Theory Construction and Causal Modeling

James Jaccard
Silver School of Social Work
New York University

Overview

First half of workshop will focus on theory construction using causal concepts. It will have a conceptual focus

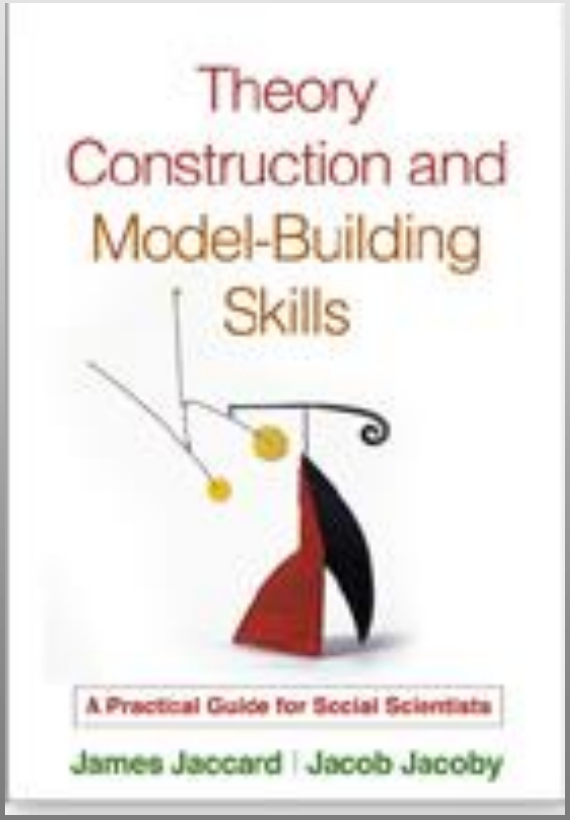
Second half of the workshop will focus on the statistical analysis of causal theories using limited information estimation frameworks

I assume you have working knowledge of multiple regression and hypothesis testing

Theory Construction

Theory Construction

Focus is on the practical mechanics of constructing theory rather than philosophy of science



Causal Theories

Most (but not all) theories rely heavily on the concept of causality, i.e., we seek to identify the *determinants* of a behavior or mental state and/or the *consequences* of a behavior or environmental/mental state

- **Why do some people suffer from depression and others do not? What are the causes of depression?**

- **Why do some people engage in unsafe sex while other people do not engage in unsafe sex? What are the causes of sexual risk behavior?**

- **Why do some people make use of mental health services but other people do not? What are the causes of service utilization?**

Causal Theories

Sometimes instead of *causes*, we are more interested in the *effects* of experiences or behaviors, but this is still part of causal analysis

- **What are the effects on a child of being raised by parents who are physically abusive?**

- **What are the effects of being adopted on a person's mental health?**

- **What are the effects of participating in a program to reduce stress on one's health?**

Building Causal Theories

Theory Construction

All causal theories work with *variables* and presumed *causal relationships* between variables.

The first step in constructing a causal theory is to identify the phenomena you want to understand or the variable whose variability you want to explain

Unintended pregnancy

Alcohol abuse and drunk driving

Child anxiety

Choosing an Outcome Variable

Choice of an outcome is often impacted by

- **Your values and what you see as important and interesting**
- **What your “major professor” or advisor studies**
- **What is fundable (ugh…)**
- **Participatory action research**

Choosing an Outcome Variable

In making your choice, be sure you can clearly answer the core question "why is this an important topic?"

Be careful about focusing on variables that are <u>too general or too abstract</u> (e.g., "adolescent risk behavior" vs. "adolescent sexual risk behavior")

In making your choice, you will want to <u>specify your target population</u> (infants, children, adolescents, young adults, adults, the elderly)

Conceptual Definitions

Whenever you introduce a variable into a model, it is important that you have a clear *conceptual definition* of that variable

Write out your definition in clear and unambiguous terms

See my theory construction book for six practical strategies for constructing clear conceptual definitions

Grand versus Mid-Level Theories

Grand theories are broad-based frameworks that identify broad classes of variables we should think about when we want to explain a phenomenon.

- **Bronfenbrenner’s ecological systems theory (individual, microsystem, mesosystem, exosystem, macrosystem)**
- **Structuralism, functionalism, materialism, systems theory, symbolic interactionism, behaviorism**

Grand theories have their place, but they typically are too general to address specific problems in ways that will yield practical, evidenced based strategies for solving problems

Grand versus Mid-Level Theories

Mid-level theories identify specific variables that are clearly defined, that are directly tied to the outcome variable, and that elucidate specific causal mechanisms that are directly testable.

Most social issues require that we ultimately address them using mid-level theories

The Building Blocks of Causal Theories

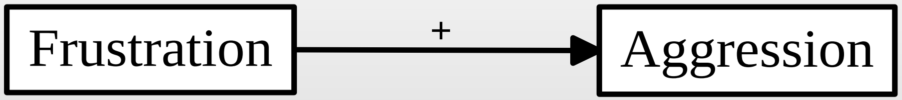
Causal Theories

Causal theories can be complicated, but at their core, there are <u>six types of relationships</u> that can be part of a causal theory

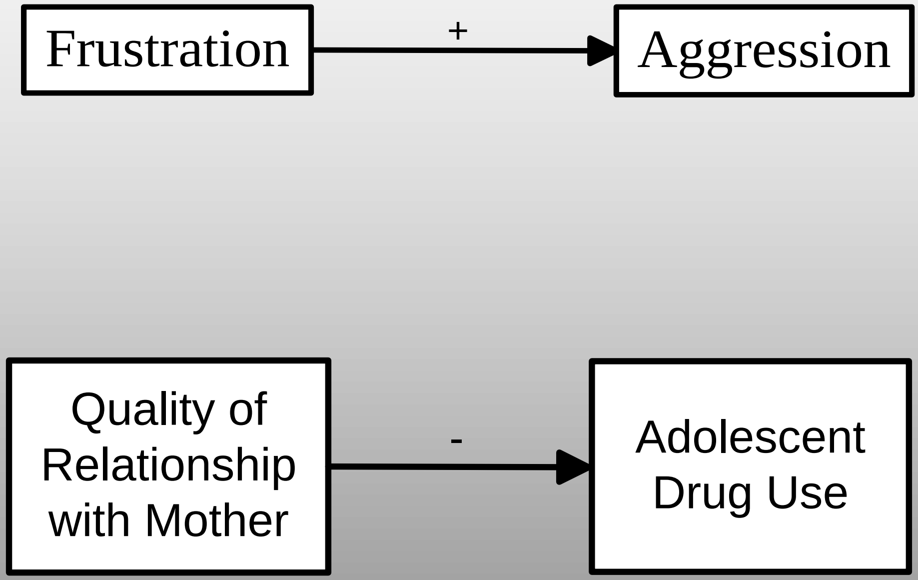
Direct Causal Relationships

A direct causal relationship is when a variable, X, has a direct causal influence on another variable, Y:

Direct Causal Relationships



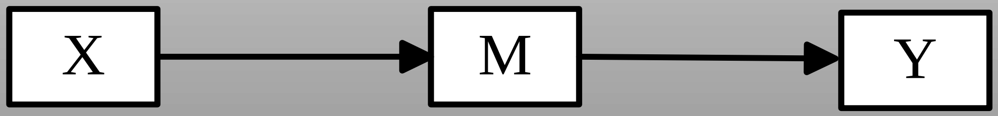
Direct Causal Relationships



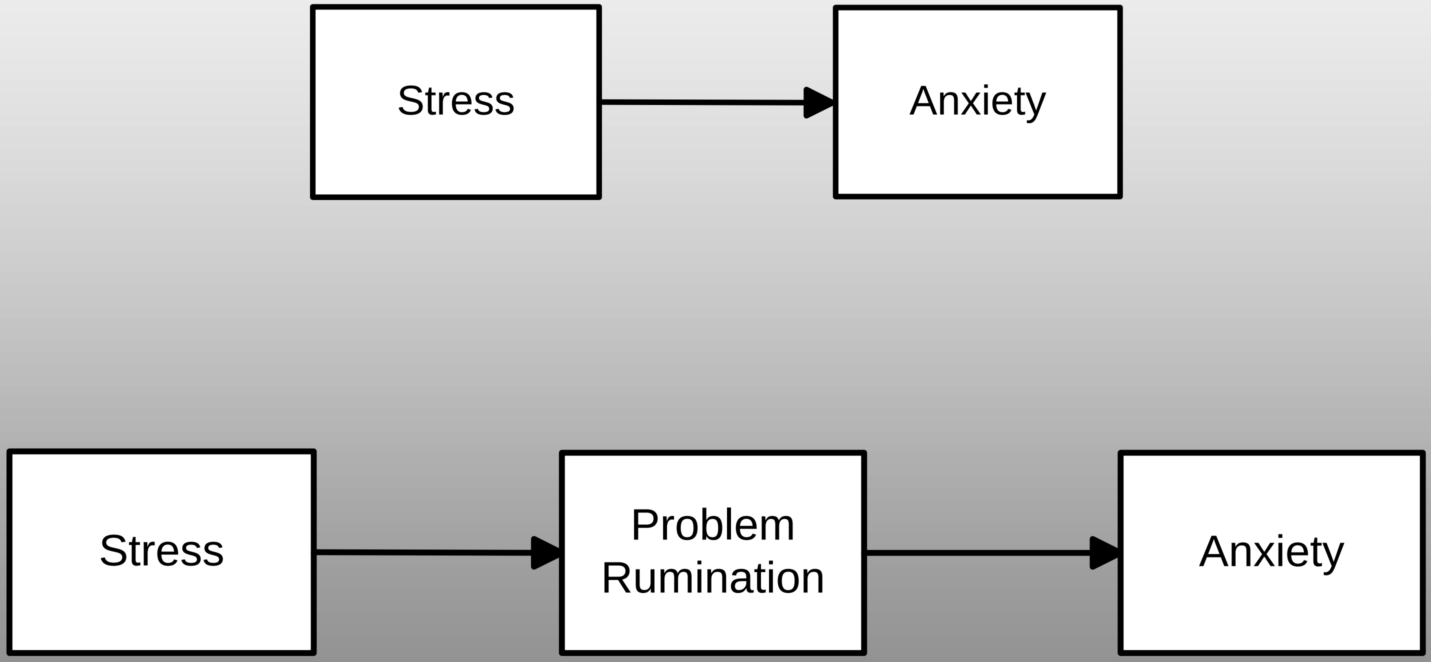
Indirect Causal Relationships

Indirect Causal Relationships

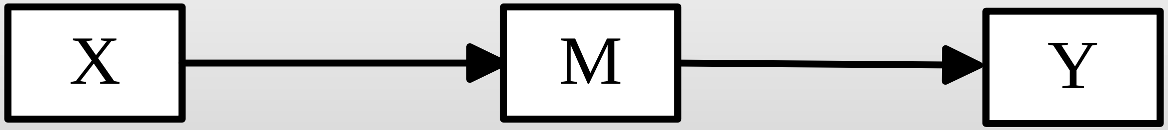
An indirect causal relationship is when a variable, X, has a causal influence on another variable, Y, through an intermediary variable, M:



Indirect Causal Relationships



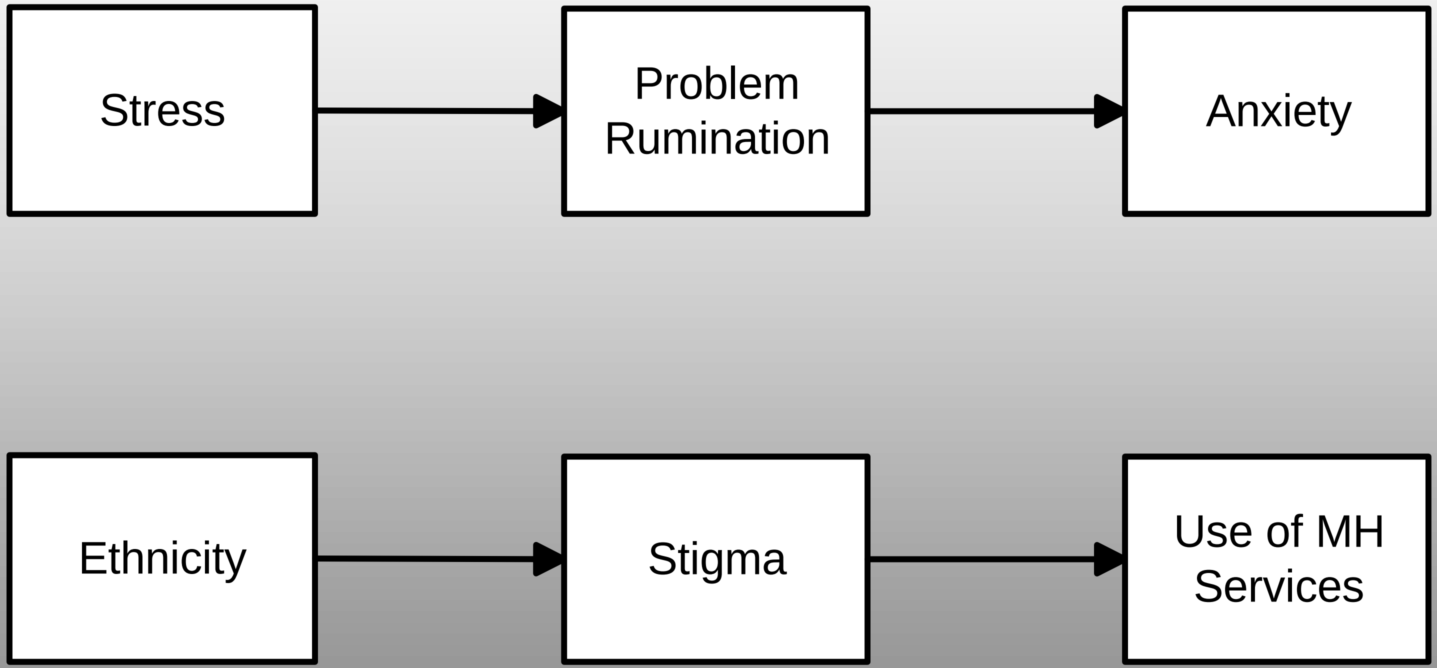
Indirect Causal Relationships



The intermediary variable is called a *mediating variable* or a *mediator*

A mediator identifies the mechanism by which X influences Y. It answers the question "why does X influence Y"

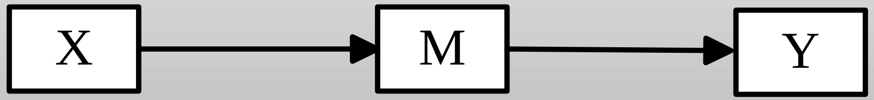
Indirect Causal Relationships



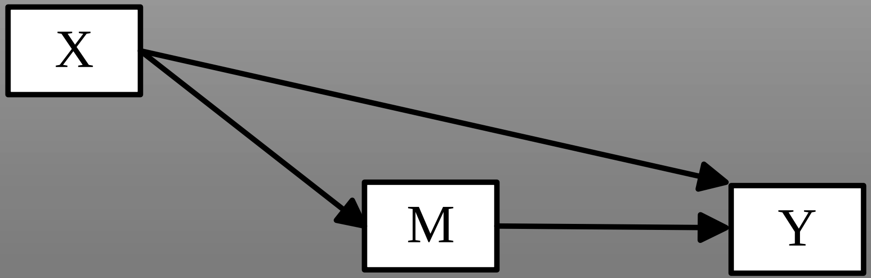
Mediated Relationships

Whenever you specify a mediated relationship, you must address in your theory the issue of complete versus partial mediation

(a) The case of complete mediation

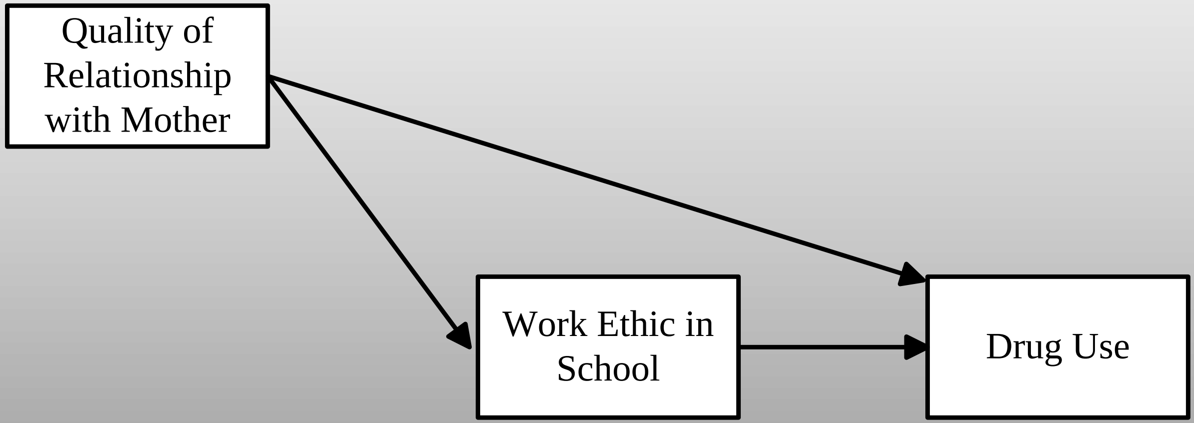


(b) The case of partial mediation



Mediated Relationships

Example of partial mediation:

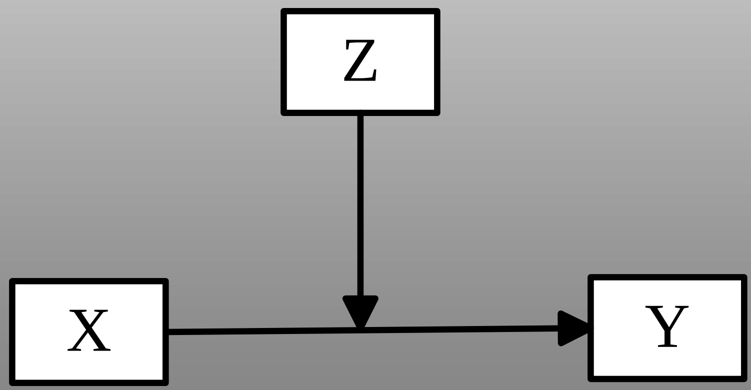


Theoretical justification: Quality of relationship also impacts likelihood adolescent hangs out with negative peers

