PSY792F SEM

Week 13 – Multilevel SEM

Mark A. Prince, PhD, MS

Advantages of SEM

- Represents hypotheses as a system of simultaneous equations
- Permits the estimation of direct and indirect regression effects
- Tests the entire model as a global hypothesis
- Separates error variance from reliable variance, automatically correcting effects for attenuation due to unreliability
- *Most statistical methods are special cases of SEM

Limitations of (traditional) SEM

- It is difficult to handle clustered (i.e., hierarchically nested) data
- Nesting violates core assumptions of SEM – making estimates untrustworthy
  - Especially the independence assumption

Advantages of MLM

- Handles an arbitrary number of nesting levels
- Partitions variance in the dependent variable to account for nesting
- Flexibly handles predictor variables at any level of measurement

Limitations of (traditional) MLM

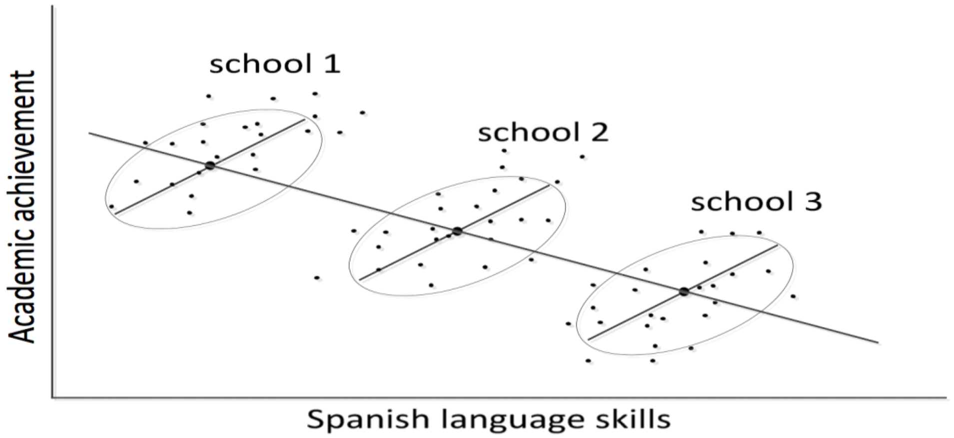
- Multivariate models are possible, but difficult to specify
- Latent variables are (nearly) impossible to include, and require all factor loadings = 1
  - Ignores measurement error
- Model fit indices are not available
  - Only a few model selection criteria
- The within and between effects are often conflated, and if steps are taken to separate them, bias is introduced
- Level two outcomes are not an option

Multilevel SEM (MSEM)

- The combination of SEM and MLM
- Not a single method, but a collection of methods designed to permit SEM-like modeling with clustered/multilevel data
- Allows us to view SEM and MLM as a special case of a more general framework
- Most critical limitations MSEM overcomes
  - SEM does not consider clustering
  - MLM assumes variables are measured without error

Gaining additional knowledge

- MLM is often used without decomposing the effects of a level-1 predictor into within and between effects.
  - Then, the level-1 slope is a weighted average of the between-cluster and within-cluster effects of the level-1 predictor
  - If these effects are different the research will be ignorant to the level-specific effects
- For example: if the within effect is negative and the between effect is positive the weighted average will be near 0
- Within schools, students with better Spanish skills had higher academic achievement. Yet, schools with the highest proportion of Spanish speakers performed poorest.



Two more examples

- Heart Attacks & Exercise:
  - Within persons
    - risk of heart attack increases with exercise
  - Between persons
    - As average exercise increases, overall risk of heart attack is lower

- Income & Crime
  - Between level (neighborhood level)
    - Higher crime is associated with lower income
  - Within level (individuals)
    - Within a neighborhood those who engage in crime may have higher income (e.g., a drug dealer in a poor neighborhood)

MSEM centering

- In MLM, to obtain within and between effects, we need to manually group mean center in a data management step, and include group means in the model specification
- In MSEM, group mean centering is done by default in a model-based way, rather than in a data management step
  - MSEM uses random intercepts (i.e., latent cluster means) in lieu of manually computed centered means
  - Thus, you do not need to center in MSEM

MSEM History

- First proposed in William Schmidt's (1969) dissertation as a way to combine SEM and MLM
- Since there have been at least 18 methods for combining SEM and MLM
  - Methods vary along dimensions of speed/efficiency, generality, flexibility, ease of implementation, software availability
- Mplus – Muthen and Asparouhov (2009)
  - Very general model of 2-level SEM
  - Developed by Muthen starting in the late 1980s
  - Accommodates random intercepts and random slopes
  - Accommodates missing data and unbalanced cluster sizes
  - Provides overall and level-specific fit indices (for fixed slope models)
  - Accommodates 2 or 3 levels
  - Separates between and within by default
  - Easily handles continuous, count, ordinal, binary data
  - Computationally efficient
  - Uses MLR which does not require normality
  - Yields robust asymptotic covariances of parameter estimates and chi-square
  - Can incorporate discrete latent variables (i.e., mixtures)
  - Allows both frequentist or Bayesian Estimation

