Longitudinal data allows for examination of whether changes in M are more likely to precede changes in Y
  - Three or more waves of data generally provide accurate representations of the temporal order of change over time that lead to more accurate conclusions about mediation
- In longitudinal data, both changes within individuals and cross-sectional relations can be investigated
  - Cross-sectional data: estimates of effects based on differences among individuals
  - Longitudinal data: estimates of effects based on changes within individuals
- Longitudinal data addresses some alternative explanations of cross-sectional mediated effects
  - Explanation cross-sectional: existence of an omitted variable that explains the relation
  - Longitudinal: measures change within a person, thereby removes some “omitted variable” explanations that are dueto static differences among individuals (e.g. genetics).

Two-wave regression models

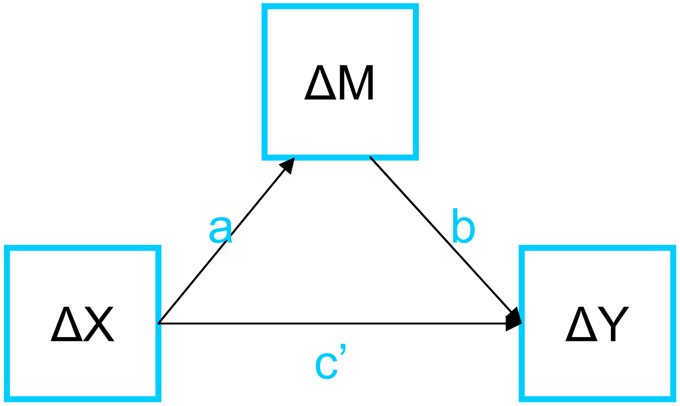
- Absolute change scores
  - $\Delta Y = Y_2 - Y_1$
  - $\Delta M = M_2 - M_1$
  - $\Delta X = X_2 - X_1$
  - Unconditional model: difference score method assumes that without an effect of an independent variable, the differences among individuals at baseline would be maintained at follow-up measurement
  - unreliable
- Analysis of Covariance
  - Baseline value of each variable is included as a covariate in the analysis
  - Conditional model: analysis of covariance method assumes that each individual's score would tend to regress to the mean of scores if unexposed to an independent variable. The rapidity that scored regress to the mean is a function of the amount of measurement error.
- Residual change score: alternative method
  1. Obtain predicted values of the wave 2 measurement using the wave 1 measure
  2. Subtract predicted wave 2 scores from actual wave 2 scores
  3. Residual score= difference between observed wave 2 score and predicted wave 2 score

Two-wave mediation model

- Absolute change scores
  - ΔY= c*ΔX
  - ΔM= a*ΔX
  - ΔY= c′*ΔX +b*ΔM
  - Mediated effect= a*b or c-c′
- Residual change scores
  - ΔY′= c*ΔX′
  - ΔM′= a*ΔX′
  - ΔY′= c′*ΔX′ +b*ΔM′

If X codes exposure to an experimental manipulation, then change in X reflects group membership