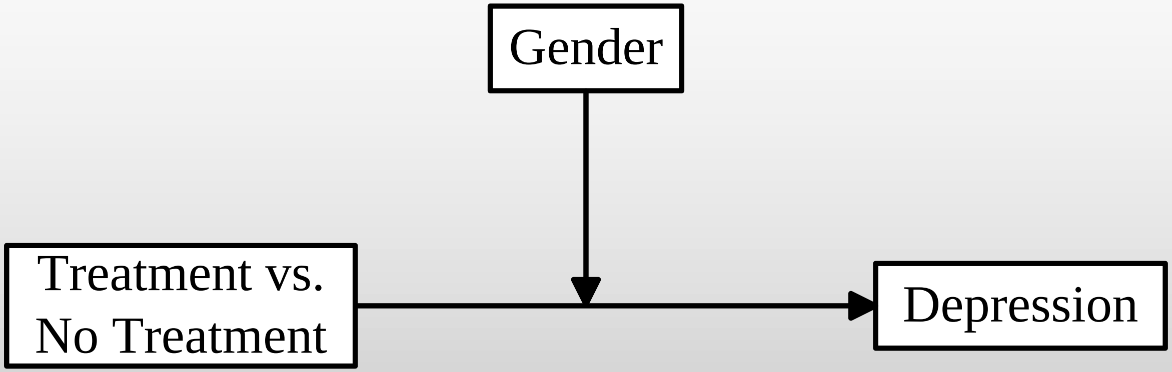
Moderated Causal Relationships

Moderated Causal Relationships

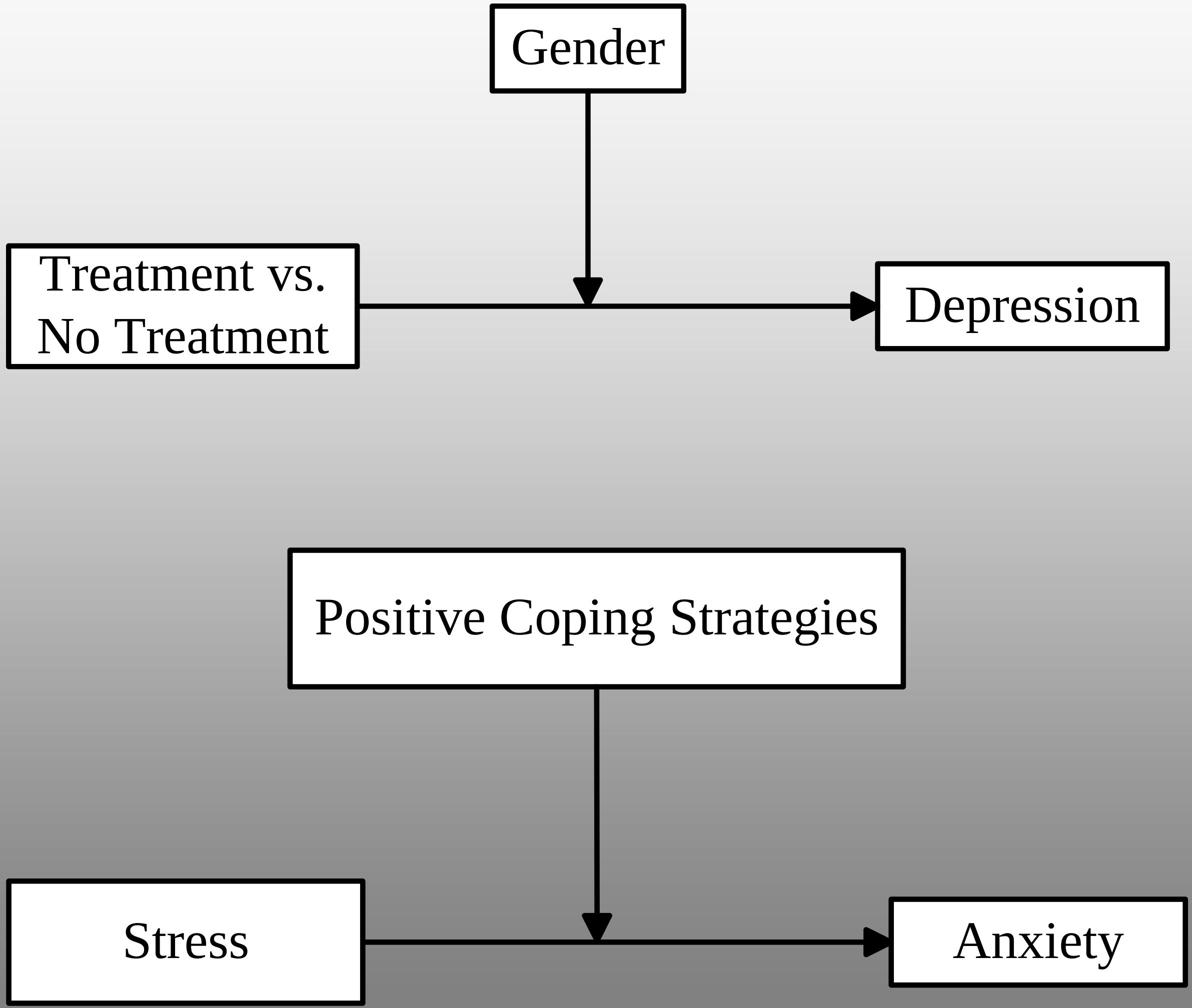
A moderated causal relationship is when the *impact* of a variable, X, on another variable, Y, differs depending on the value of a third variable, Z



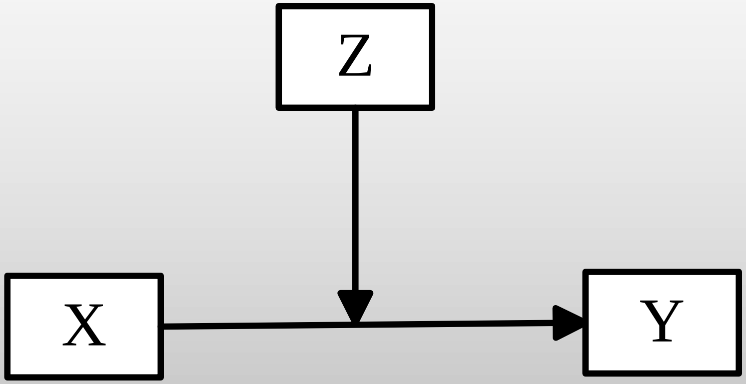
Moderated Causal Relationships



Moderated Causal Relationships



Moderated Causal Relationships



The variable that "moderates" the relationship is called a *moderator* variable.

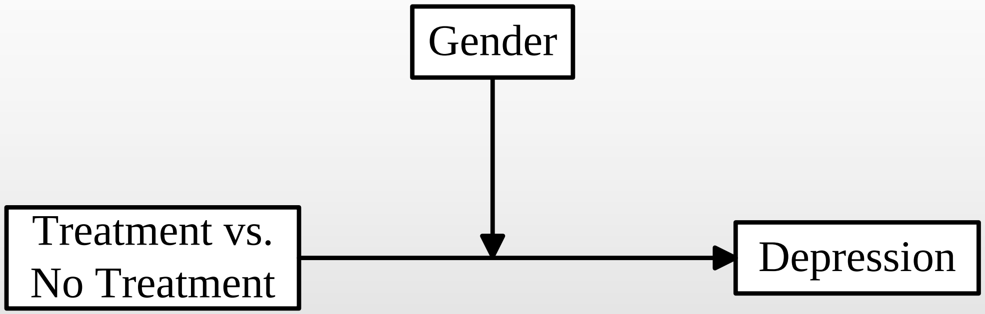
Be careful not to confuse the terminology "moderator" with "mediator." They are very different causal dynamics!

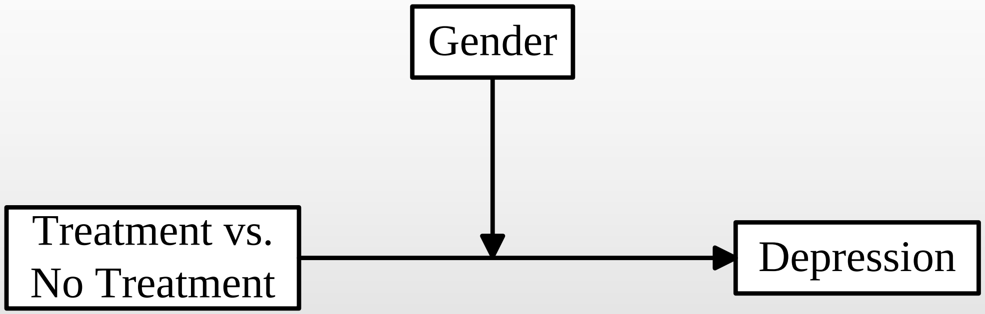
Moderated Causal Relationships

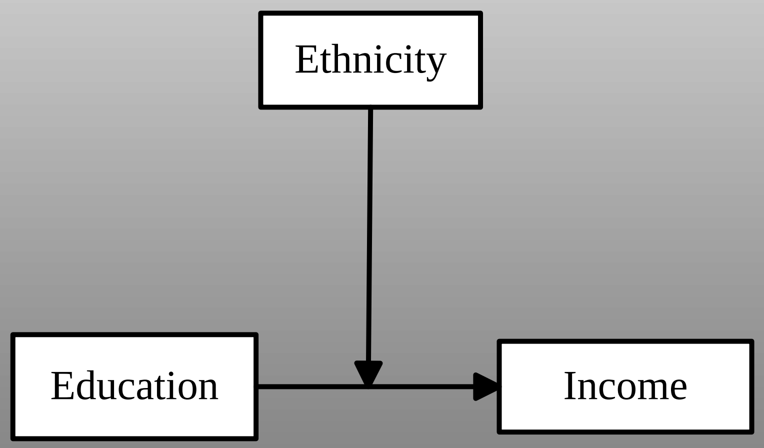
Moderators identify subpopulations and conditions where causal relationships exist (or are stronger) versus subpopulations and conditions where they do not exist (or are weaker)

Common moderating individual difference variables are

Age/grade	**Gender**
Ethnicity	**Social class**


Gender
Treatment vs.
No Treatment
Depression


Gender
Treatment vs.
No Treatment
Depression


Ethnicity
Education
Income

Moderated Causal Relationships

Mediation asks the question *why*. Moderation asks the questions for *who*, *where* and *when*

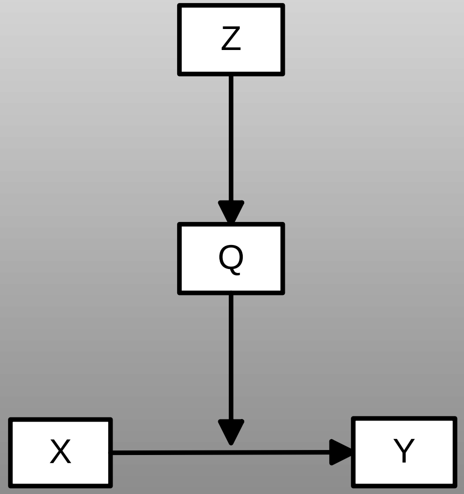
Who? (for whom does this apply and for whom not)

Where? (where does this apply and where not)

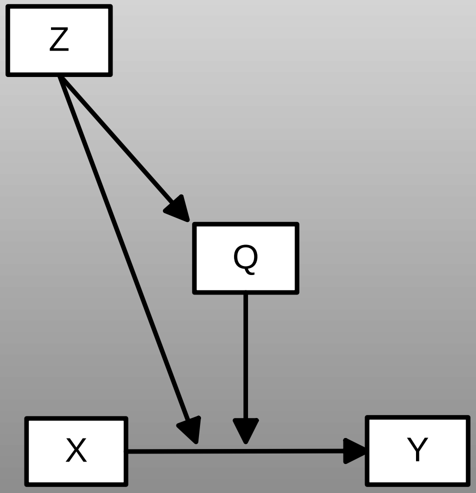
When? (when does this apply and when not)

Can Combine Mediation and Moderation in Causal Models

Mediated Moderation

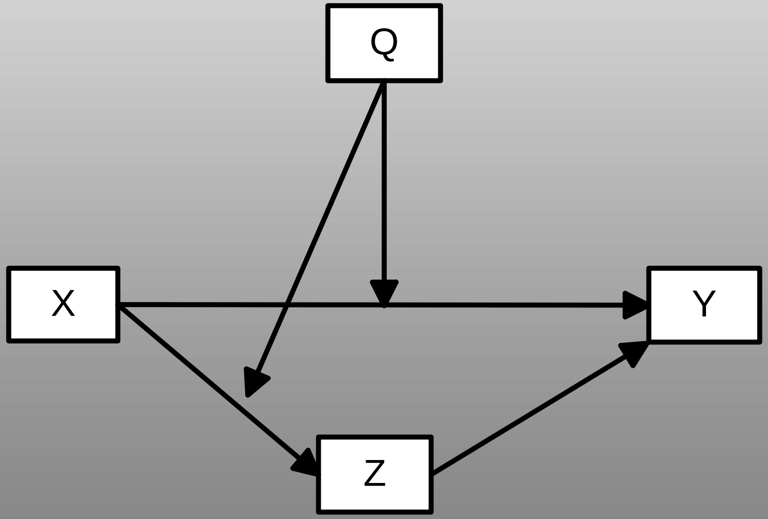


Partial Mediated Moderation

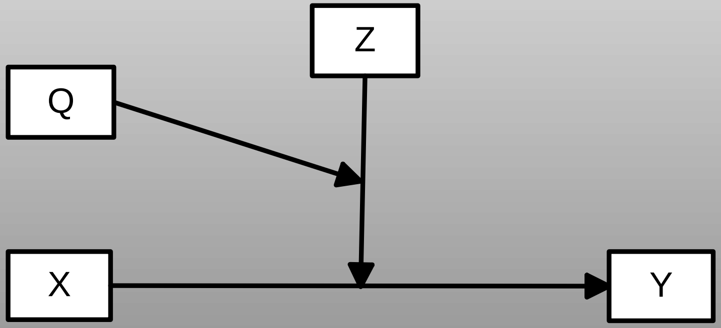


Can Combine Mediation and Moderation in Causal Models

Moderated Mediation



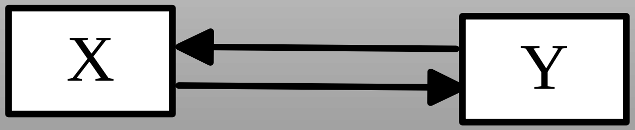
Moderated Moderation



Bidirectional Causal Relationships

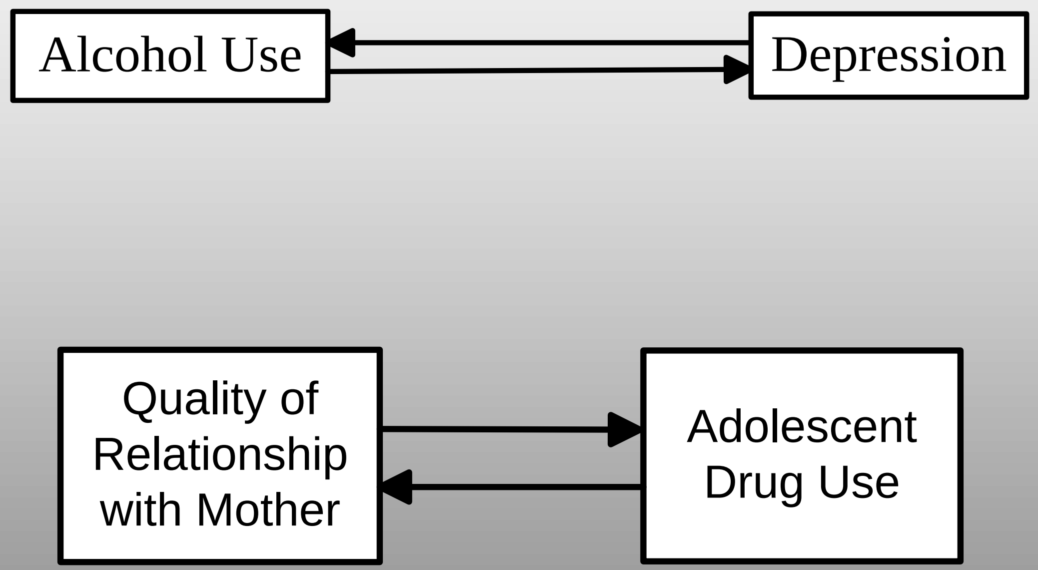
Bidirectional Causal Relationships

A bidirectional causal relationship is when a variable, X, has a causal influence on another variable, Y, and that effect, Y, has a "simultaneous" impact on X:



Bidirectional Causal Relationships

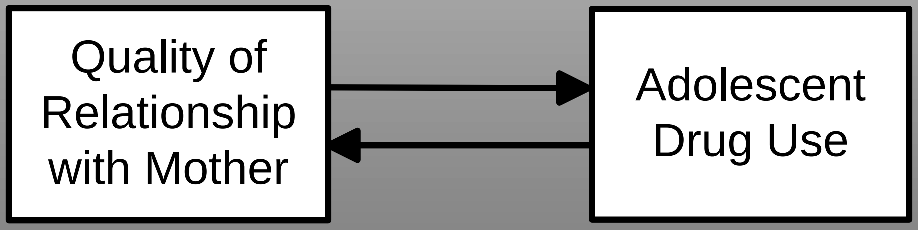
Bidirectional Causal Relationships



Reciprocal Causality

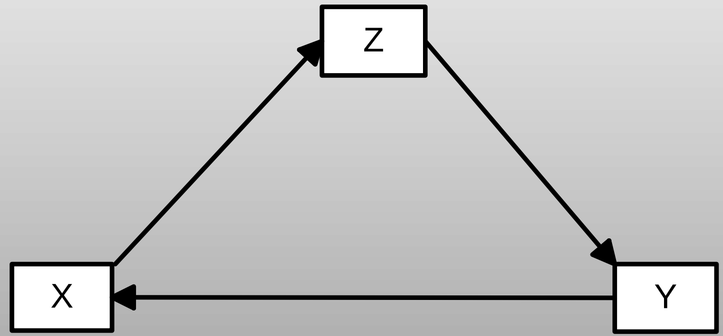
Technically, there is no such thing as reciprocal causality