SEM Equations

$$\mathbf{Y}_i = \boldsymbol{\nu} + \boldsymbol{\Lambda}\boldsymbol{\eta}_i + \mathbf{K}\mathbf{X}_i + \boldsymbol{\varepsilon}_i$$ ← measurement model

$$\boldsymbol{\eta}_i = \boldsymbol{\alpha} + \mathbf{B}\boldsymbol{\eta}_i + \boldsymbol{\Gamma}\mathbf{X}_i + \boldsymbol{\zeta}_i$$ ← structural model

$$\boldsymbol{\varepsilon}_i \sim MVN(\mathbf{0}, \boldsymbol{\Theta})$$

$$\boldsymbol{\zeta}_i \sim MVN(\mathbf{0}, \boldsymbol{\Psi})$$

MSEM Equations

MSEM begins with the single-level equations from earlier and simply adds *j* subscripts to denote that contents of these matrices may vary across clusters:

$$\mathbf{Y}_{ij} = \boldsymbol{\nu}_j + \boldsymbol{\Lambda}_j \boldsymbol{\eta}_{ij} + \mathbf{K}_j \mathbf{X}_{ij} + \boldsymbol{\varepsilon}_{ij}$$ ← Level-1 measurement model

$$\boldsymbol{\eta}_{ij} = \boldsymbol{\alpha}_j + \mathbf{B}_j \boldsymbol{\eta}_{ij} + \boldsymbol{\Gamma}_j \mathbf{X}_{ij} + \boldsymbol{\zeta}_{ij}$$ ← Level-1 structural model

$$\boldsymbol{\eta}_j = \boldsymbol{\mu} + \boldsymbol{\beta}\boldsymbol{\eta}_j + \boldsymbol{\gamma}\mathbf{X}_j + \boldsymbol{\zeta}_j$$ ← Level-2 structural model (NEW!)

$\boldsymbol{\eta}_j$ is an $(s \times 1)$ vector of level-2 random coefficients

$\boldsymbol{\mu}$ is an $(s \times 1)$ vector of means of level-2 random coefficients

$\boldsymbol{\beta}$ is an $(s \times s)$ matrix of level-2 structural regression slopes

$\boldsymbol{\gamma}$ contains slopes for level-2 exogenous covariates in the vector $\mathbf{X}_j$

$\boldsymbol{\zeta}_j$ is a vector of level-2 error terms (random effects)

MSEM equations continued

$$\mathbf{Y}_{ij} = \boldsymbol{\nu}_j + \boldsymbol{\Lambda}_j \boldsymbol{\eta}_{ij} + \mathbf{K}_j \mathbf{X}_{ij} + \boldsymbol{\varepsilon}_{ij}$$ Level-1 measurement model

$$\boldsymbol{\eta}_{ij} = \boldsymbol{\alpha}_j + \mathbf{B}_j \boldsymbol{\eta}_{ij} + \boldsymbol{\Gamma}_j \mathbf{X}_{ij} + \boldsymbol{\zeta}_{ij}$$ Level-1 structural model

$$\boldsymbol{\eta}_j = \boldsymbol{\mu} + \boldsymbol{\beta}\boldsymbol{\eta}_j + \boldsymbol{\gamma}\mathbf{X}_j + \boldsymbol{\zeta}_j$$ Level-2 structural model (NEW!)

Furthermore,

$$\boldsymbol{\varepsilon}_{ij} \sim MVN(0, \boldsymbol{\Theta})$$ Item residual (co)variances

$$\boldsymbol{\zeta}_{ij} \sim MVN(0, \boldsymbol{\Psi})$$ Level-1 structural residual covariances

$$\boldsymbol{\zeta}_j \sim MVN(0, \boldsymbol{\psi})$$ Level-2 structural residual covariances (NEW!)

MSEM equations continued 2

$$\mathbf{Y}_{ij} = \mathbf{\Lambda}\boldsymbol{\eta}_{ij}$$

Links observed variables to "between" and "within" latent variables (LVs)

$$\boldsymbol{\eta}_{ij} = \boldsymbol{\alpha}_j + \mathbf{B}_j\boldsymbol{\eta}_{ij} + \boldsymbol{\zeta}_{ij}$$

Links "within" LVs to each other, and "between" LVs to random intercepts

$$\boldsymbol{\eta}_j = \boldsymbol{\mu} + \boldsymbol{\beta}\boldsymbol{\eta}_j + \boldsymbol{\zeta}_j$$

Links "between" LVs to each other

Almost all MSEM models can be expressed using a reduced set of equations that eliminate some matrices.