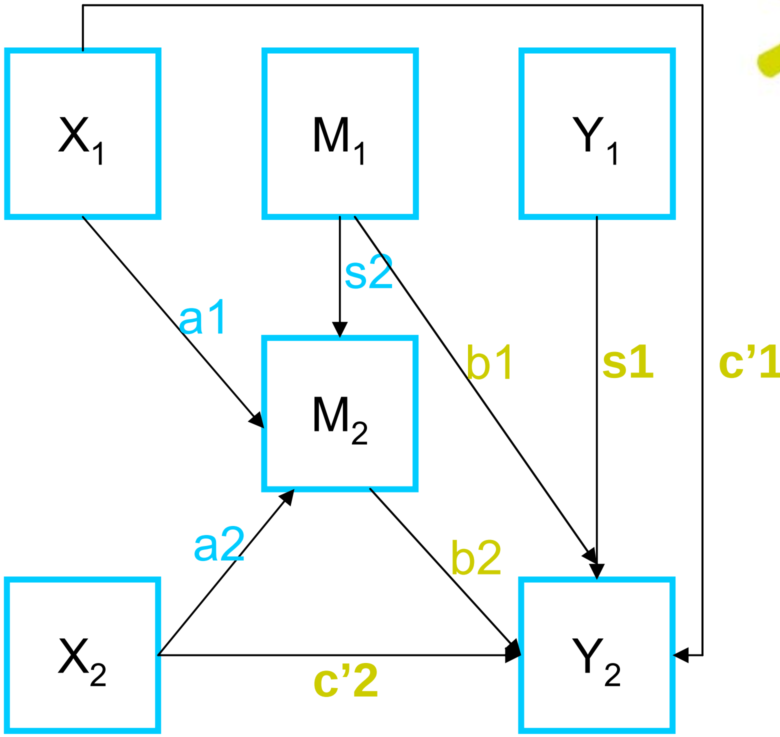
ΔY'= residual change score of Y



ENERGY

Two-wave mediation models

- Analyses of covariance
  - $Y_2 = c_1X_1 + c_2X_2 + s_1Y_1$
  - $M_2 = a_1X_1 + a_2X_2 + s_2M_1$
  - $Y_2 = c_1'X_1 + c_2'X_2 + b_1M_1 + b_2M_2 + s_1Y_1$



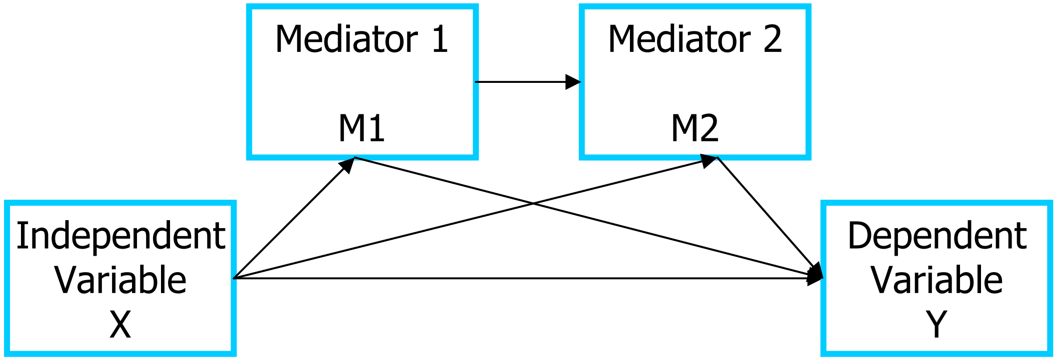
More complex models

- Still no true causal relations
- Three-or-more-waves models
  - Auto-regressive models
  - Cross-lagged autoregressive models
  - Latent Growth Curve Models
    - Models growth over time + individual differences over time
- Other statistical programs needed (Mplus, LISREL)

3.4. Complex mediation (and moderation) models

Advanced mediation models

- Latent Growth curve models
- Multilevel mediation models
  - Schools
  - neighborhoods
- Mediation of a mediated effect (path models)
- Mediation in Structural Equation Models
  - More complex models
  - Takes measurement errors into account



To summarize…

- With a **categorical outcome variable**, the product-of-coefficient test and the earlier standard error calculation, can be used to conduct a mediation analyses
- The difference in coefficient test can give distorted results

- **Moderation of a mediated effect** can be executed by estimating separate mediation models; in which mediation models for each moderator-based subgroup are estimated, the equivalence of the ab point estimates across the subgroups are compared and divided by the pooled standard error.

- Two-wave **longitudinal mediation analyses** can be conducted by using absolute change scores, residual change scores or by using the analyses of covariance model. The absolute change score are less reliable. The residual change scores and covariance model yield similar results

- For more complex mediation models, SPSS is not sufficient, and more computer intensive programs such as Mplus are needed.

Exercises 3

Exercises

- Practice with your own data, OR
- Exercise 3.1: longitudinal mediation analyses
- Exercise 3.2: moderation of a mediated effect
- Exercise 3.3: mediation analyses with a categorical outcome variable