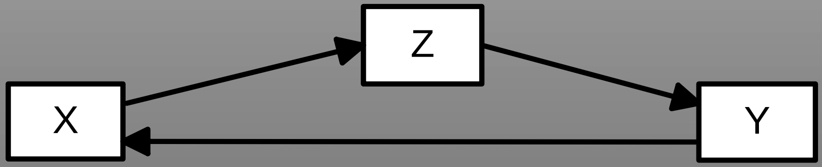
Which translates into:



Feedback Loops

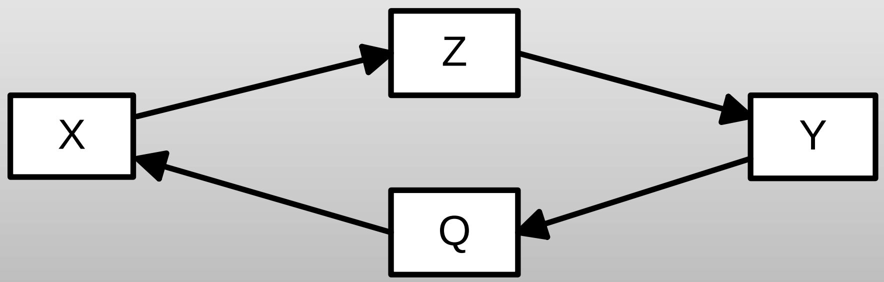
Feedback loops in models are simply a combination of bi-directional effects and indirect effects

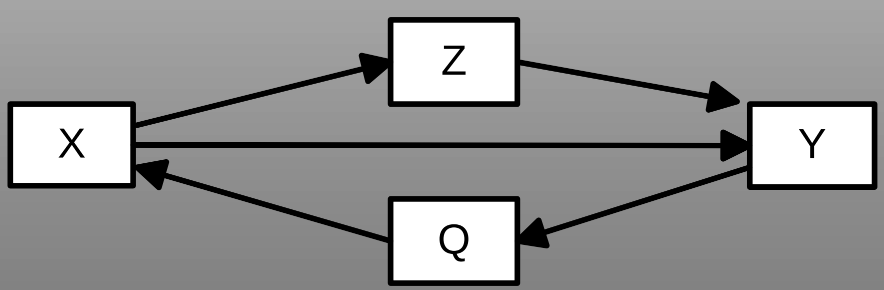




Feedback Loops

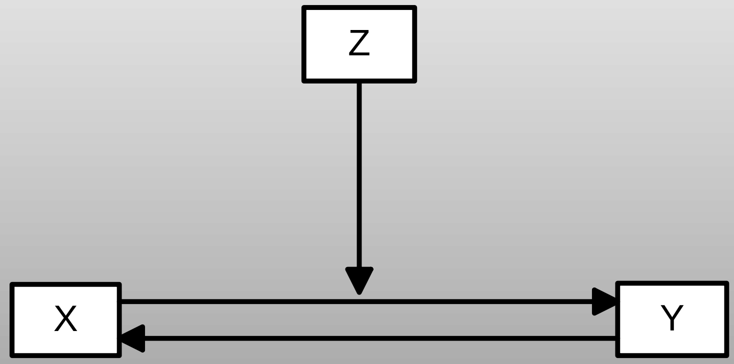
We can have more than one mediator

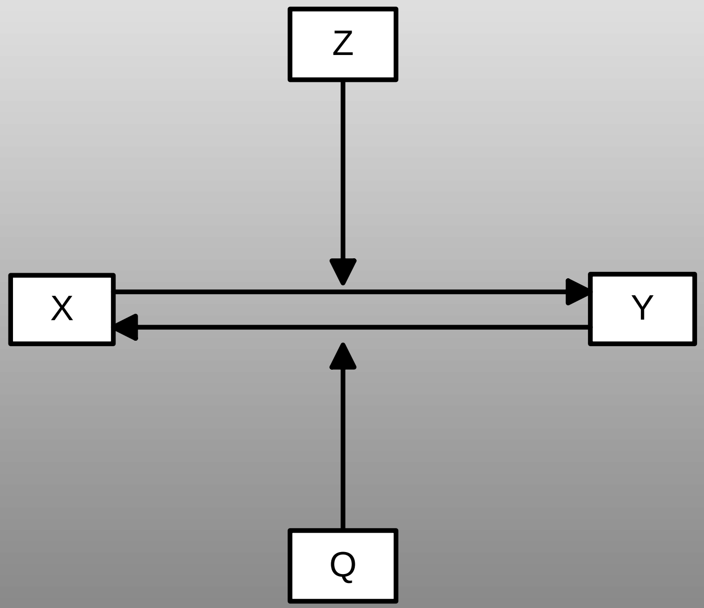




Bidirectional Causal Relationships

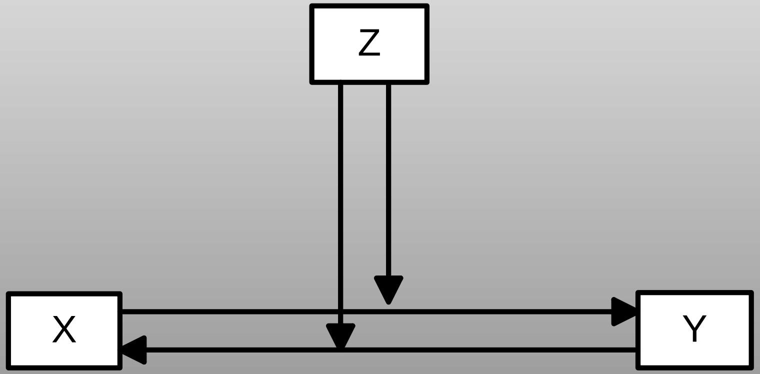
We can also add moderators





Bidirectional Causal Relationships

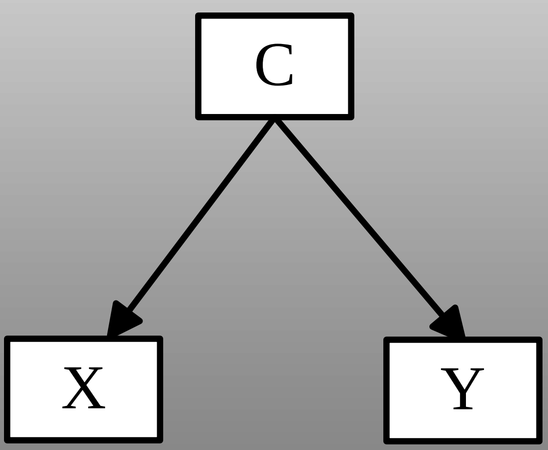
Here is an example of one moderator variable with two moderated relationships



Spurious Relationship

Spurious Relationship

A *spurious relationship* is one where two variables that are not causally related share a common cause:



Spurious Relationship

Examples

There is a relationship between shoe size and verbal intelligence: The bigger your feet, the smarter you are!

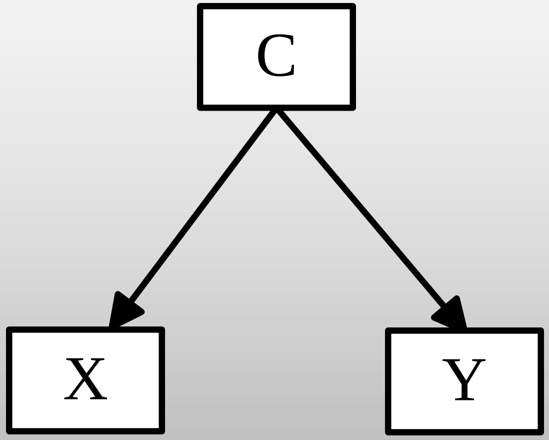
Spurious Relationship

Examples

There is a relationship between shoe size and verbal intelligence: The bigger your feet, the smarter you are!

There is a relationship between how long your hair is and how tall you are: The shorter your hair, the taller you grow!

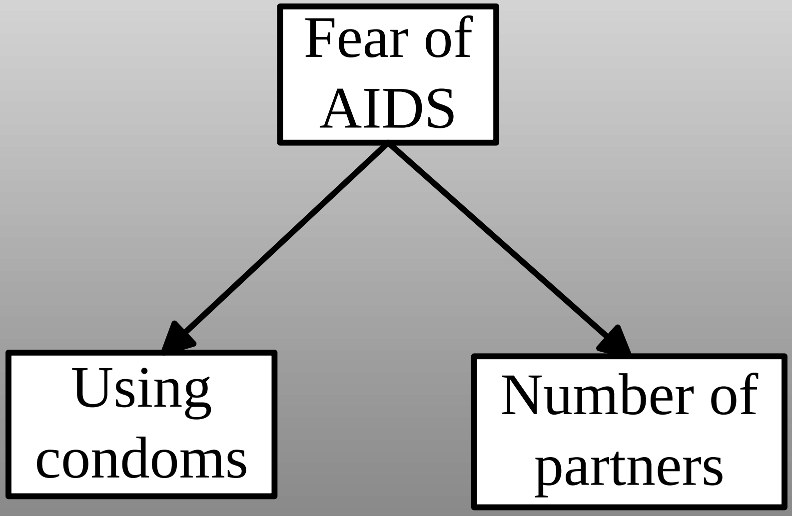
Spurious Relationship



The common cause is called a *confound*, but this term is used to characterize other scenarios in the social sciences as well

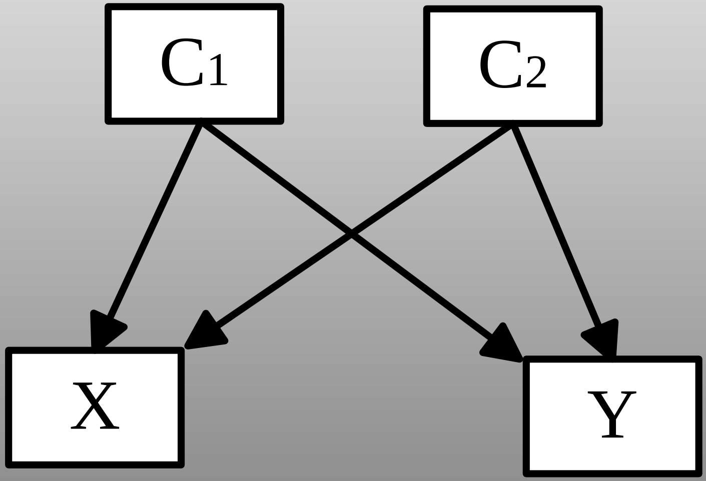
Spurious Relationship

Spurious relationships are not necessarily “bad” or misleading. They can be conceptually meaningful



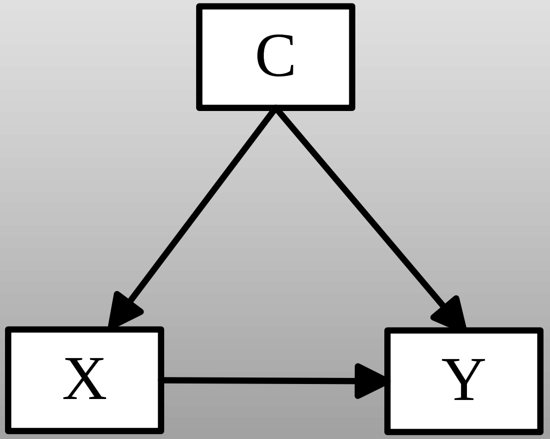
Spurious Relationship

There can be more than one confound in a model:



Spurious Relationship

Relationships can be *both* causal and spurious



Unanalyzed Relationships

Unanalyzed Relationships

An unanalyzed relationship is one where the possibility of two variables being correlated is acknowledged, but there is no interest in describing the causal dynamics between the variables, if there are any. The relationship will remain "unanalyzed."

Summary of Core Relationships

Direct causal relations

Indirect (mediated) causal relations

Moderated causal relations

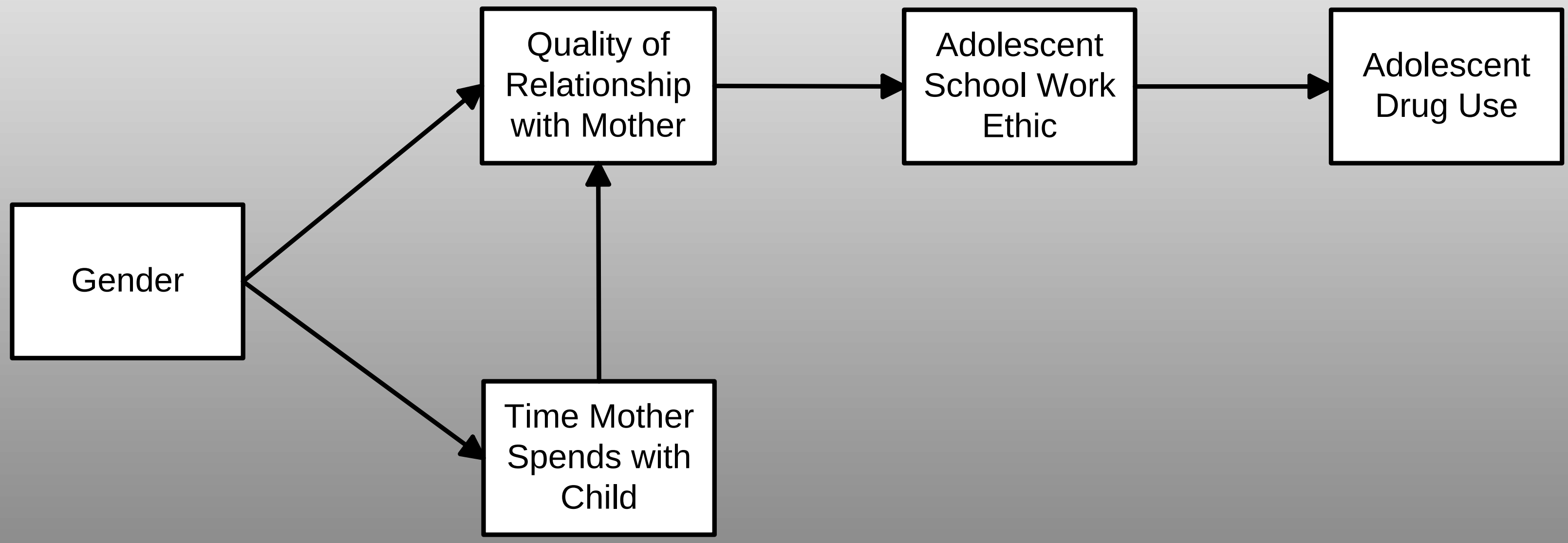
Bidirectional causal relations

Spurious relations

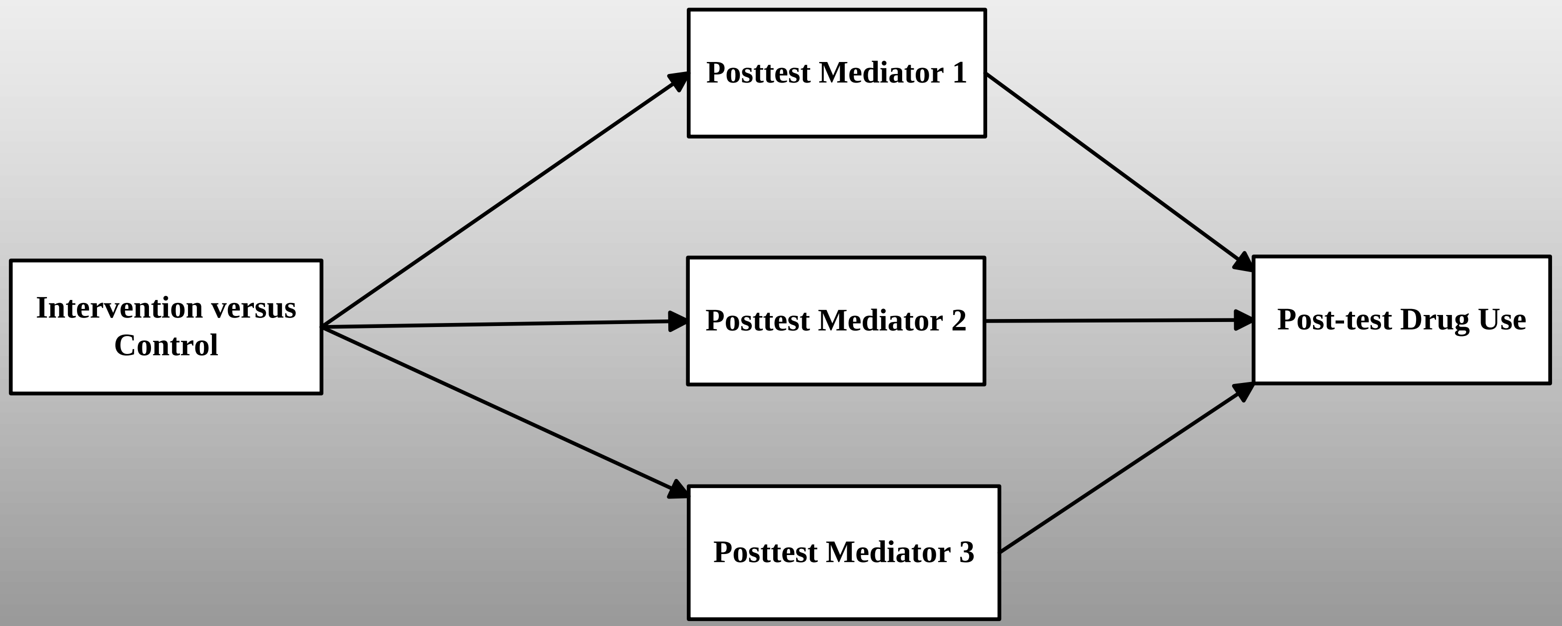
Unanalyzed relations

Causal Theories

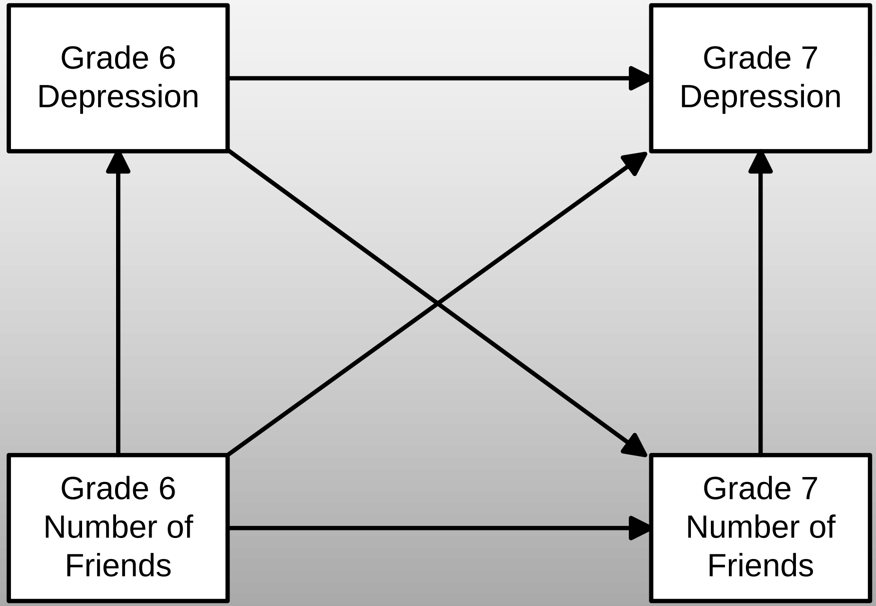
We put all these ideas together to build complex theories of phenomena. Here is one example:



An Intervention With Three Mediators

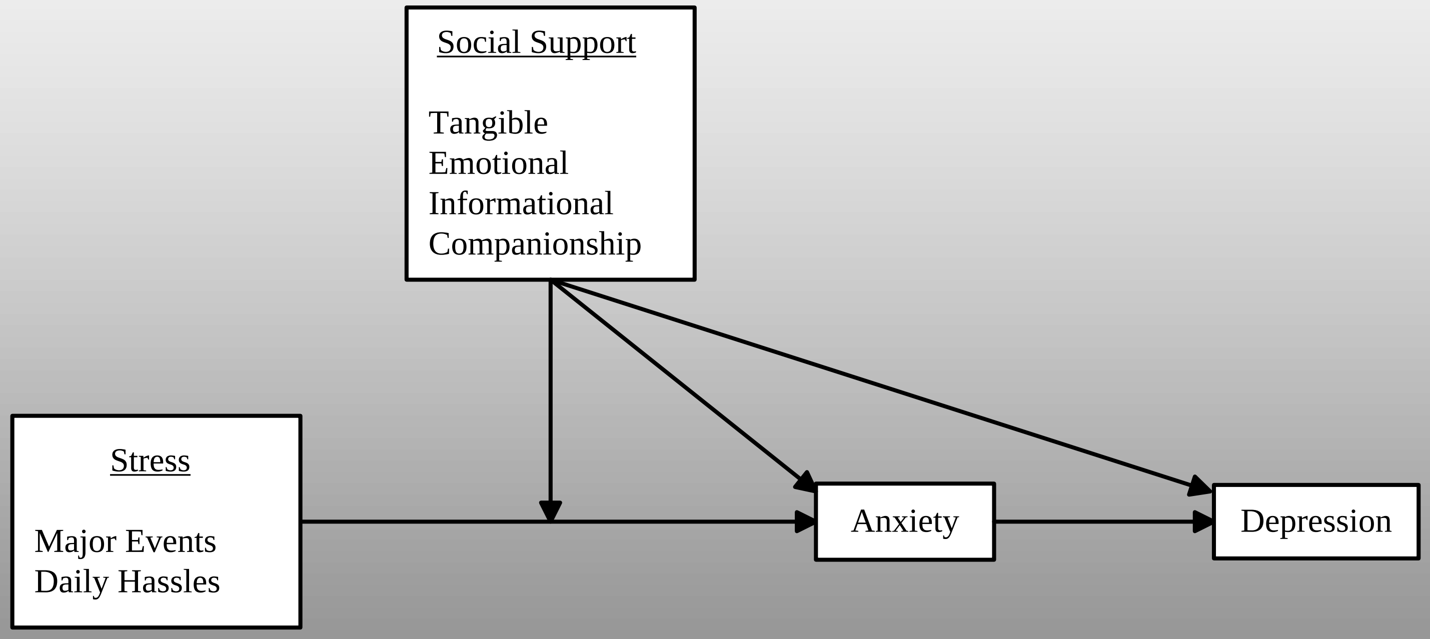


A Longitudinal Model



This model has autoregressive effects, contemporaneous effects, and lag effects

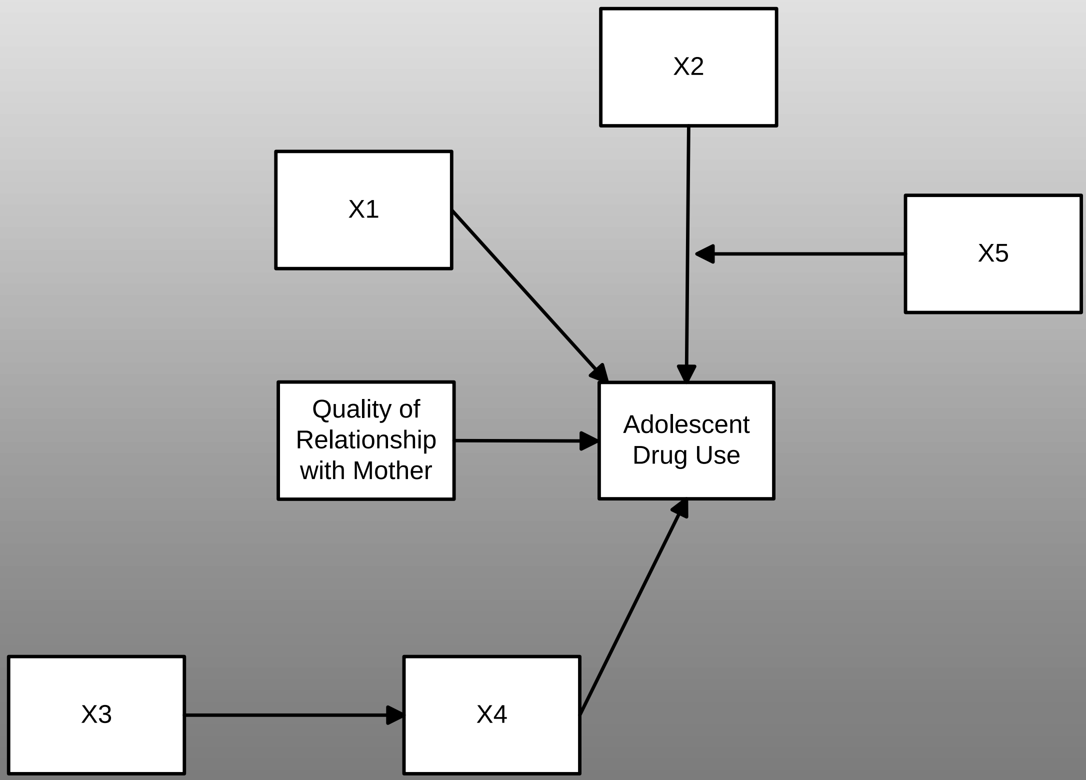
Portraying Complex Models



Building Your Own Original Theory

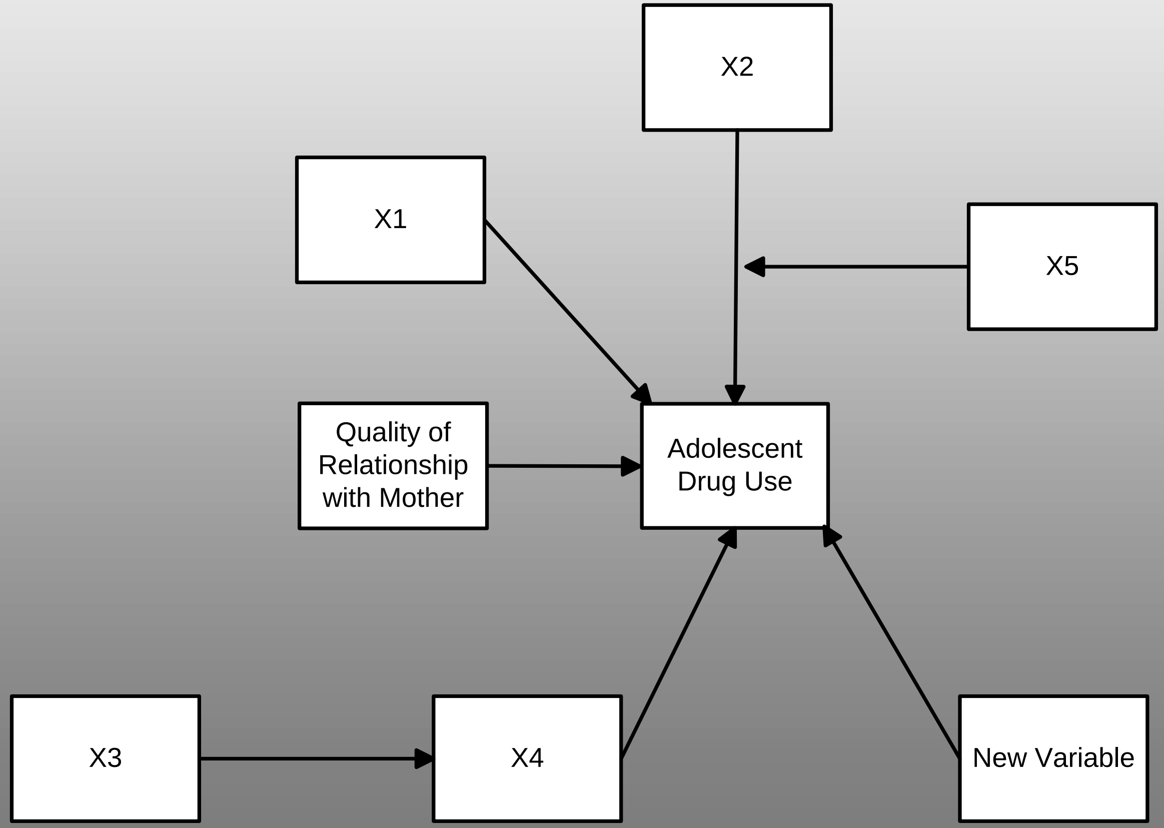
Building Your Theory

After identifying your outcome, read the extant literature on it and draw an influence diagram that captures what the current "state of knowledge" has to say:



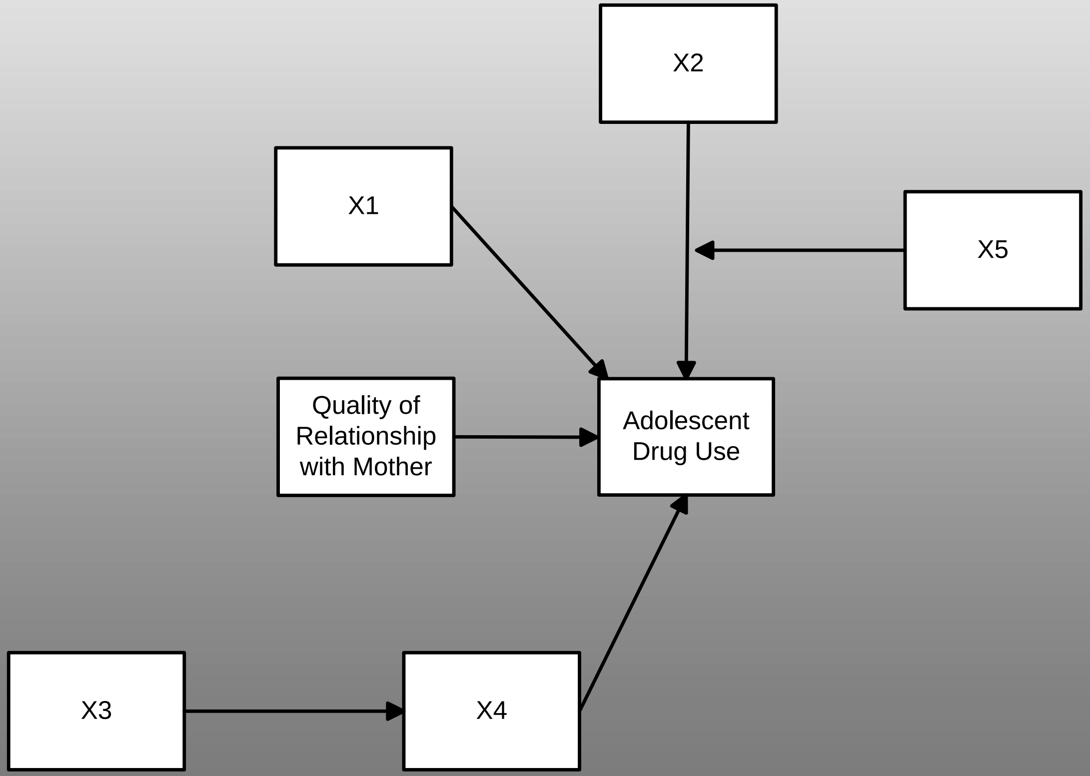
Building Your Theory

If you think of a new variable that the literature has ignored, then add it to the system; that is a theoretical contribution!



Building Your Theory

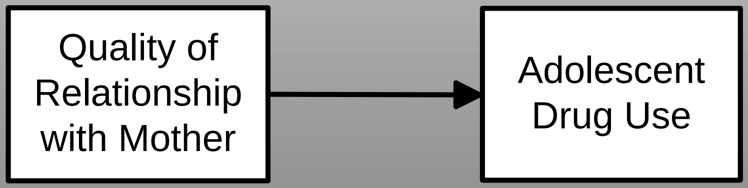
Or, take an existing relationship and expand on it in novel ways to expand and strengthen the theory



Building Your Theory by Adding Mediation

Consider turning the direct relationships into an indirect relationship

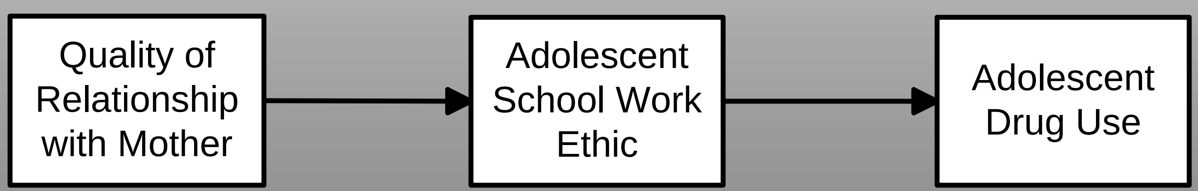
There are three strategies for turning a direct relationship into an indirect or mediated relationship



Building Your Theory by Adding Mediation

<u>Strategy 1</u>: Use the *why heuristic*. i.e. state *why* one variable influences the other variable. The articulation will reveal the mediator

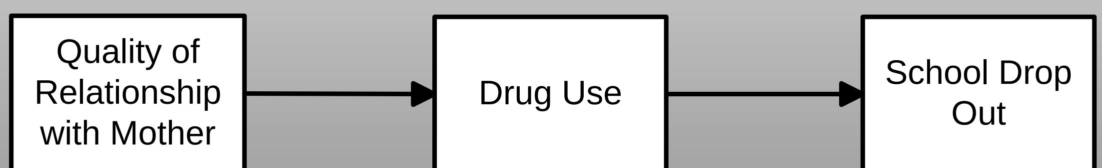
Example: Why does a poor relationship with parents lead to drug use?