One common simplification is to treat all variables as if they were endogenous (dependent), which removes $\mathbf{K}_j\mathbf{X}_{ij}$, $\mathbf{\Gamma}_j\mathbf{X}_{ij}$, and $\boldsymbol{\gamma}\mathbf{X}_j$ from the model.

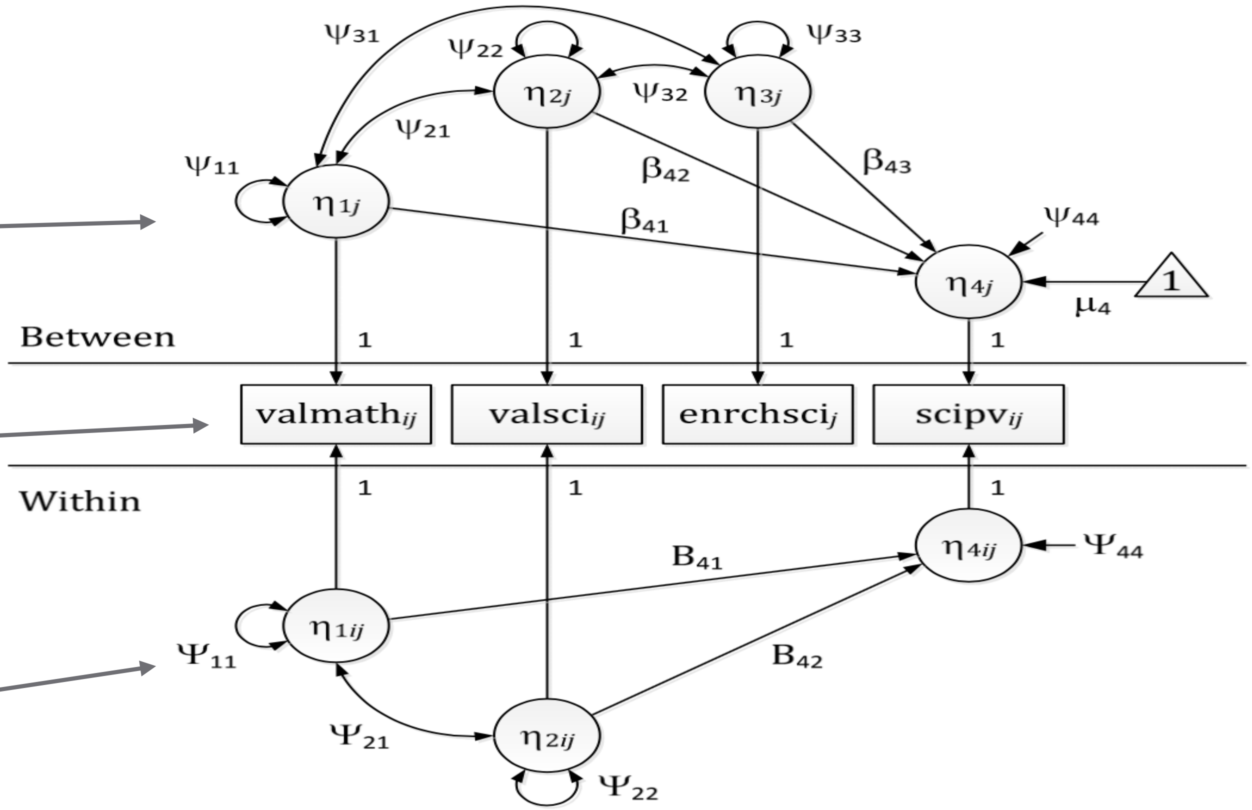
Item-level intercepts and residual variances usually can be estimated as part of the "between" model, which removes $\boldsymbol{\nu}_j$ and $\boldsymbol{\varepsilon}_{ij}$.

New Path Diagrams
(Preacher et al., 2010)

