Structural Equation Modeling

Intro to SEM

Psy 524

Ainsworth

AKA

- SEM – Structural Equation Modeling
- CSA – Covariance Structure Analysis
- Causal Models
- Simultaneous Equations
- Path Analysis
- Confirmatory Factor Analysis

SEM in a nutshell

- Combination of factor analysis and regression
  - Continuous and discrete predictors and outcomes
  - Relationships among measured or latent variables
- Direct link between *Path Diagrams* and equations and fit statistics
- Models contain both measurement and path models

Jargon

- Measured variable
  - Observed variables, indicators or manifest variables in an SEM design
  - Predictors and outcomes in path analysis
  - Squares in the diagram
- Latent Variable
  - Un-observable variable in the model, factor, construct
  - Construct driving measured variables in the measurement model
  - Circles in the diagram

Jargon

- Error or E
  - Variance left over after prediction of a measured variable
- Disturbance or D
  - Variance left over after prediction of a factor
- Exogenous Variable
  - Variable that predicts other variables
- Endogenous Variables
  - A variable that is predicted by another variable
  - A predicted variable is endogenous even if it in turn predicts another variable

Jargon

- Measurement Model
  - The part of the model that relates indicators to latent factors
  - The measurement model is the factor analytic part of SEM
- Path model
  - This is the part of the model that relates variable or factors to one another (prediction)
  - If no factors are in the model then only path model exists between indicators

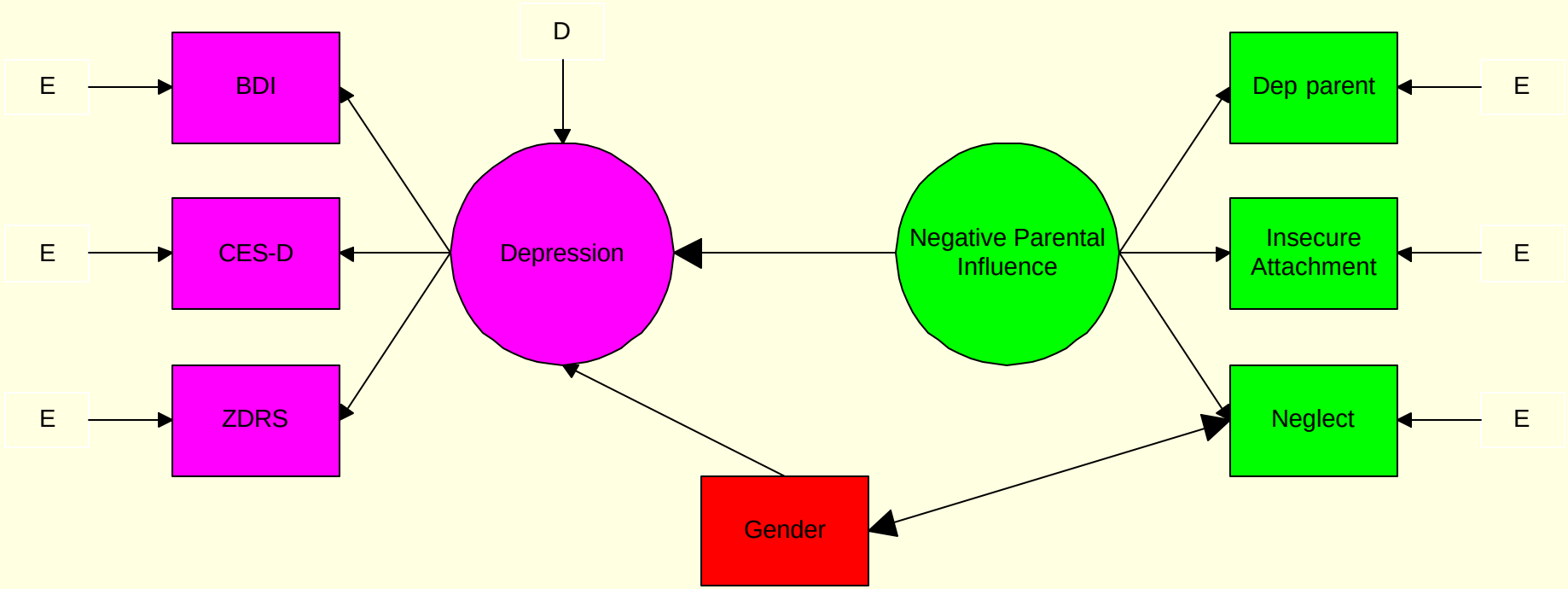
Jargon

- Direct Effect
  - Regression coefficients of direct prediction
- Indirect Effect
  - Mediating effect of x1 on y through x2
- Confirmatory Factor Analysis
- Covariance Structure
  - Relationships based on variance and covariance
- Mean Structure
  - Includes means (intercepts) into the model

Diagram elements

- Single-headed arrow ?
  - This is prediction
  - Regression Coefficient or factor loading
- Double headed arrow ?
  - This is correlation
- Missing Paths
  - Hypothesized absence of relationship
  - Can also set path to zero

Path Diagram



SEM questions

- Does the model produce an estimated population covariance matrix that “fits” the sample data?
  - SEM calculates many indices of fit; close fit, absolute fit, etc.
- Which model best fits the data?
- What is the percent of variance in the variables explained by the factors?
- What is the reliability of the indicators?
- What are the parameter estimates from the model?

SEM questions

- Are there any indirect or mediating effects in the model?
- Are there group differences?
  - Multigroup models
- Can change in the variance (or mean) be tracked over time?
  - Growth Curve or Latent Growth Curve Analysis

SEM questions

- Can a model be estimated with individual and group level components?
  - Multilevel Models
- Can latent categorical variables be estimated?
  - Mixture models
- Can a latent group membership be estimated from continuous and discrete variables?
  - Latent Class Analysis

SEM questions

- Can we predict the rate at which people will drop out of a study or end treatment?
  - Discrete-time survival mixture analysis
- Can these techniques be combined into a huge mess?
  - Multiple group multilevel growth curve latent class analysis???????

SEM limitations

- SEM is a confirmatory approach
  - You need to have established theory about the relationships
  - Cannot be used to explore possible relationships when you have more than a handful of variables
  - Exploratory methods (e.g. model modification) can be used on top of the original theory
  - SEM is not causal; experimental design = cause

SEM limitations

- SEM is often thought of as strictly correlational but can be used (like regression) with experimental data if you know how to use it.
  - Mediation and manipulation can be tested
- SEM is by far a very fancy technique but this does not make up for a bad experiment and the data can only be generalized to the population at hand

SEM limitations

- Biggest limitation is sample size
  - It needs to be large to get stable estimates of the covariances/correlations
  - 200 subjects for small to medium sized model
  - A minimum of 10 subjects per estimated parameter
  - Also affected by effect size and required power

SEM limitations

- Missing data
  - Can be dealt with in the typical ways (e.g. regression, EM algorithm, etc.) through SPSS and data screening
  - Most SEM programs will estimate missing data and run the model simultaneously
- Multivariate Normality and no outliers
  - Screen for univariate and multivariate outliers
  - SEM programs have tests for multi-normality
  - SEM programs have corrected estimators when there's a violation

SEM limitations

- Linearity
- No multicollinearity/singularity
- Residuals Covariances (R minus reproduced R)
  - Should be small
  - Centered around zero
  - Symmetric distribution of errors
  - If asymmetric than some covariances are being estimated better than others