Building Your Theory by Adding Mediation

<u>Strategy 2</u>: ***Identify an effect of an effect.*** **Think of your outcome variable as a direct cause of another variable**

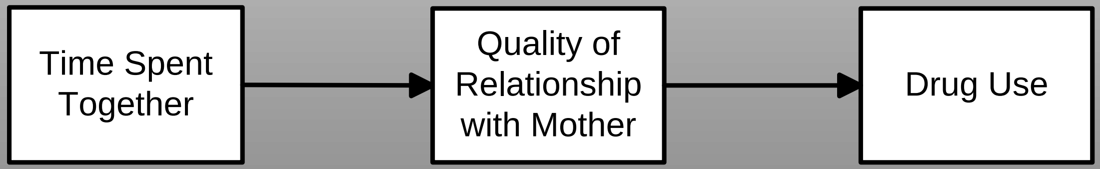
Example: Drug use leads to dropping out of school



Building Your Theory by Adding Mediation

<u>Strategy 3</u>: ***Identify a cause of a cause*. Think of your cause as an outcome variable and identify a direct cause of it**

Example: Time mother and child spend together affects the quality of the relationship with the mother



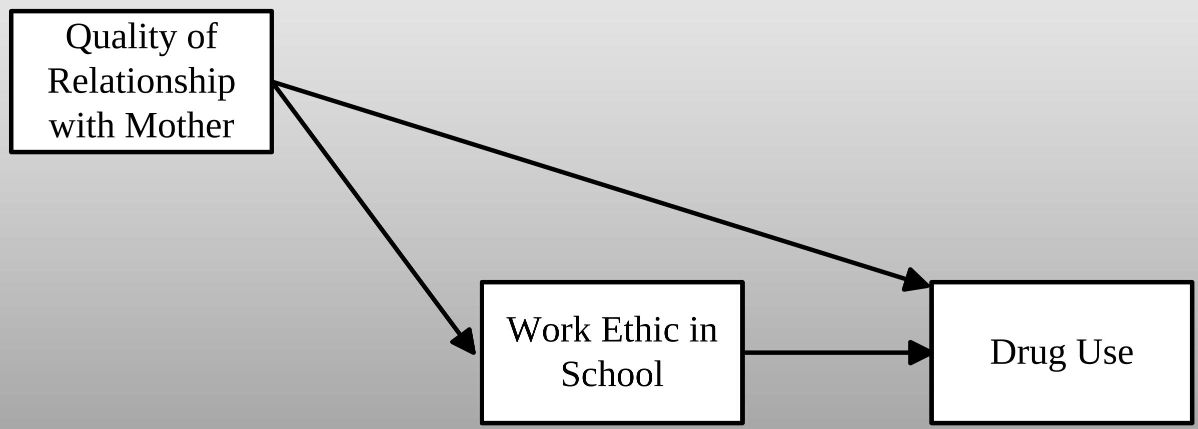
Building Your Theory by Adding Mediation

We often end up with mediated relationships with multiple links in the mediational chain

We are interested in "new" relationships within the chain that represent theoretical innovations

Building Your Theory by Adding Mediation

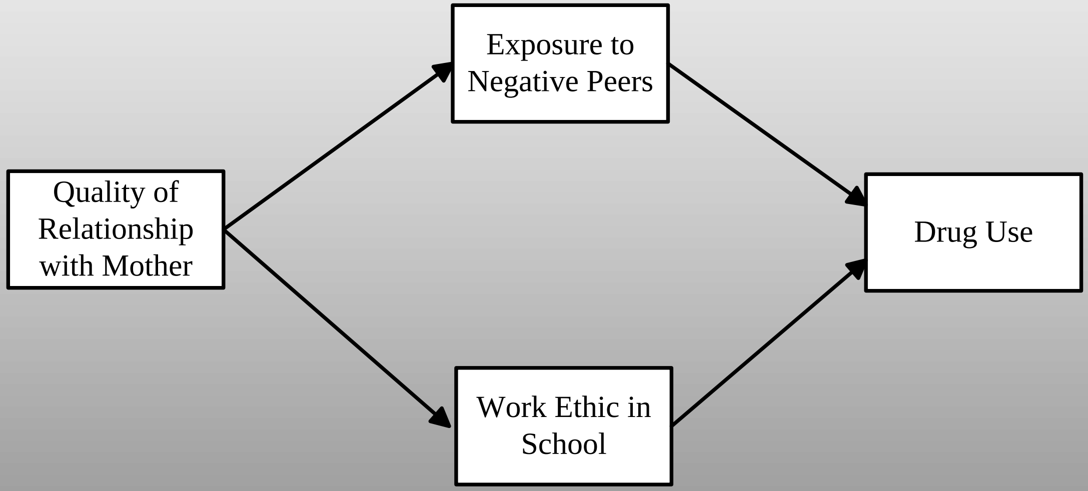
For each mediated relationship, specify if there is partial or complete mediation:



Theoretical justification: Quality of relationship also impacts likelihood adolescent hangs out with negative peers

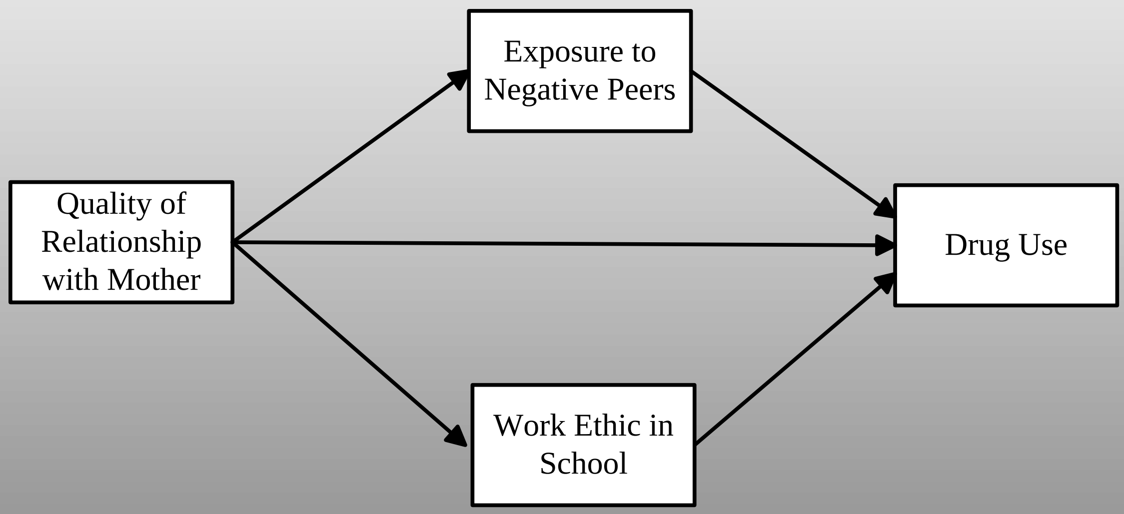
Building Your Theory by Adding Mediation

Why not:



Building Your Theory by Adding Mediation

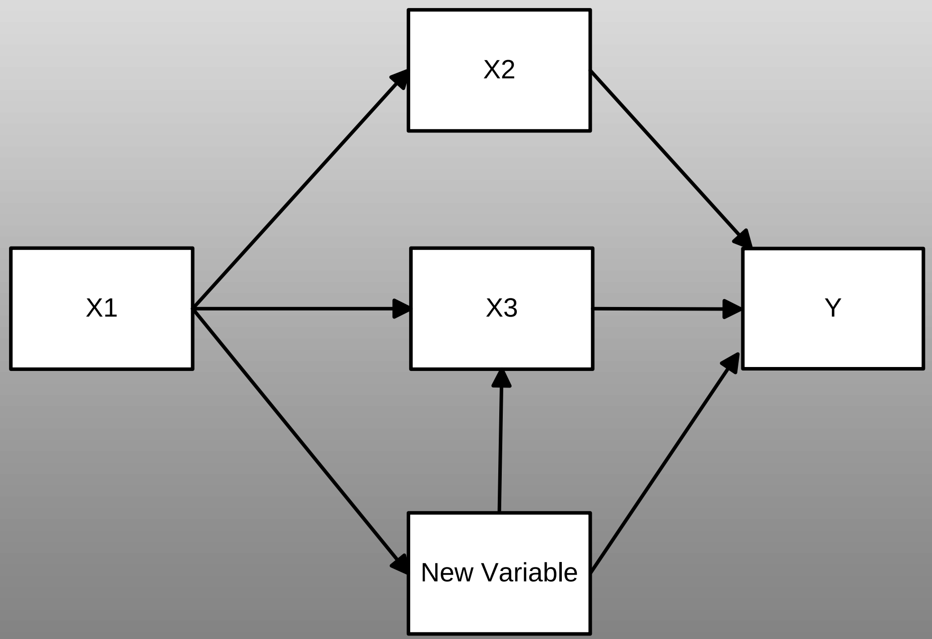
Can do so, but then must address issues of partial versus complete mediation for these mediators:



The process is never-ending. At some point, we need to close out the system

Building Your Theory by Adding Mediation

Maybe the literature has identified two mediators, but you think of a mediator that no one has considered. This is a theoretical contribution – a new mechanism



Build Your Theory by Adding Bidirectional Effects

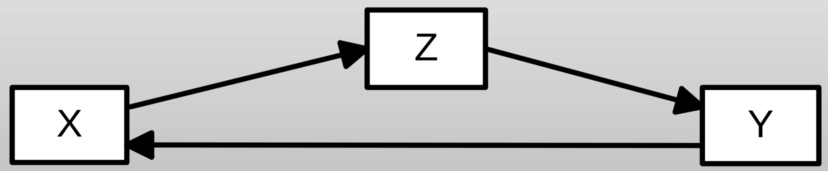
Consider turning the direct relationships into a reciprocal causal effect

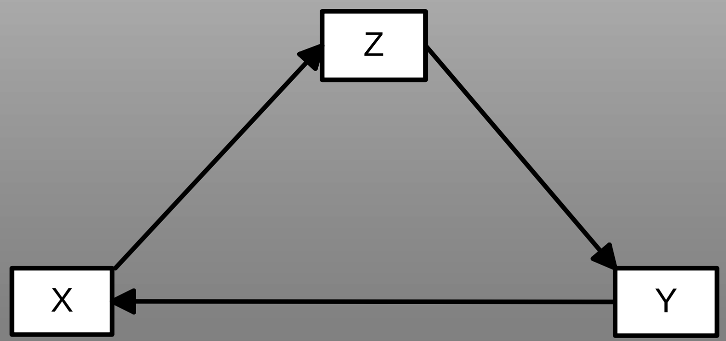
Ask yourself if it is feasible for the causal direction to go the other way

- **Gender impacts depression**
- **Religiosity impacts adolescent sexual behavior**
- **Childhood trauma impacts adolescent depression**

Build Your Theory by Adding Bidirectional Effects

Consider turning the bidirectional relationship into a feedback loop by adding a mediator to it





Build Your Theory by Adding Moderated Effects

Pick a direct relationship to focus on and then try to identify a moderator of it

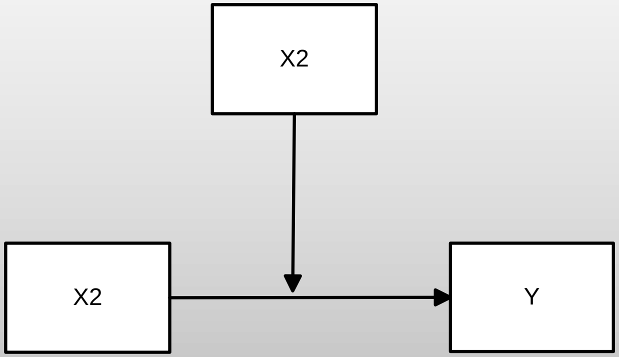
To identify moderators, ask the questions and their negations:

Who? (for whom does this apply and for whom not)

Where? (where does this apply and where not)

When? (when does this apply and when not)

Build Your Theory by Adding Moderated Effects



After doing so, consider adding mediated moderation, moderated moderation, or moderated mediation

Common Mistakes

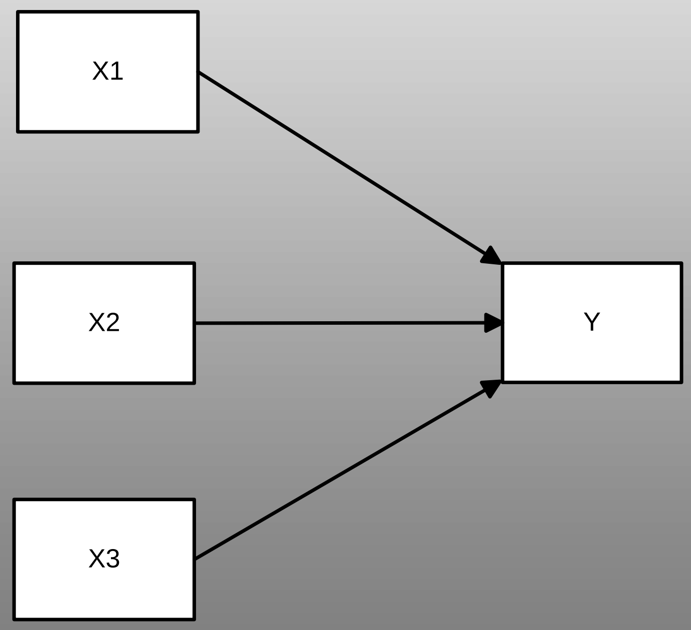
Model is too complex – too many variables

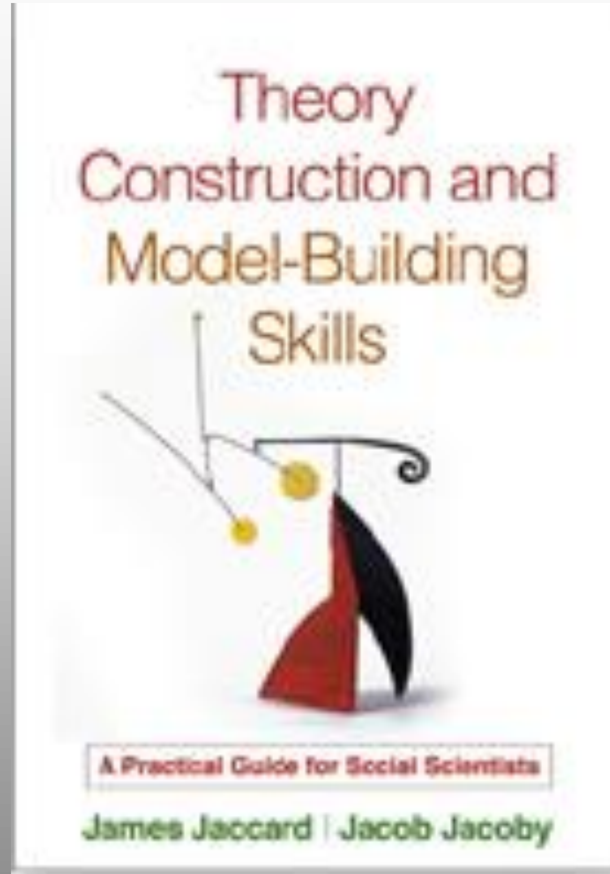
Variables are not clearly defined

Variables are too abstract and too general

Common Mistakes

Know that when you run a single equation model (one regression equation, one logistic equation, one survival analysis), the model you are assuming is this:




Theory
Construction and
Model-Building
Skills
A Practical Guide for Social Scientists
James Jaccard | Jacob Jacoby

Analyzing Causal Models

Analyzing Causal Models

Most people associate causal modeling with *structural equation modeling* (SEM; also called *path analysis*)

There are two general forms of SEM analysis, *limited information estimation* and *full information estimation*. Most people are familiar with just the latter

Each has strengths and weaknesses. I will be showing you limited information estimation.

Analyzing Causal Models

When social scientists address causal models, they rely heavily on the linear model, a mathematical function that describes the relationship between two variables

The model is usually applied when characterizing the relationship between two continuous variables, but as we will see, it can be adapted to model a wide range of variable types

We will be using multiple regression analysis, so I want to briefly review it with you.

Multiple Regression Analysis

The Linear Model

The general form of the linear model is

$$Y = \alpha + \beta X + \varepsilon$$

or, using sample notation,

$$Y = a + B X + e$$

If I think annual income is a linear function of the number of years of education people have, I would write the equation

$$\text{Income} = \alpha + \beta \text{ Education} + \varepsilon$$

The Linear Model

$$\text{Income} = \alpha + \beta\ \text{Education} + \varepsilon$$

There are three key parameters in the linear model

α = the intercept

β = the regression coefficient or the slope

ε = the error term, the residual term, or the disturbance term

Each parameter has a numerical value associated with it that is subject to meaningful interpretation

The Linear Model

$$\text{Income} = \alpha + \beta\ \text{Education} + \varepsilon$$

Suppose that I collect data on annual income and the number of years of education of people and analyze the data using the SPSS regression package. I might observe the following regression equation:

$$\text{Income} = 10{,}000 + 2{,}000\ \beta + e$$

What do these numbers mean?

The Linear Model

$Y = a + B X + e$

Income = 10,000 + 2,000 β + e

***a* is the predicted mean on Y when X = 0. In this case, it is 10,000.**

Intercepts are useful but they are rarely interpreted in social science research. I am not going to spend much time on them in this workshop

The Linear Model

Education	Mean Income	SD of Income
8	26,000	5,000
9	28,000	5,000
10	30,000	5,000
11	32,000	5,000
12	34,000	5,000

$$Y = a + B X + e$$

$$\text{Income} = 10{,}000 + 2{,}000\ \text{Education} + e$$

B tells us for every one unit that X increases, how much the mean on Y is predicted to change

The Linear Model

$$Y = \alpha + \beta X + \varepsilon$$

$$\text{Income} = \alpha + \beta\,\text{Education} + \varepsilon$$

The ε represent factors other than X that impact Y. They are often called residuals, errors, or disturbances.

The magnitude of the disturbances is often represented as 1 minus the squared correlation between X and Y.

The Linear Model

$$Y = \alpha + \beta X + \varepsilon$$

$$\text{Income} = \alpha + \beta\ \text{Education} + \varepsilon$$

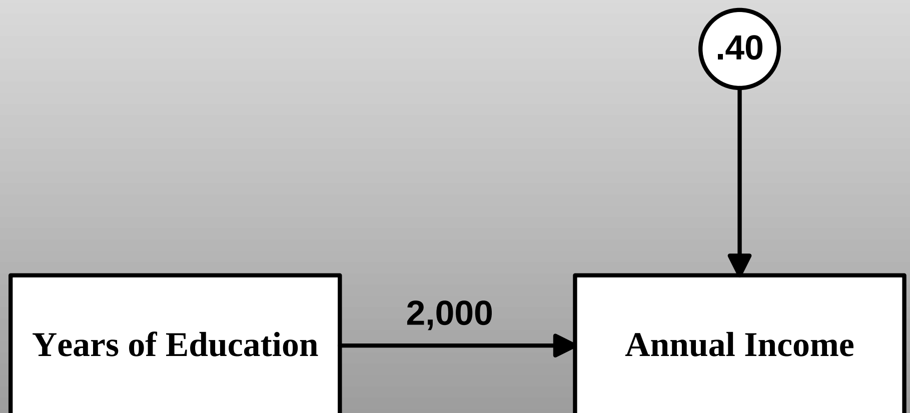
If I tell you the "disturbance value" is 0.30, then this means that X accounts for 70% of the variance in Y and 30% is due to other factors (or is "unexplained").

If I tell you the "disturbance value" is 0.80, then this means that X accounts for 20% of the variance in Y and 80% is due to other factors.