Between component latent variables

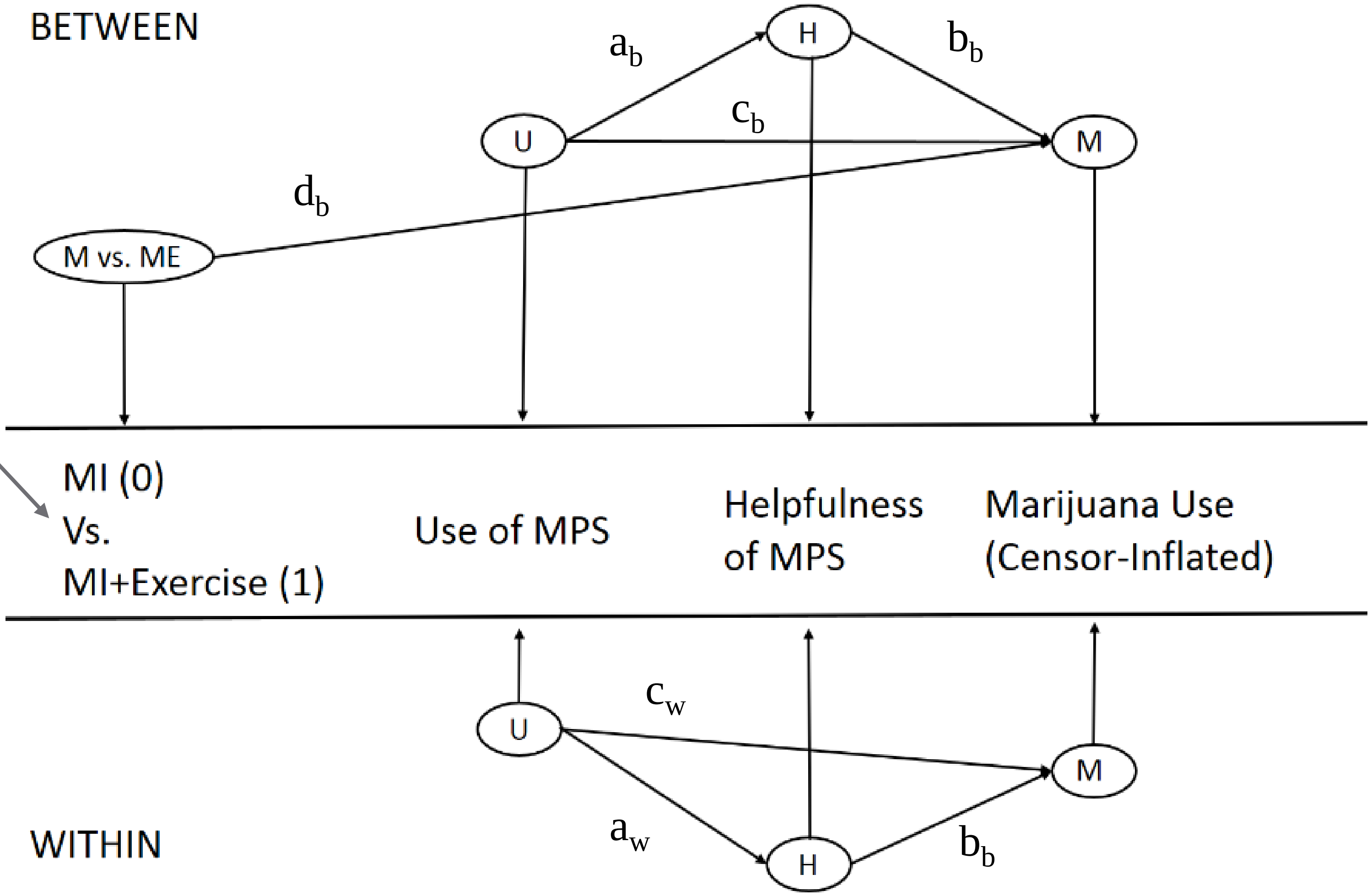
Observed variables

