A Guide to Conducting SEM in Counseling Research

Stephanie Crockett, PhD, NCC
Oakland University
crockett@oakland.edu
Workshop Outline

• Introduction to SEM
  – Definitions, Advantages, Statistical Software, Misconceptions/Misapplications, Key Terms

• 5 Steps for Conducting SEM
  – Model Specification, Identification, Estimation, Testing and Modification

• Additional Considerations
  – Sample size, Data screening, Testing mediation effects

• Advanced SEM models

• Best Practices for Publishing
INTRODUCTION TO SEM
Structural Equation Modeling

- SEM is a second-generation multivariate analysis technique that is used to determine the extent to which an a priori theoretical model is supported by the non-experimental sample data (Raykov & Marcoulides, 2000; Schumacker & Lomax, 2010).
  - Allows the researcher to understand the complex relationships that occur among observed variables and latent constructs.
Advantages to Using SEM

• Test the plausibility of complex theoretical models.
• Simultaneously analyze the direct and indirect effects of multiple variables
  – Test mediation and moderator effects
• Take into account measurement error in the model’s observed variables
• Compare viable, alternate models to the theoretical model
• Provide a path diagram, or visual representation, of the hypothesized relationships among variables
Statistical Software

- LISREL1
- AMOS
- R
- EQS1,
- Mx, Mplus1
- Ramona
- SEPATH1

- Many of the SEM software programs allow researchers to statistically analyze raw data and provide procedures for managing missing data, outliers, and variable transformations.
- Many SEM software programs allow the option to construct a path diagram that can be translated by the software program into the mathematical equations needed for analysis.
- [James Gaskin SEM Tutorials](#)
The Problem

• SEM has become the most widely used advanced statistical method in the social sciences (Martens, 2005; Schumacker & Lomax, 2010)
  – Increased accessibility to user-friendly SEM software
• The number of published articles utilizing SEM methods has increased over the past 25 years, as well as the number of journals publishing SEM related articles (Hershberger, 2003)
• Many researchers have little, if any, formal SEM background, potentially leading to misapplications and publications of questionable utility (Mueller & Hancock, 2008; Steiger, 2001; Tomarken & Waller, 2005).
Common Misconceptions and Misapplications

• Conceptual issues:
  – use of SEM to test causal hypotheses (Mueller & Hancock, 2008)
  – use SEM as an exploratory technique (Schreiber, Nora, Stage, Barlow, & King, 2006)
  – poor theoretical specification of the SEM model (Guo, Perron, & Gillespie, 2009; Martens, 2005)
  – failure to specify and test alternative models (McDonald & Ho, 2002; Tomarken & Waller, 2005)
Common Misconceptions and Misapplications

• Execution and interpretation issues:
  – all-in-one SEM analysis (Martens, 2005; McDonald & Ho, 2002; Mueller & Hancock, 2008)
  – failing to assess for multivariate normality (Guo, Perron, & Gillespie, 2009; Martens, 2005)
  – sample size and power (Tomarken & Waller, 2005)
  – misinterpreting goodness-of-fit (Tomarken & Waller, 2005)
  – failure to provide a full report of parameter estimates (Hoyle & Paner, 1995; MacCallum & Austin, 2000; McDonald & Ho, 2002; Martens, 2005)
  – errors related to model modification (Guo, Perron, & Gillespie, 2009; Martens, 2005)
So here’s what you need to remember. . .

• Examines the relationships among variables, but cannot establish causal effects.

• Confirmatory technique and is most advantageous when the researcher has an a priori theoretical model to test.
  – It is not an exploratory technique and “is ill suited for exploring and identifying relationships” among variables (Kelloway, 1998, p. 7).

• Practically, it requires large sample sizes, and additional statistical software

• Learning a new terminology
Basic SEM Terms to Know

• Exogenous variable
• Endogenous variable
• Latent construct
• Observed variable
  – Parceling, subscales, single items (see: Little et al., 2002)
• Direct/indirect effects
  – Mediator variable
  – Full vs. partial mediation
CONDUCTING SEM
5 Steps for Conducting SEM

• Model Specification
• Model Identification
• Model Estimation
• Model Testing
• Model Modification

*See Crockett (2012) for a detailed overview
MODEL SPECIFICATION
Overview: Model Specification

• Occurs prior to data collection and analysis.
• Involves the development of a theoretical model using applicable, related theory and research to determine variables of interest and the relationships among them (Cooley, 1978).
• A path diagram can be constructed to visually represent the hypothesized relationships among variables in the theoretical model.
Does SEM align with my research purpose and questions?

- Research purpose and questions dictate the use of SEM
  - Model testing
  - Confirmatory, not exploratory
    - Should be sufficient information on the theoretical framework for the model being tested
  - Inclusion of latent variables
    - Example: Working Alliance Inventory

(Boomsma, 2000; Hoyle & Panter, 1995)
Example Research Questions

• Does the therapeutic working alliance partially mediate the relationship between counselor multicultural competence and client counseling outcomes?