(terminology of eta squared)

The Linear Model

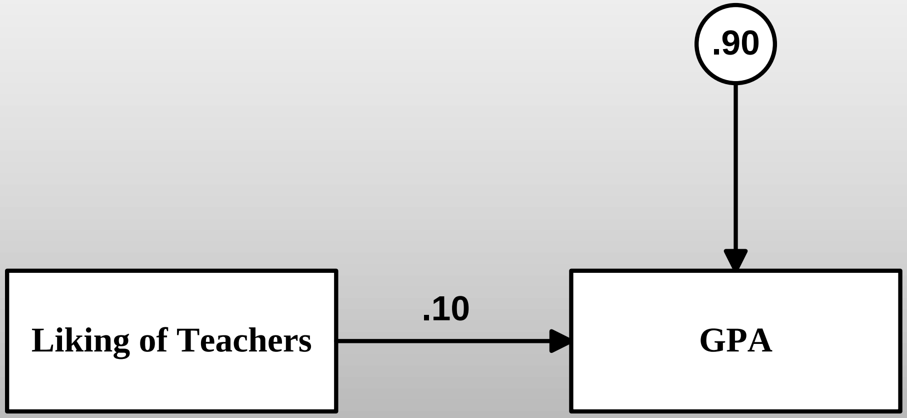
I can represent the key results of the linear analysis as applied to a set of data using path diagrams in the framework of causal modeling and it would look like this:



(note the absence of the intercept term)

The Linear Model

Here is another example:

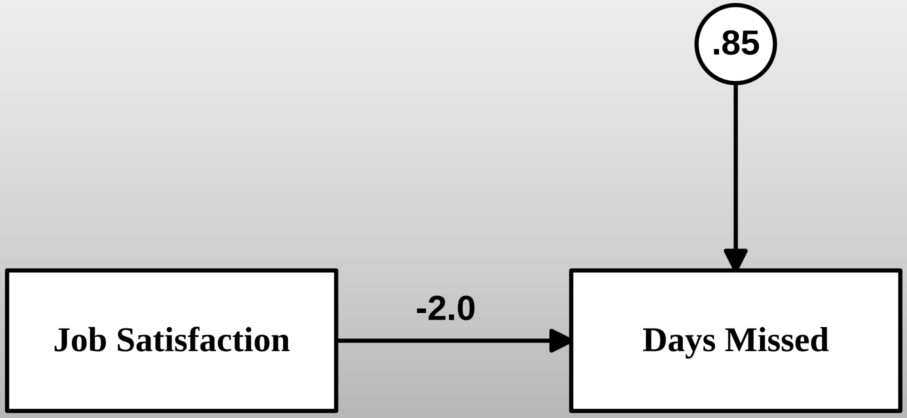


In general, how much do you like your teachers in your school?

1 = Very much dislike them
2 = Moderately dislike them
3 = Neither
4 = Moderately like them
5 = Very much like them

The Linear Model

And yet another example:

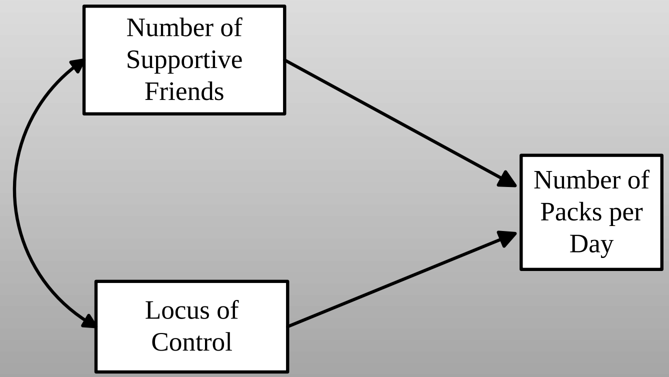


How satisfied are you with your job and the work you do for your job?

1 = Very dissatisfied
2 = Moderately dissatisfied
3 = Neither
4 = Moderately satisfied
5 = Very satisfied

Multiple Regression

We can expand the analysis to include multiple determinants or "predictors," and this is called *multiple regression analysis*



$$\textbf{Packs} = \alpha + \beta_1 \textbf{ Friends} + \beta_2 \textbf{ Control} + \varepsilon$$

$$Y = \alpha + \beta_1 X + \beta_2 Z + \varepsilon$$

Multiple Regression

$$\text{Packs} = \alpha + \beta_1 \text{ Friends} + \beta_2 \text{ Control} + \varepsilon$$

Suppose that I collect data on smoking, the number of supportive friends and locus of control and analyze the data using the SPSS regression package. I might observe the following regression equation:

$$\text{Packs} = 1.0 + -0.20 \text{ Friends} + -0.10 \text{ Control} + \varepsilon$$

What do these numbers mean?

Multiple Regression

Packs = α + β_1 Friends + β_2 Control + ε

Packs = 1.0 + -0.20 Friends + -0.10 Control + ε

A regression coefficient is the number of units that the mean of Y is predicted to change, given a one unit increase in the target predictor, *holding all other predictors constant*

Multiple Regression

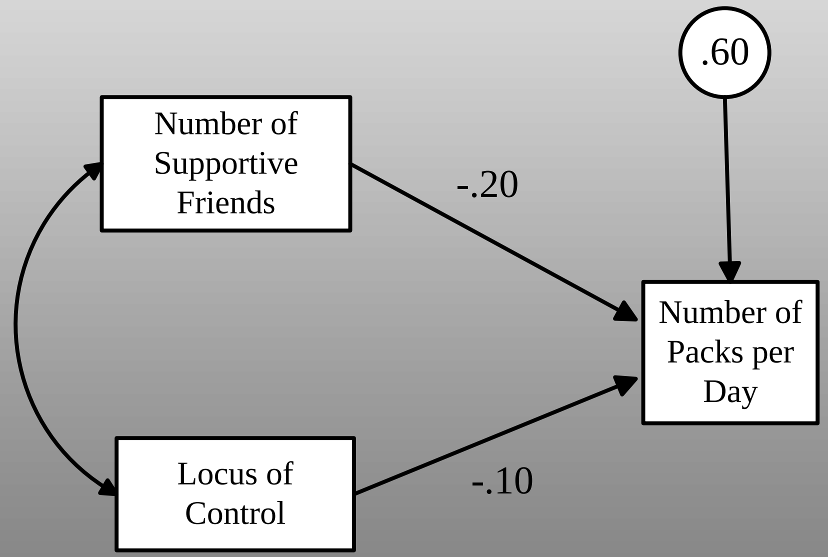
$$\text{Packs} = \alpha + \beta_1 \text{Friends} + \beta_2 \text{Control} + \varepsilon$$

$$\text{Packs} = 1.0 + -0.20 \text{ Friends} + -0.10 \text{ Control} + \varepsilon$$

R squared or eta squared is the proportion of variability in Y that all predictors, considered simultaneously, account for. The ‘disturbance value’ is the proportion of unexplained variance in Y

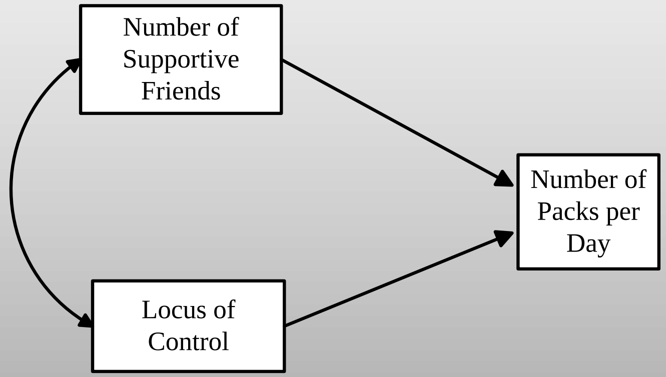
Multiple Regression

I can represent the key results of the linear analysis as applied to a set of data using path diagrams in the framework of causal modeling and it would look like this:



Representing Causal Models as Equations

There is a key concept I am using here. I draw my causal model:

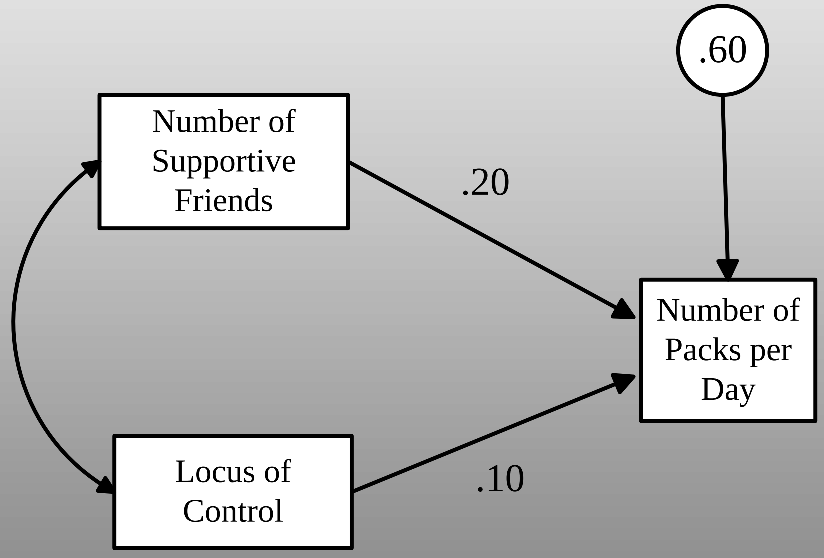


I then represent it as a "multiple regression" problem (i.e., as a linear equation):

$$\textbf{Packs} = \alpha + \beta_1 \textbf{ Friends} + \beta_2 \textbf{ Control} + \varepsilon$$

Representing Causal Models as Equations

Then I use SPSS or SAS to estimate key coefficients in the model using standard multiple regression methods:

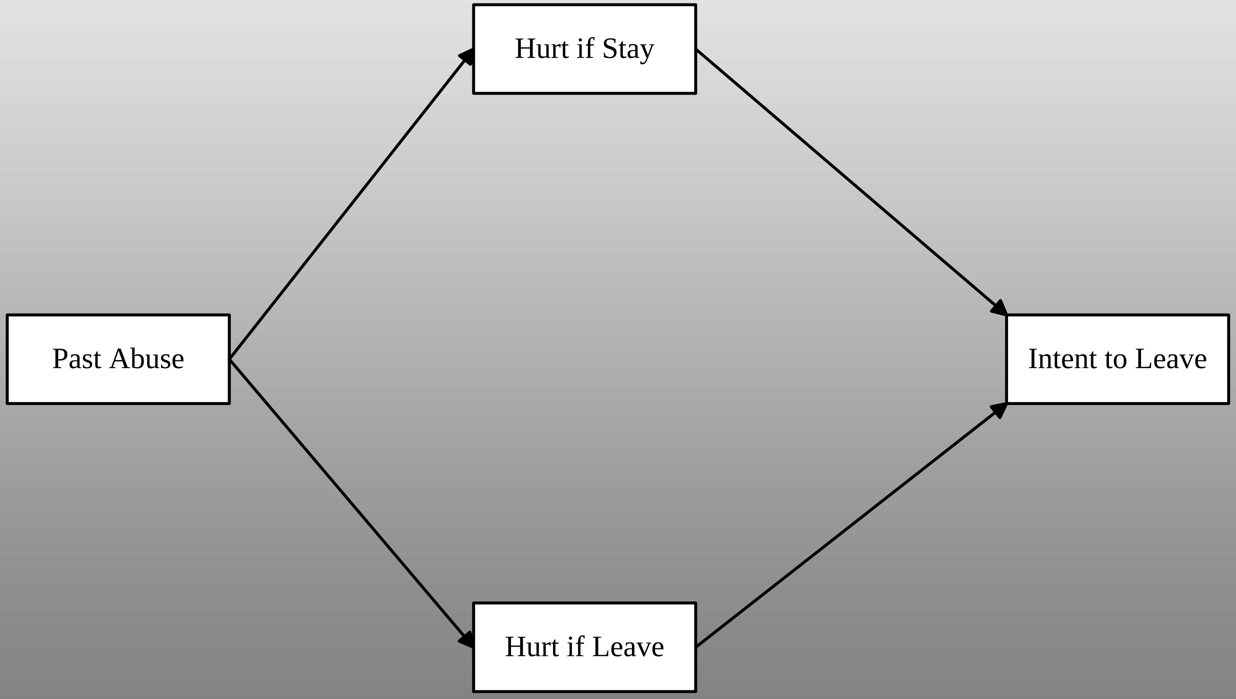


and interpret the coefficients accordingly

Limited Information SEM

Limited Information SEM

To analyze a causal model, we first draw a path diagram of that model:



Abuse Example

<u>Intention to leave relationship in the next 3 months</u>

I intend to leave this relationship in the next 3 months

1 = strongly disagree
2 = moderately disagree
3 = neither
4 = moderately agree
5 = strongly agree

<u>Physically hurt if try to leave</u>

If I try to leave the relationship in the next 3 months, he will physically hurt me

1 = strongly disagree
2 = moderately disagree
3 = neither
4 = moderately agree
5 = strongly agree

Abuse Example

Physically hurt if stay

If I stay in this relationship for the next 3 months, he will probably physically hurt me

1 = strongly disagree
2 = moderately disagree
3 = neither
4 = moderately agree
5 = strongly agree

Past Abuse

How many times over the past 3 months has he physically hurt you?

____ number of times

Abuse Example

We identify endogenous and exogenous variables in the system.

***Endogenous variable* is any variable with a causal arrow going to it**

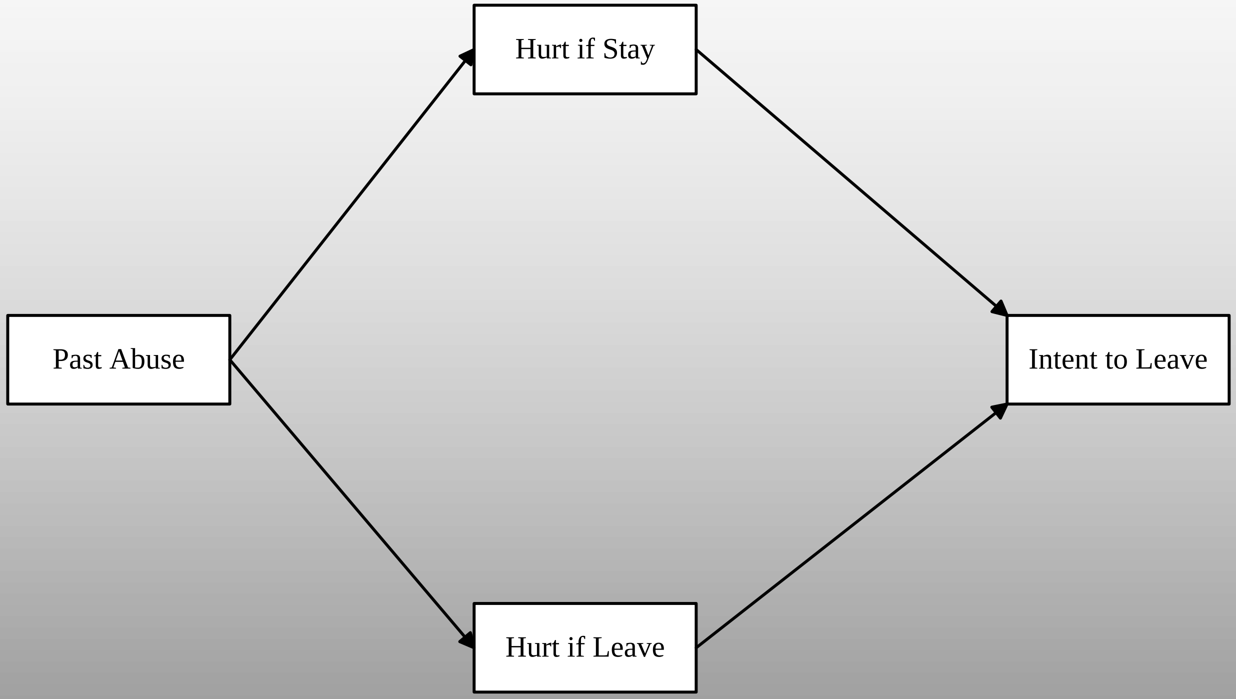
***Exogenous variable* is any variable with no causal arrow going to it**

Next, we translate the model into a set of linear equations

There is one equation for each endogenous variable

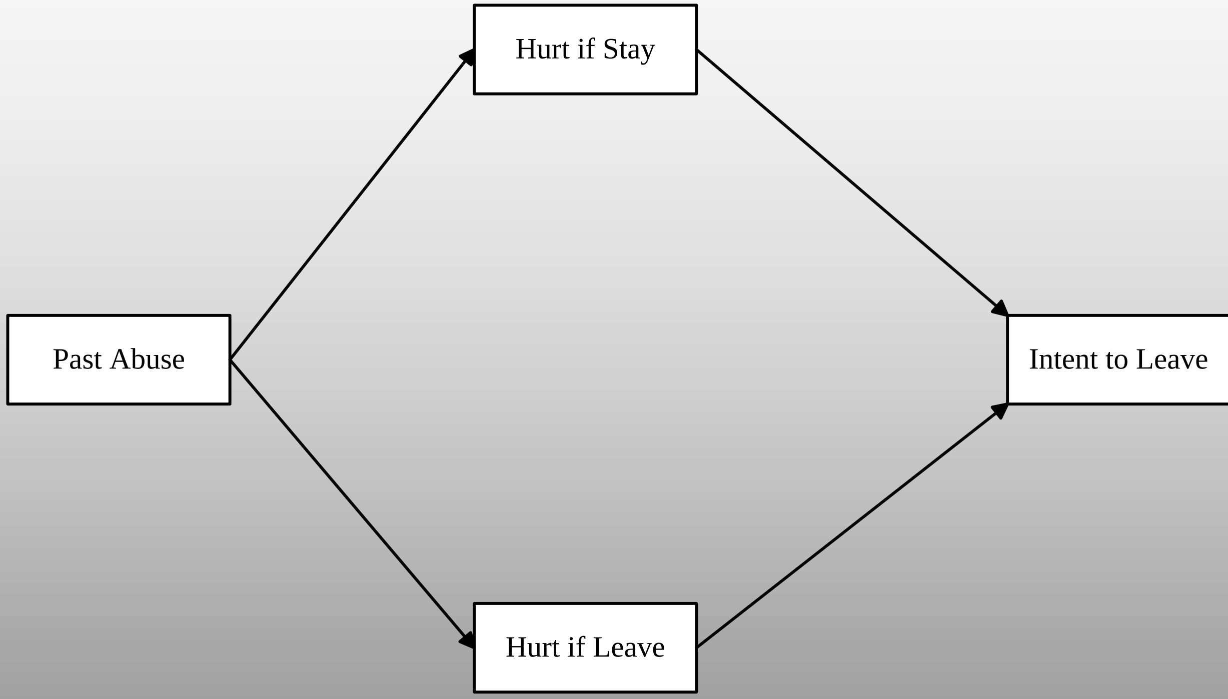
The Y variable is the endogenous variable. The Xs are all variables that have an arrow going directly to Y