Within component latent variables

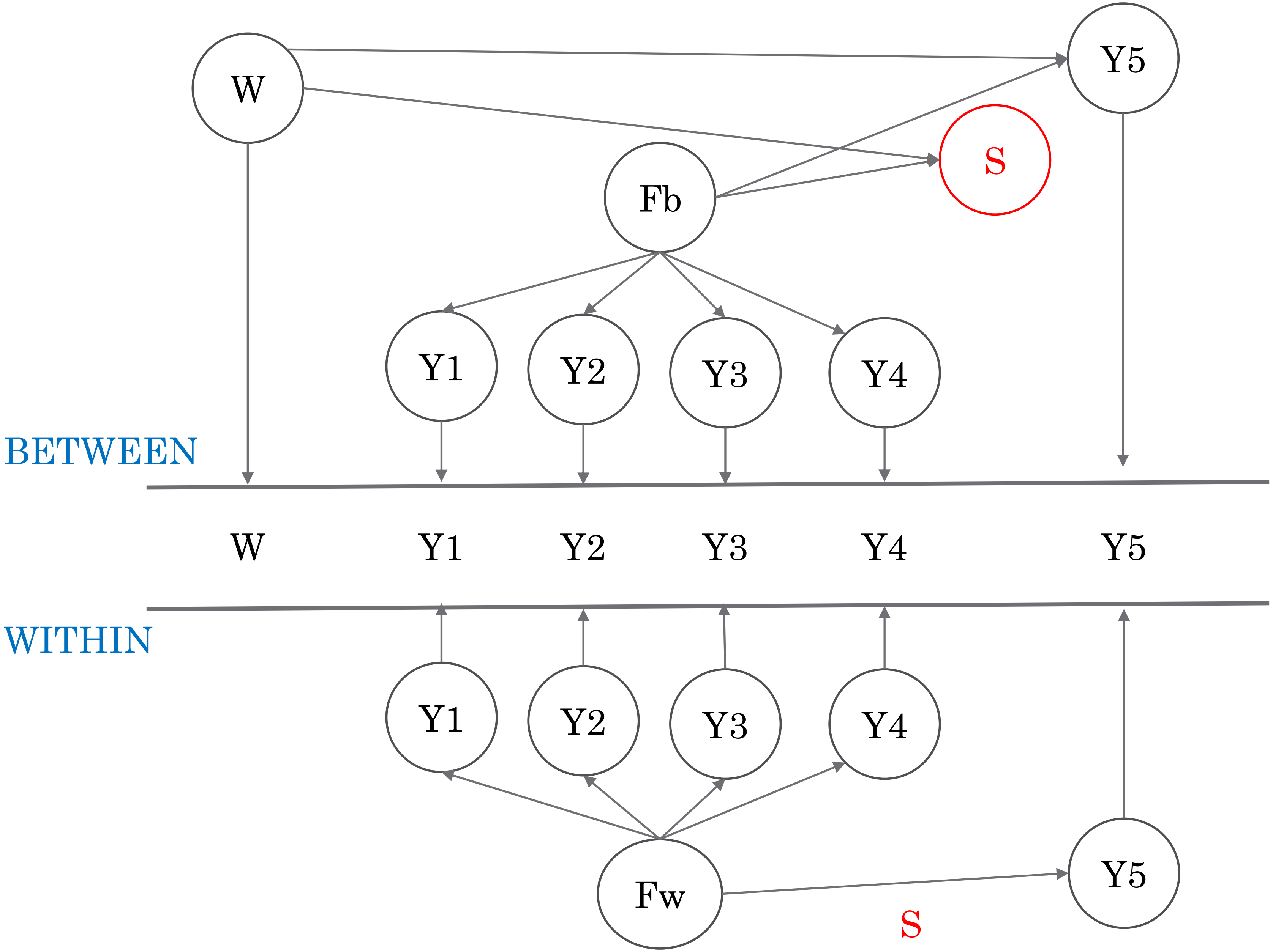


MSEM Perceived Helpfulness of MPS (EMA Ratings)

Note: Treatment Condition only has a between component



MSEM with latent factors

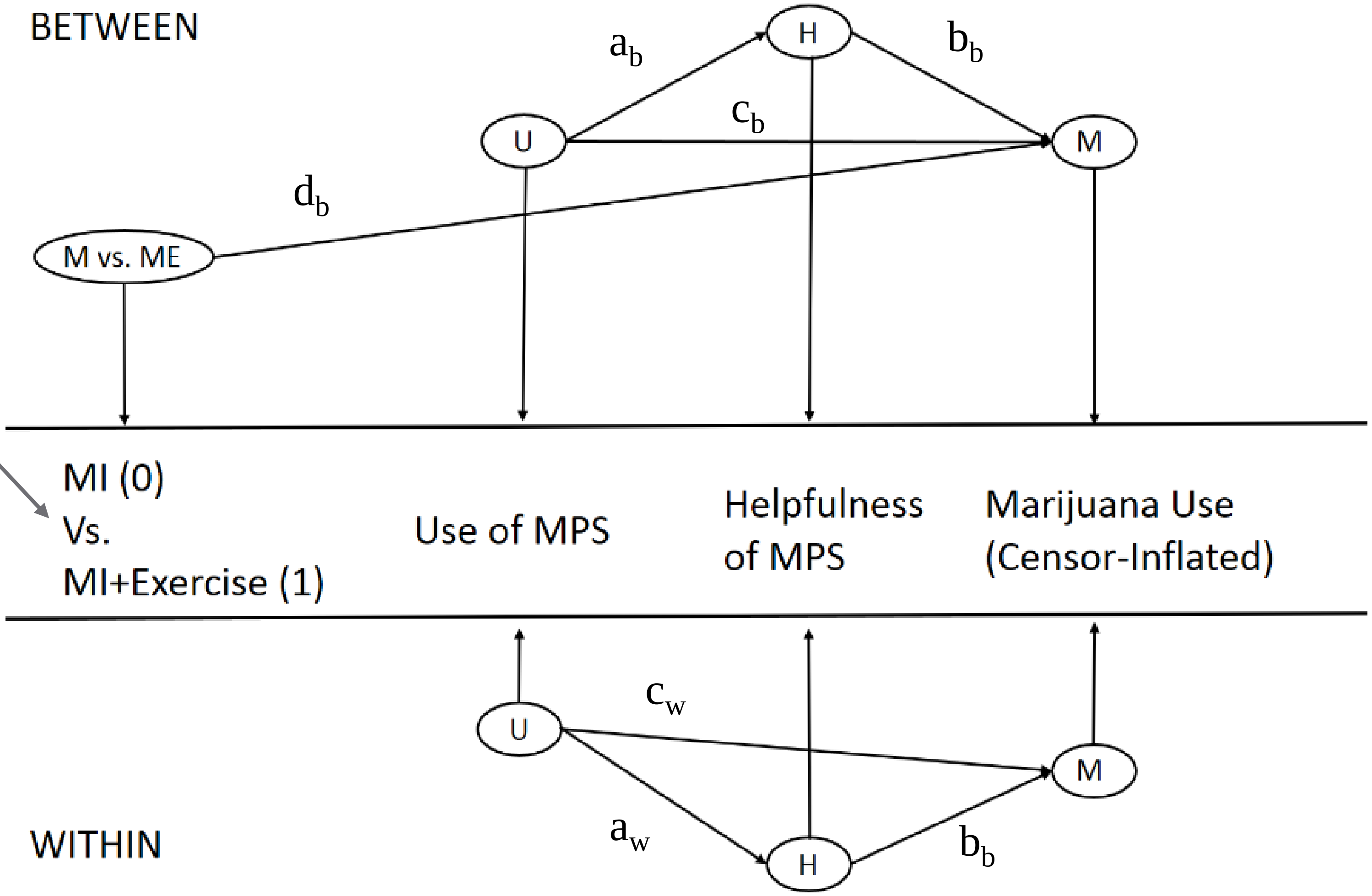


How to write the code…

- You have essentially already learned all the code
  - Use the code from SEM, MLM, Mixtures
    - On, with, by, %within%, %between%, s|, constraints
- Example:
  - We'll walk through the marijuana example from the previous slide
    - But, we'll save the mediation tests for next week.
- Method
  - Participants participated in one of two treatment conditions
    - Motivational Interviewing (MI) or MI+exercise
  - Provided data on marijuana quantity, use of marijuana protective strategies (MPS), and helpfulness of MPS at every use episode (could be multiple times per day) across 21 days.

MSEM Perceived Helpfulness of MPS (EMA Ratings)

Note: Treatment Condition only has a between component



```
VARIABLE:
NAMES ARE
ID TOTALJOINTS USESTRAT USEOTHSTRAT HELPFULSTRAT
LIKELYUSESTRAT condition;

MISSING = ALL (-999);

USEVARIABLES ARE
ID TOTALJOINTS  USESTRAT  HELPFULSTRAT condition
cuse chelp  ;

Between are cuse chelp condition;
Within are USESTRAT HELPFULSTRAT ;

Cluster is ID;

Define:
 cuse = cluster_mean (usestrat);
 chelp = cluster_mean (helpfulstrat);

Analysis:    Type is Twolevel;
```

When a variable is listed as within or
Between all of its variance is put on that
Level – unlisted variables are treated
On both levels

Manually disaggregated variables
To use on between level
Necessary because of mediation test

No random specification – thus
fixed effects model

```
Model:
%WITHIN%
TOTALJOINTS on usestrat (cw);
TOTALJOINTS on helpfulstrat (bw);
helpfulstrat on usestrat (aw);
```

This code specifies the within level
Mediation paths (i.e., the triangle on the
Bottom)

```
%BETWEEN%
TOTALJOINTS on cuse (cb);
TOTALJOINTS on chelp (bb);
chelp on cuse (ab);
totaljoints on condition;
```

This code specifies the between level
Mediation paths (i.e., the triangle on the
Top) and the direct effect from condition
To marijuana use (which we are treating
As normal for simplicity)

ICC

There are 34 clusters with an average size of 26 (i.e., the typical participant reported 26 use episodes in 3 weeks)

```
Number of missing data patterns             2
Number of clusters                         34

Average cluster size       26.000

Estimated Intraclass Correlations for the Y Variables

           Intraclass              Intraclass
Variable  Correlation   Variable  Correlation

TOTALJOI     0.306      HELPFULS     0.000
```

31% of the variance in marijuana use is between Persons and 69% is within persons