Abuse Example



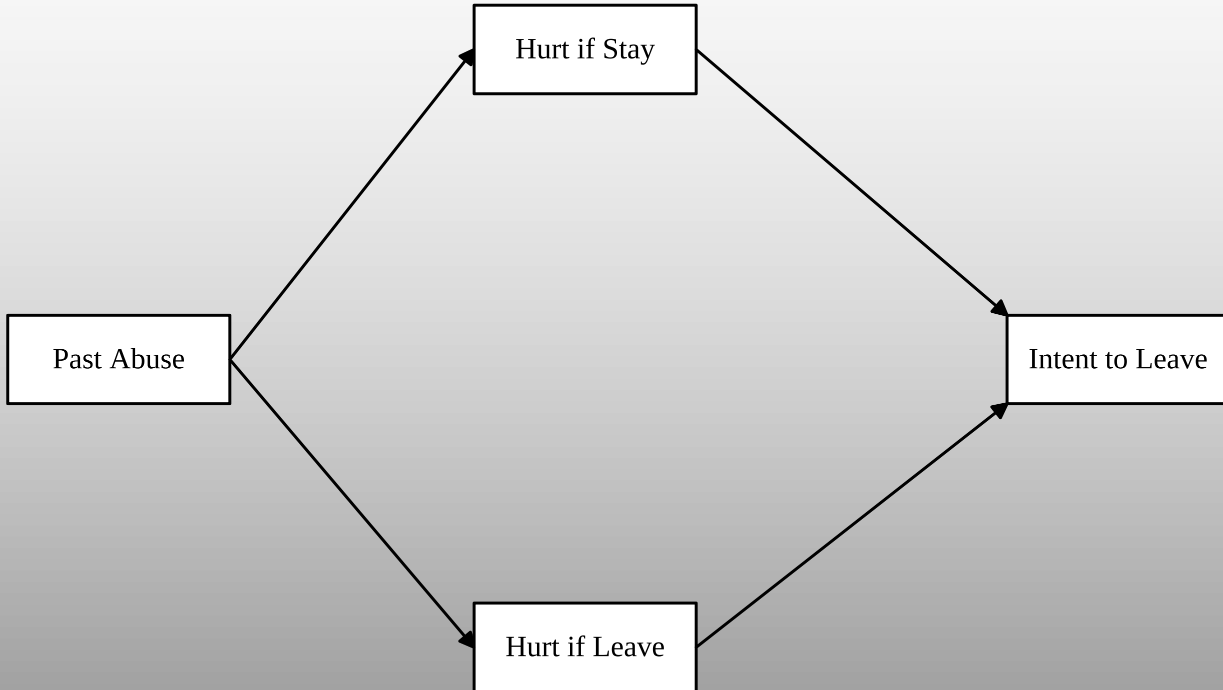
$$\text{Intent} = \alpha_1 + \beta_1 \text{Stay} + \beta_2 \text{Leave} + \varepsilon_3$$

Abuse Example



$$\textbf{Stay} = \alpha_2 + \beta_3\ \textbf{Abuse} + \varepsilon_1$$

Abuse Example



$$\textbf{Leave} = \alpha_3 + \beta_4\,\textbf{Abuse} + \varepsilon_2$$

Abuse Example

The equations are:

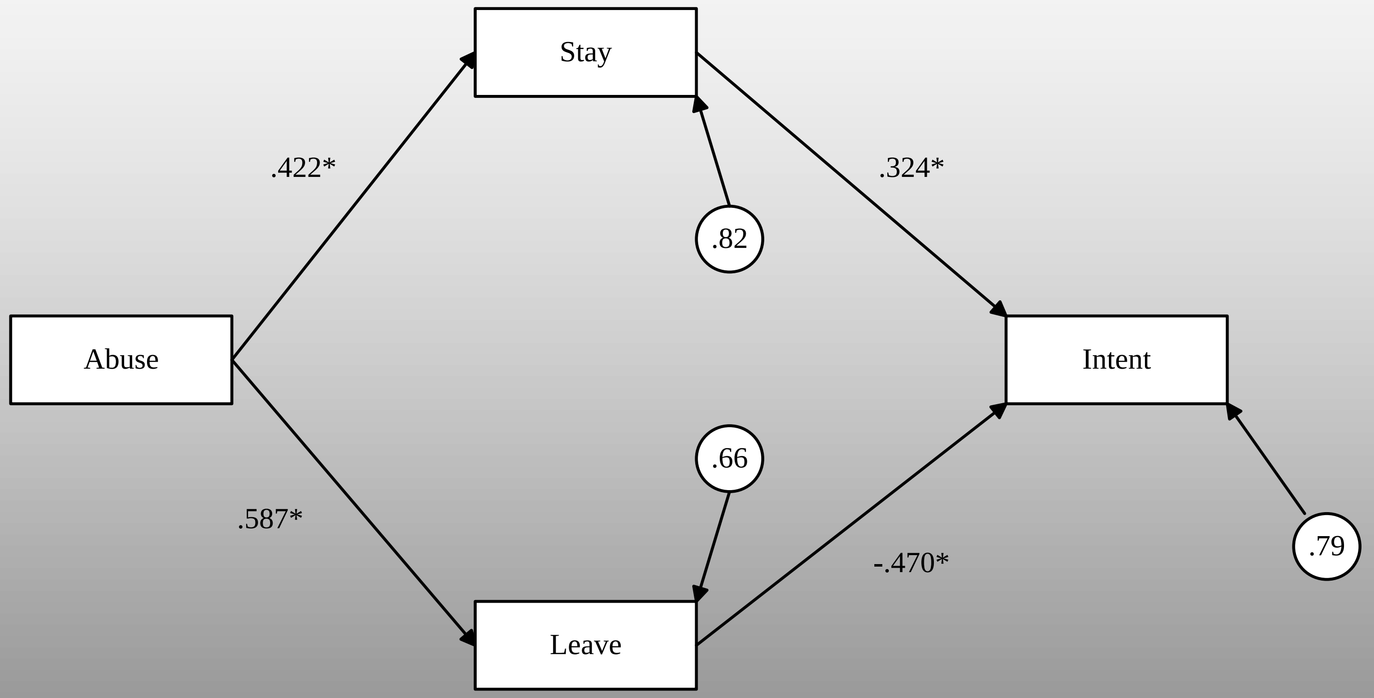
$$\text{Intent} = \alpha_1 + \beta_1 \text{ Stay} + \beta_2 \text{ Leave} + \varepsilon_3$$

$$\text{Stay} = \alpha_2 + \beta_3 \text{ Abuse} + \varepsilon_1$$

$$\text{Leave} = \alpha_3 + \beta_4 \text{ Abuse} + \varepsilon_2$$

I can then estimate all the parameters using the SPSS multiple regression program, one equation at a time

Abuse Example



Abuse Example

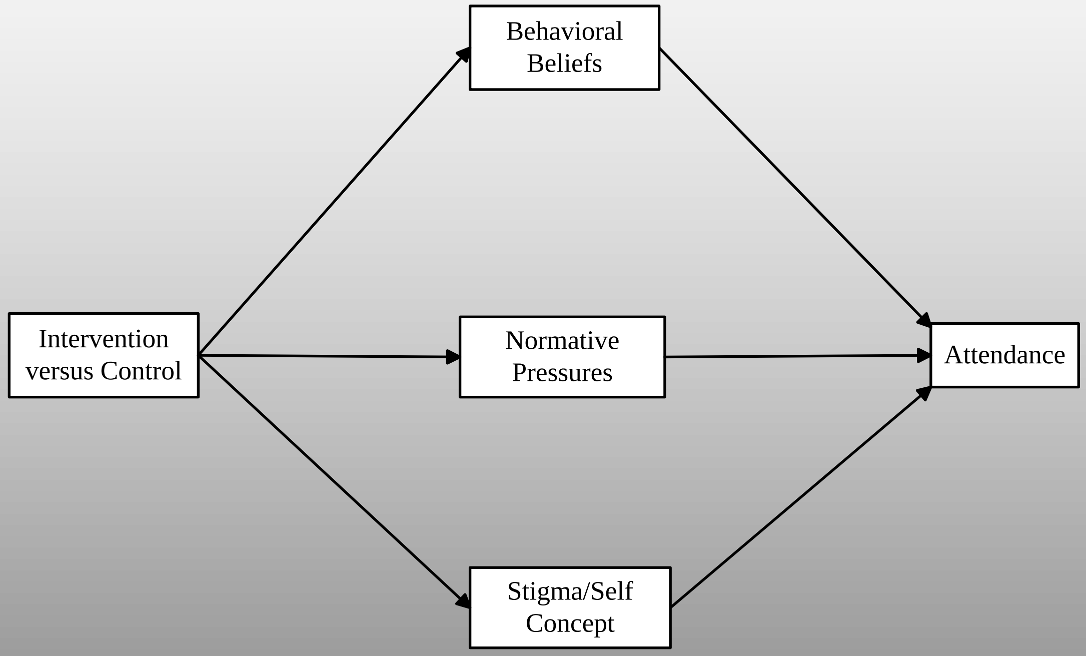
$$\text{Intent} = \alpha_1 + \beta_1 \text{ Stay} + \beta_2 \text{ Leave} + \varepsilon_3$$

$$\text{Stay} = \alpha_2 + \beta_3 \text{ Abuse} + \varepsilon_1$$

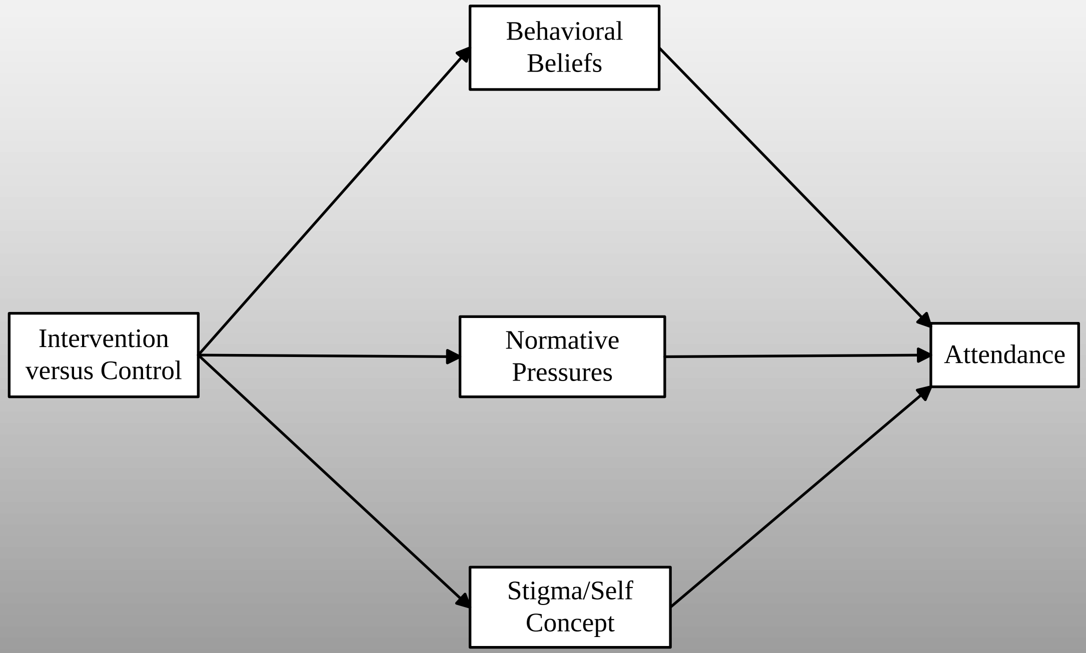
$$\text{Leave} = \alpha_3 + \beta_4 \text{ Abuse} + \varepsilon_2$$

Full information estimation estimates all coefficients simultaneously in one step. Limited information estimation chops the model up and estimates each equation separately

Intervention Example

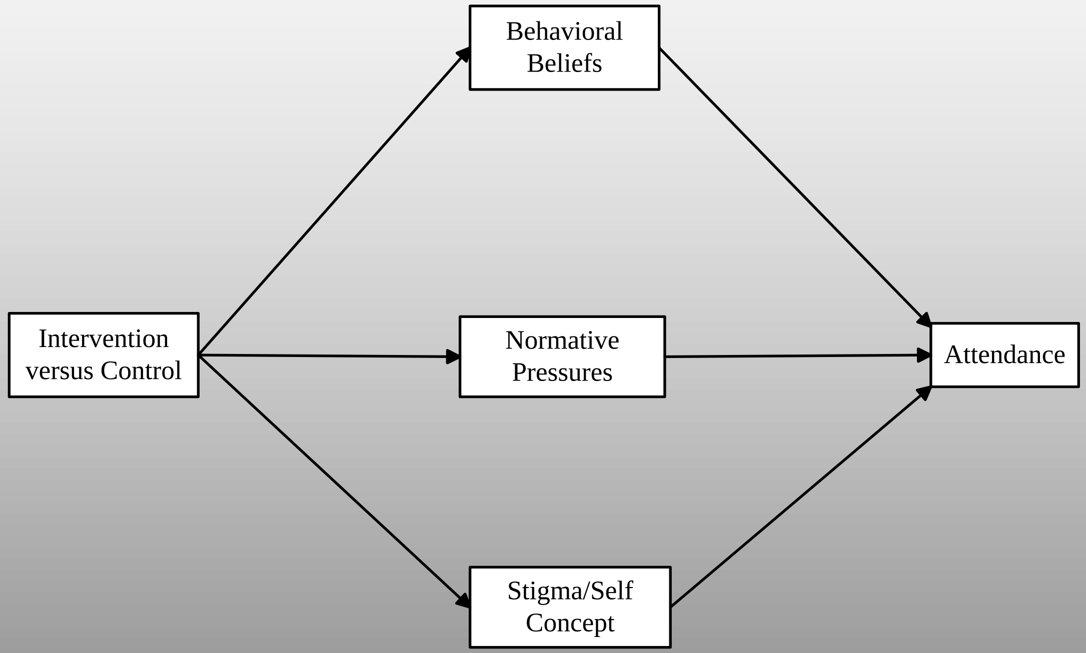


Intervention Example



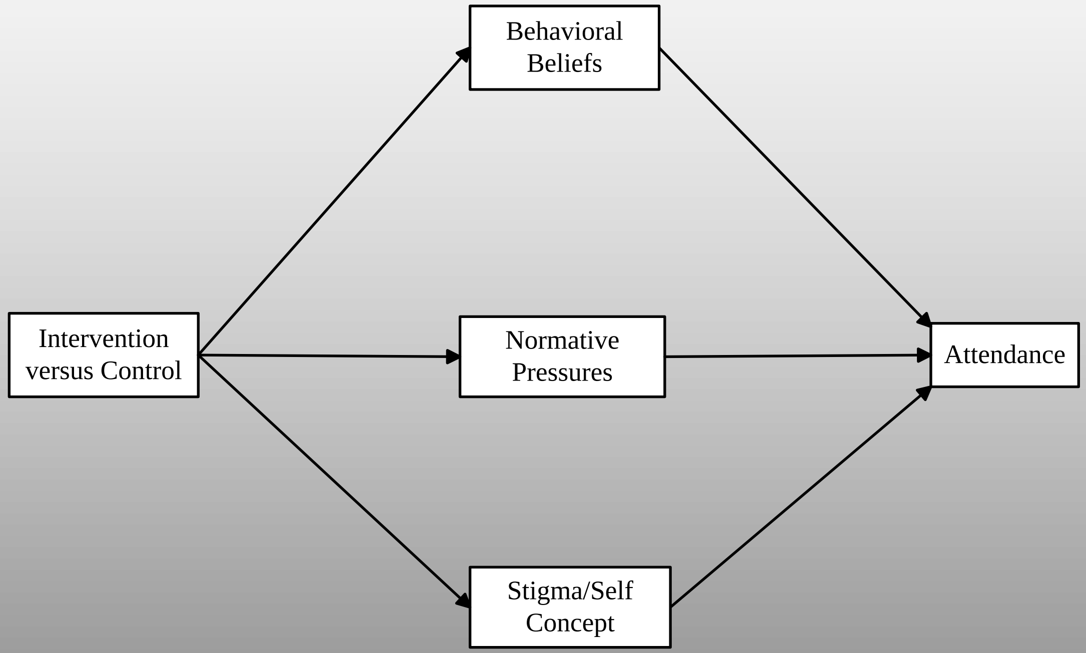
$$\textbf{Attendance} = \alpha_1 + \beta_1 \textbf{ Beliefs} + \beta_2 \textbf{ Norms} + \beta_3 \textbf{ Stigma} + \varepsilon_4$$

Intervention Example



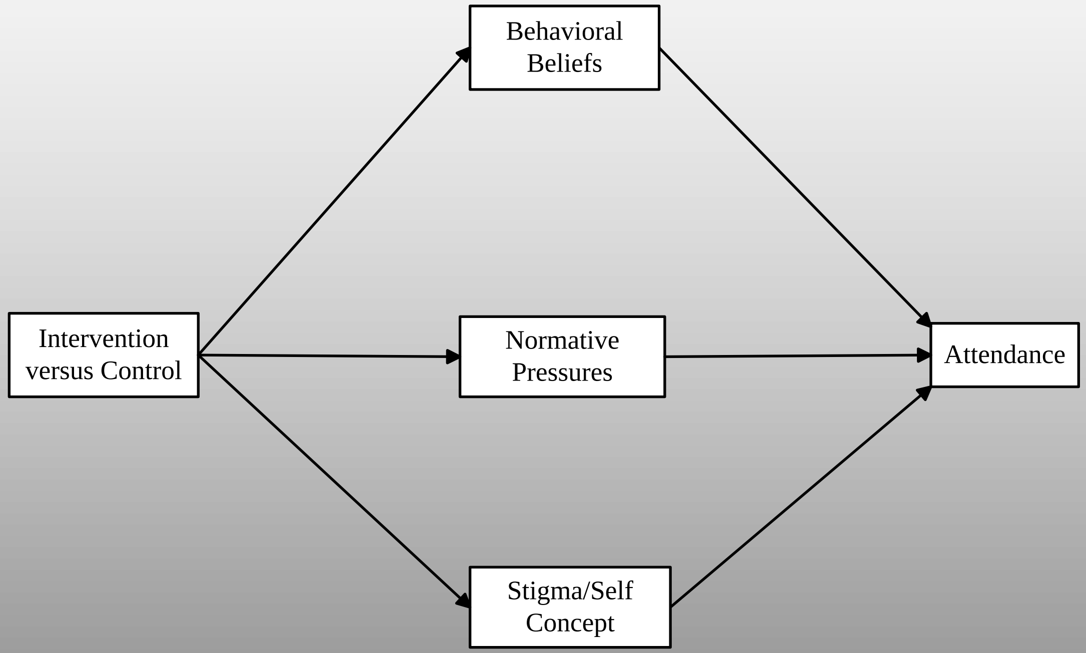
$$\textbf{Beliefs} = \alpha_2 + \beta_4 \,\textbf{Intervention} + \varepsilon_1$$

Intervention Example



$$\text{Norms} = \alpha_3 + \beta_5\,\text{Intervention} + \varepsilon_2$$

Intervention Example



$$\textbf{Stigma} = \alpha_4 + \beta_6\ \textbf{Intervention} + \varepsilon_3$$

Intervention Example

Putting them all together, I get:

$$\text{Attendance} = \alpha_1 + \beta_1 \text{Beliefs} + \beta_2 \text{Norms} + \beta_3 \text{Stigma} + \varepsilon_4$$

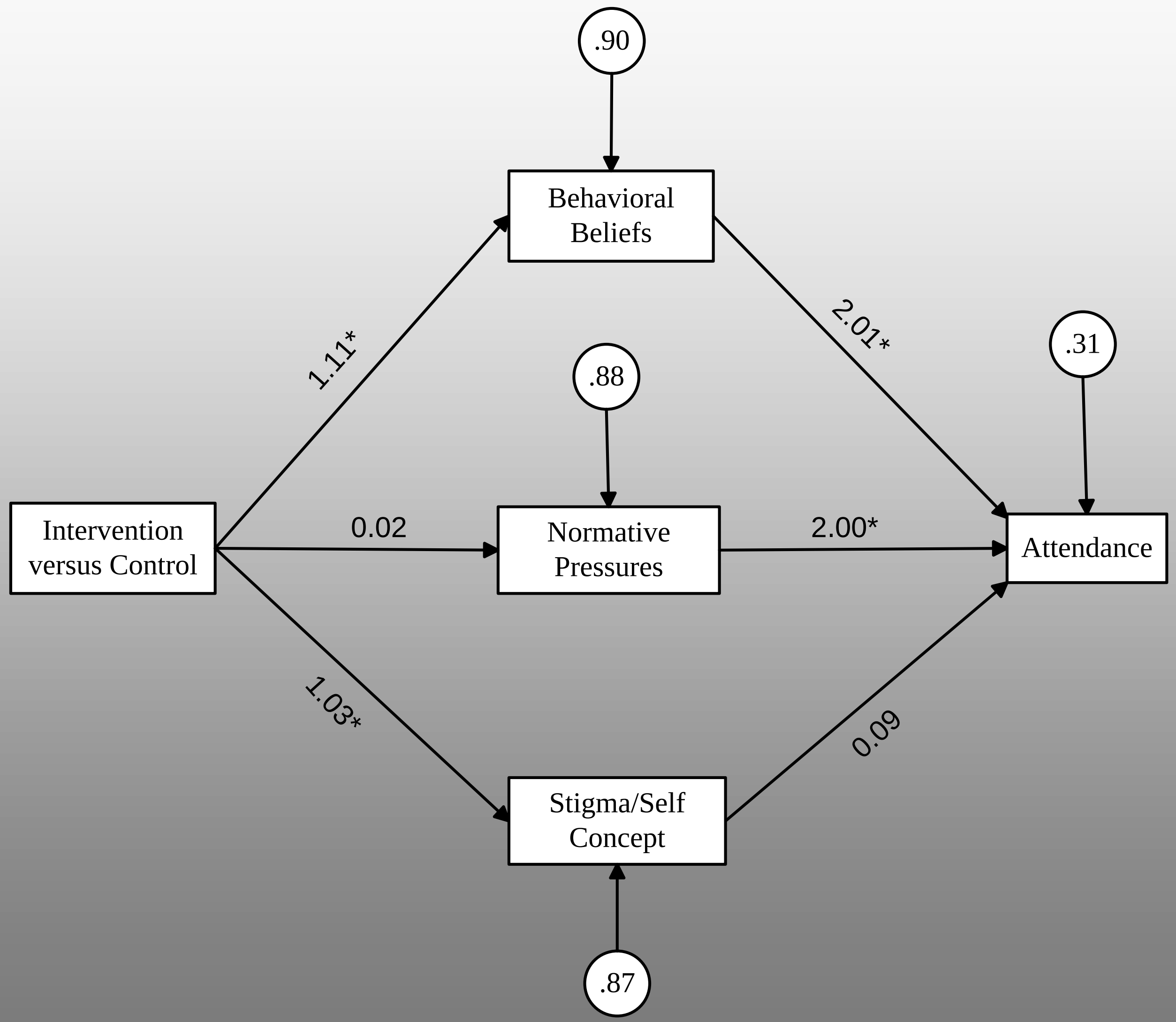
$$\text{Beliefs} = \alpha_2 + \beta_4 \text{Intervention} + \varepsilon_1$$

$$\text{Norms} = \alpha_3 + \beta_5 \text{Intervention} + \varepsilon_2$$

$$\text{Stigma} = \alpha_4 + \beta_6 \text{Intervention} + \varepsilon_3$$

I can then estimate all the parameters using the SPSS multiple regression program, working on one equation at a time

Intervention Example



Example Write-Up: Abuse Example

Method of Analysis

The model in Figure 1 was tested using a limited information estimation strategy in conjunction with ordinary least squares (OLS) regression as applied to each of the three linear equations implied by the model. Unstandardized regression coefficients are reported in conjunction with standardized disturbance terms reflecting the proportion of unexplained variance for each endogenous variable.

Results

Figure 1 presents the results of the three regression analyses in the form of unstandardized path coefficients and the proportion of unexplained variance for each endogenous variable. As predicted, all four of the path coefficients were statistically significant ($p < 0.05$). The model is consistent with an impact of physical abuse on both the belief that staying in the relationship will result in future physical abuse and the belief that trying to leave the relationship will result in physical violence. For every one additional reported act of physical abuse in the couple's history over the course of the past 3 months, the mean belief that trying to

Example Write-Up

leave the relationship will result in physical violence is predicted to increase by 0.37 rating scale units. History of physical abuse accounts for 34% of the variance in this belief and the associated standard error of estimate associated with the disturbance term was 0.41. Similarly, for every one additional reported act of physical abuse in the couple's history over the course of the past 3 months, the mean belief that staying in the relationship will result in future physical abuse is predicted to increase by 0.25 rating scale units. History of physical abuse accounts for 18% of the variance in this belief and the associated standard error of estimate with the disturbance term was 0.52.

The beliefs associated with being harmed if one stays or tries to leave the relationship were both statistically significantly ($p < 0.05$) related to the intent to leave the relationship, accounting for 21% of the variation in intentions (the associated standard error of estimate was 0.44). For every one unit that the belief of being hurt if one stays in the relationship increased, the mean intent to leave the relationship increased by 0.32 rating scale units. For every one unit that the belief of being hurt if one tries to leave the relationship increased, the mean intent to leave the relationship decreased by 0.47 rating scale units.

Example Write-Up

Both of the beliefs in the model have the role of being mediators of the relationship between past abuse and the intent to leave the relationship in the next three months, as reflected by the joint significance test (MacKinnon et al., 2008).

Covariates

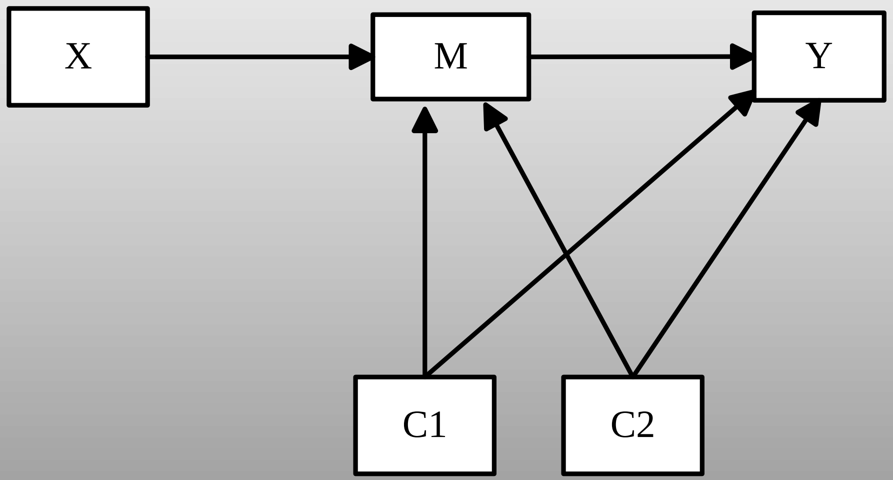
Covariates

We often have variables that are not part of our formal theory but that we want to control for methodological reasons.

We omit these variables from our formal presentations but we include them in the analysis, as per traditional covariate control in multiple regression

Traditionally, one controls for the covariates for each endogenous variable in the system, but this can be over-ridden by theoretical considerations

Covariates



Limited Information SEM

Some General Advice

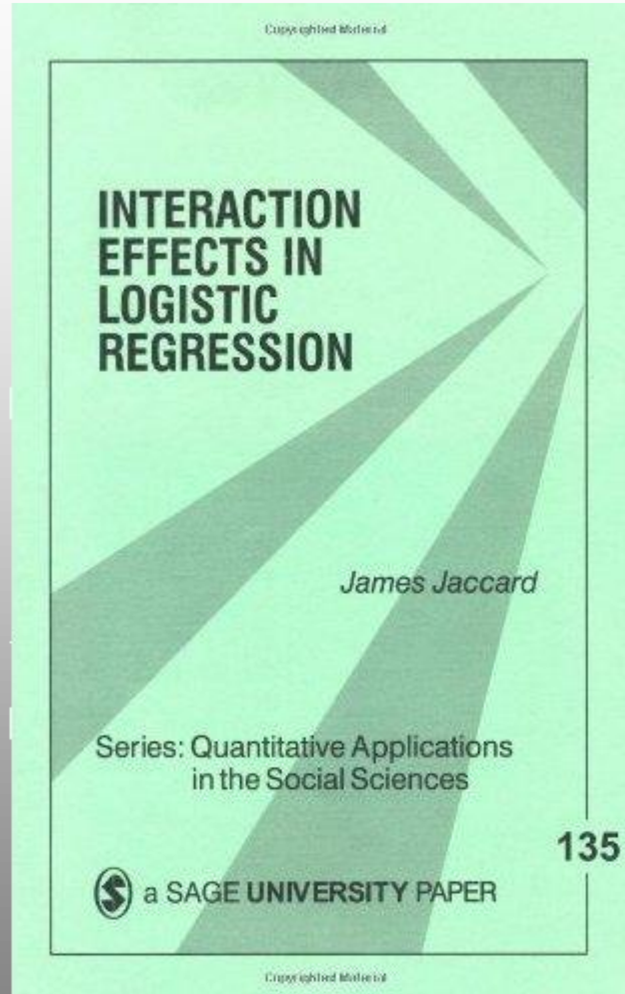
Avoid hierarchical regression

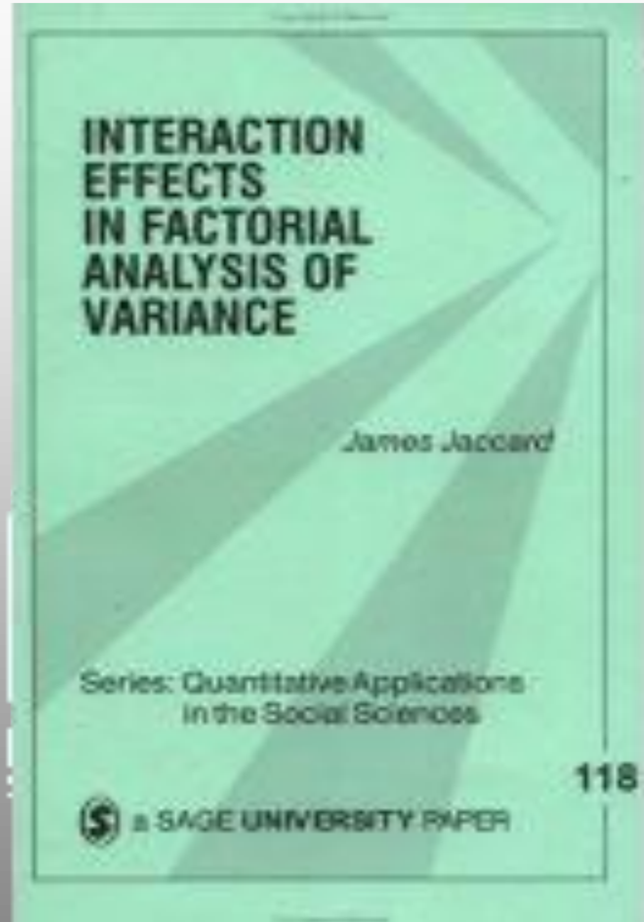
Be wary of standardized regression coefficients

(Use averaging instead of summation of items)

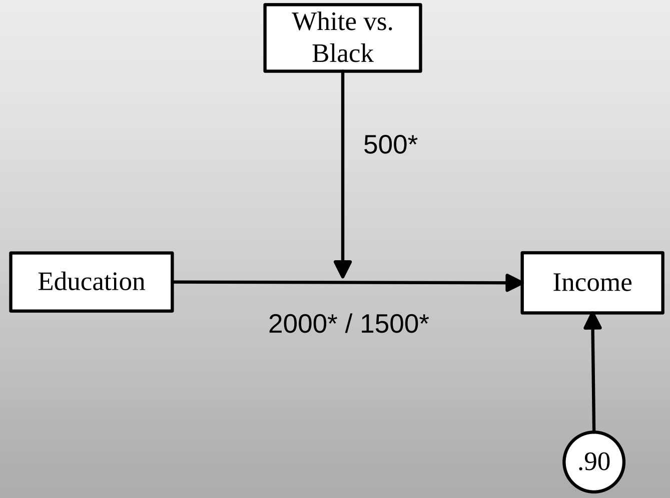
Moderator Analysis


INTERACTION EFFECTS IN MULTIPLE REGRESSION
SECOND EDITION
James Jaccard
Robert Turrisi
Series: Quantitative Applications in the Social Sciences
72
a SAGE UNIVERSITY PAPER


Copyrighted Material
INTERACTION EFFECTS IN LOGISTIC REGRESSION
James Jaccard
Series: Quantitative Applications in the Social Sciences
135
a SAGE UNIVERSITY PAPER
Copyrighted Material


INTERACTION
EFFECTS
IN FACTORIAL
ANALYSIS OF
VARIANCE
James Jaccard
Series: Quantitative Applications
in the Social Sciences
118
a SAGE UNIVERSITY PAPER

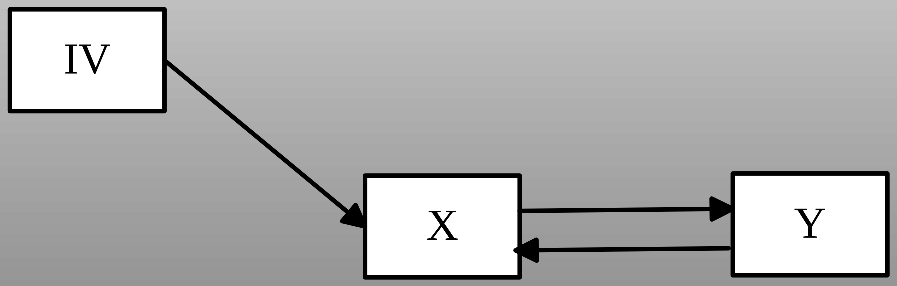
Moderator Analysis



Reciprocal Causation

Uses two stage least squares regression (TSLS) or maximum likelihood methods

Requires instrumental variables



Full Information versus Limited Information Estimation

Full Information versus Limited Information

Advantages of Full Information Estimation

Parameter estimates are more efficient

Can handle complex error structures more elegantly

Yields a more complete set of model fit indices

Can deal with latent variables and measurement error

Full Information versus Limited Information

<u>Advantages of Limited Information</u>

Less affected by model misspecification

More flexible in terms of mixing analytic methods (e.g., count regression, logistic regression, robust regression, ordinal regression, quantile regression, non-linear regression)

Less sample size demanding

Uses familiar methods across disciplines (e.g., dummy variables, polynomial regression)

Some References

Jaccard, J., Guilamo-Ramos, V. Johansson, M. & Bouris, A. (2006). Multiple regression analyses in clinical child and adolescent psychology. Journal of Clinical Child and Adolescent Psychology, 35, 456-479.

MacKinnon,D. P., Lockwood, C. M., Hoffman, J. M.,West, S. G., & Sheets, V. (2002).A comparison of methods to test the significance of the mediated effect. Psychological Methods, 7(1), 83–104.

Bollen, K.A. (2001). Two-stage least squares and latent variable models: Simultaneous estimation and robustness to misspecifications. In R. Cudeck, S. Du Toit, & D. Sörbom (Eds.), Structural equation modeling: Present and future (pp. 119–138). Lincolnwood, IL: Scientific Software.

Bollen, K.A. (1996). A limited information estimator for LISREL models with and without heteroscedasticity. In G.A. Marcoulides, & R.E. Schumacker (Eds.), Advanced structural equation modeling (pp. 227–241). Mahwah, NJ: Erlbaum.