100% of the variance of helpfulness is within Persons, because we specified it that way

```
MODEL FIT INFORMATION

Number of Free Parameters                       14

Loglikelihood

          H0 Value                       -3216.328
          H0 Scaling Correction Factor      7.6584
            for MLR
          H1 Value                       -3216.161
          H1 Scaling Correction Factor      7.2070
            for MLR

Information Criteria

          Akaike (AIC)                    6460.657
          Bayesian (BIC)                  6527.639
          Sample-Size Adjusted BIC        6483.178
            (n* = (n + 2) / 24)

Chi-Square Test of Model Fit

          Value                              0.378*
          Degrees of Freedom                     1
          P-Value                           0.5385
          Scaling Correction Factor         0.8867
            for MLR

*   The chi-square value for MLM, MLMV, MLR, ULSMV, WLSM and WLSMV cannot be used
    for chi-square difference testing in the regular way.  MLM, MLR and WLSM
    chi-square difference testing is described on the Mplus website.  MLMV, WLSMV,
    and ULSMV difference testing is done using the DIFFTEST option.

RMSEA (Root Mean Square Error Of Approximation)

          Estimate                           0.000

CFI/TLI

          CFI                                1.000
          TLI                                1.011

Chi-Square Test of Model Fit for the Baseline Model

          Value                            472.108
          Degrees of Freedom                     8
          P-Value                           0.0000

SRMR (Standardized Root Mean Square Residual)

          Value for Within                   0.000
          Value for Between                  0.023
```

This should all look familiar.
At the top are comparative fit indices
At the bottom are model specific indices

MODEL RESULTS

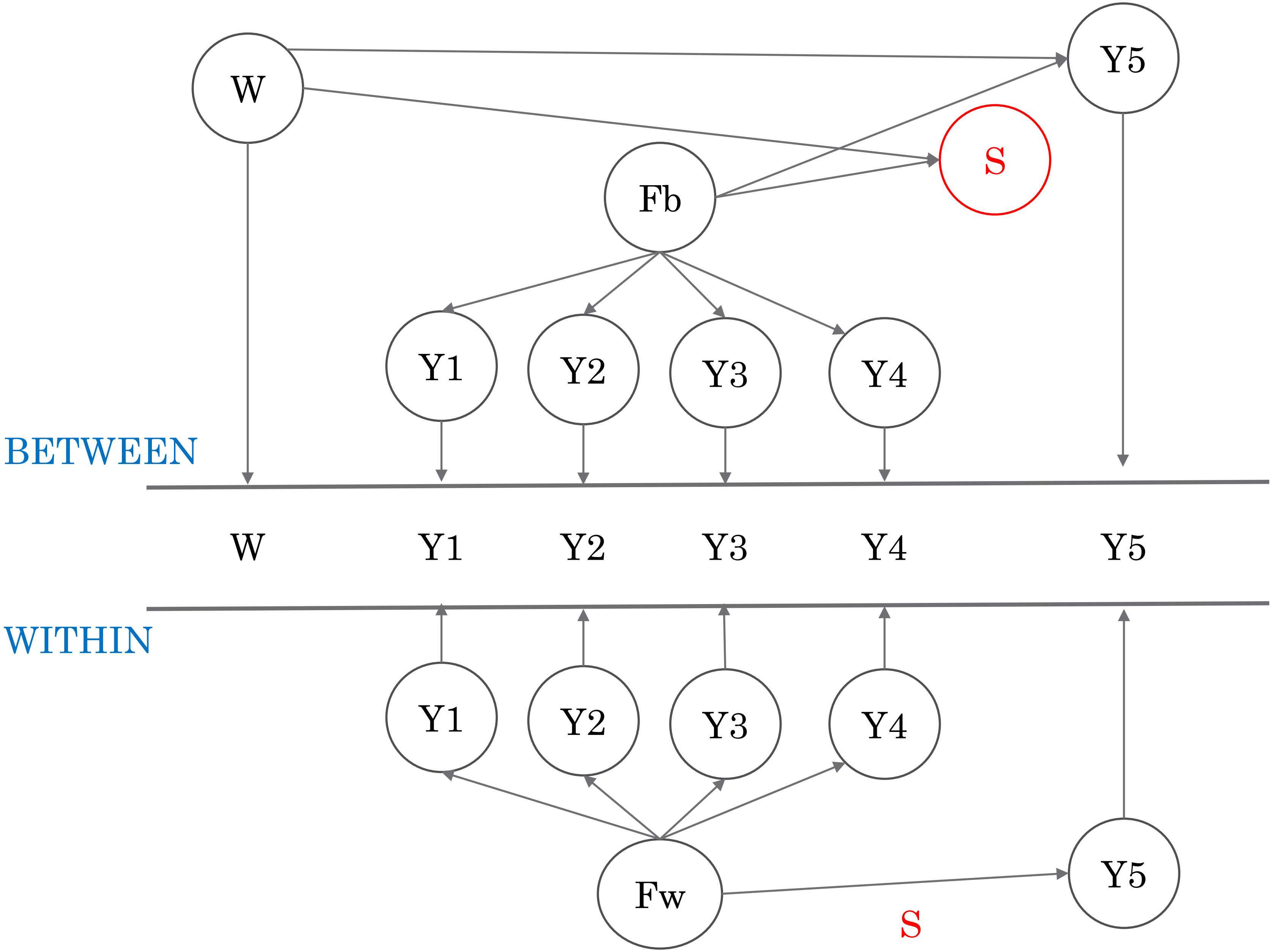
	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
Within Level				
TOTALJOINT ON				
USESTRAT	0.109	0.178	0.611	0.541
HELPFULSTR	-0.059	0.027	-2.176	0.030
HELPFULSTR ON				
USESTRAT	-3.185	0.393	-8.096	0.000
Intercepts				
HELPFULSTR	9.777	0.615	15.906	0.000
Residual Variances				
TOTALJOINT	1.064	0.409	2.601	0.009
HELPFULSTR	3.789	0.370	10.239	0.000
Between Level				
TOTALJOINT ON				
CUSE	0.093	0.657	0.141	0.888
CHELP	0.016	0.108	0.151	0.880
CONDITION	0.046	0.263	0.174	0.862
CHELP ON				
CUSE	-4.137	0.729	-5.674	0.000
Intercepts				
CHELP	11.268	1.077	10.463	0.000
TOTALJOINT	1.514	1.459	1.038	0.299
Residual Variances				
CHELP	1.524	0.364	4.189	0.000
TOTALJOINT	0.483	0.103	4.708	0.000

Because everything is fixed, we can use typical interpretations

Within:
During a given use episode, the helpfulness of protective strategies was negatively associated with marijuana use, but the sheer use of strategies was not

Between:
For the average person, across episodes neither having a tendency to use strategies or having a tendency to rate strategies as helpful were predictive of marijuana use. Treatment did not have on effect on marijuana use.

MSEM with latent factors



MSEM with latent variables code

```
TITLE:  this is an example of a two-level SEM with continuous
factor indicators and a random slope for a factor

DATA:   FILE IS ex9.10.dat;

VARIABLE:
NAMES ARE y1-y5 w clus;

BETWEEN = w;
CLUSTER = clus;

ANALYSIS:
TYPE = TWOLEVEL RANDOM;

ALGORITHM = INTEGRATION;
INTEGRATION = 10;

MODEL:
%WITHIN%
fw BY y1-y4;
s | y5 ON fw;

%BETWEEN%
fb BY y1-y4;
y1-y4@0;
y5 s ON fb w;

OUTPUT:
TECH1 TECH8;
```

W is a between only variable

Allow random intercepts and slopes

Specifying the within level latent variable

Creating a random slope

Specifying the between level latent variable

Fixing between level random variance to 0

Between level regression paths

```
MODEL RESULTS

                                                    Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value

Within Level

 FW       BY
    Y1                 1.000      0.000    999.000    999.000
    Y2                 1.286      0.119     10.813      0.000
    Y3                 1.199      0.098     12.282      0.000
    Y4                 1.262      0.096     13.201      0.000

 Variances
    FW                 0.732      0.112      6.539      0.000

 Residual Variances
    Y1                 1.041      0.083     12.502      0.000
    Y2                 0.911      0.060     15.276      0.000
    Y3                 0.983      0.081     12.151      0.000
    Y4                 0.995      0.069     14.392      0.000
    Y5                 0.475      0.046     10.400      0.000

Between Level

 FB       BY
    Y1                 1.000      0.000    999.000    999.000
    Y2                 0.978      0.094     10.403      0.000
    Y3                 1.085      0.093     11.628      0.000
    Y4                 1.007      0.125      8.045      0.000

 S          ON
    FB                -0.029      0.192     -0.151      0.880

 S          ON
    W                  0.366      0.109      3.350      0.001

 Y5         ON
    FB                 0.450      0.142      3.181      0.001

 Y5         ON
    W                  0.630      0.072      8.801      0.000

 Intercepts
    Y1                -0.074      0.096     -0.765      0.444
    Y2                -0.015      0.096     -0.154      0.878
    Y3                 0.005      0.103      0.044      0.965
    Y4                -0.029      0.099     -0.294      0.769
    Y5                -0.026      0.089     -0.286      0.775
    S                  1.232      0.122     10.070      0.000

 Variances
    FB                 0.455      0.111      4.110      0.000

 Residual Variances
    Y1                 0.000      0.000    999.000    999.000
    Y2                 0.000      0.000    999.000    999.000
    Y3                 0.000      0.000    999.000    999.000
    Y4                 0.000      0.000    999.000    999.000
    Y5                 0.203      0.057      3.554      0.000
    S                  0.579      0.124      4.657      0.000
```

Within factor loadings:
Factor structure on a typical assessment occasion
(think states)

Between factor loadings:
Factor structure across assessments
(think traits)

Between regressions

Average within level slope
Also, the regression of Y5 on Fw

How to write up the results…

- Same as always.
  - Data decisions
  - Hypotheses
  - Model fit (if available)
  - Model building
  - Within and between specific parameter estimates
- Be careful with the language you use.