• Is the model, where the therapeutic working alliance is a mediator variable, the best fit for the data?
Variables and Measures

• **Exogenous variable: Counselor multicultural competence**
  – Measured by the Multicultural Awareness, Knowledge, and Skills Survey (MAKSS) Revised (Kim, Cartwright, Asay, & D’Andrea, 2003)
  – 33 items; 10-Awareness, 13-Knowledge, 10-Skills

• **Mediator variable: Therapeutic working alliance**
  – Measured by the Working Alliance Inventory-Short Form (WAI-SF; Hatcher & Gillaspy, 2006).
  – 12 items; 4-Goals, 4-Tasks, 4-Bond

• **Endogenous variable: Client counseling outcomes**
  – Measured by the OQ-45.2 (Lambert, Hansen, Umpress, Lunnen, Okiishi, & Burlingame, 2003)
  – 45 items; 25-symptom distress, 11-interpersonal relations, 9-social role
Model Specification-2 Step Process

• Measurement Model Specification
  – involves the identification of observed variables that comprise each of the model’s latent constructs.

• Structural Model Specification
  – specifies relationships among the latent variables in the theoretical model.

*It is imperative that such relationships are indicated prior to model estimation and testing as SEM is a confirmatory technique.
Writing Equations for the Models

• Measurement Model Example for WAI
  – Tasks = function of working alliance + error
  – Goals = function of working alliance + error
  – Bond = function of working alliance + error

• Structural Model
  – Counseling outcomes = structure coefficient\textsubscript{1} \times Counselor Multicultural Competence + error
  – Counseling outcomes = structure coefficient\textsubscript{2} \times Working Alliance + error
  – Working Alliance = structure coefficient\textsubscript{3} \times Counselor Multicultural competence + error
Constructing a Path Diagram

Path Diagram Symbols

- **Observed Variable**
- **Latent Variable**
- **Unidirectional path**
- **Covariance among two IVs**
- **Error Term (Measurement error)**
Measurement Model

Awareness
Knowledge
Skill
Task
Goal
Bond
Symptom
Distress
Interpersonal Relationships
Social Role
Counselor Multicultural Competence
Therapeutic Working Alliance
Counseling Outcomes
Full Theoretical, Structural Model

Therapeutic Working Alliance

Counselor Multicultural Competence

Aim
Goal
Bond

Skill
Knowledge
Awareness

Symptom Distress
Interpersonal Relationships
Social Role

Counseling Outcomes
SEM WORKSHOP EXERCISE

• Work individually to complete steps 1-3
  – **STEP ONE**: Articulate your theoretical model and identify your research question(s) below.
  – **STEP TWO**: If you can, identify the measurements you will use to answer your research questions. How many items does each measure have? What is the factor structure of the instrument (e.g., subscales)?
  – **STEP THREE**: Construct a structural path diagram of your theoretical model. If you can, construct the measurement model of your theoretical model.
MODEL IDENTIFICATION
Overview: Model Identification

• Concerned with whether a unique solution to the model can be generated.

• It must be theoretically possible to establish a unique estimate for each parameter (Kelloway, 1998; Schumacker & Lomax, 2010)
  – Consider: \( x + y = 10 \)
  – Free, fixed, and constrained parameters
Overview: Model Identification

• O’Brien (1994) exerted that the measurement model is most likely identified when:
  – there are two or more latent variables, each with at least three indicators that load on it, the errors of these indicators are not correlated, and each indicator loads on only one factor.

• Reference variable: a causal path from each latent variable to a corresponding observed variable must be fixed at one (Kline, 2010)
Overview: Model Identification

• Bollen (1989) outlined a widely used set of rules for the identification of *structural models*: the recursive rule and the t rule.

• **t rule:** The structural model must have more “known” pieces of information than “unknown” pieces in order to find unique solutions.
  - **KNOWN:** $p(p + 1)/2$, where $p$ is equal to the number of observed variables.
  - **UNKNOWN:** is equal to the number of free parameters to be estimated in the model

• More reading on model identification:
SEM WORKSHOP EXERCISE

• In pairs or small groups:
  – Share your research questions and model with the group
  – Solicit feedback on your research questions and model
  – Discuss and answer Step 4 with your group members

• **STEP FOUR:** What steps can you take to ensure that both models are identified? Explain.
MODEL ESTIMATION
Overview of Model Estimation

• SEM analysis programs use an iterative procedure, often referred to as a fitting function, to minimize the differences between the estimated theoretical covariance matrix $P$ and the observed covariance matrix $S$.
  
  – Specifically, the iterative procedure attempts to improve the preliminary parameter estimates with subsequent calculation cycles.
  
  – The final parameter estimates represent the best fit to observed covariance matrix $S$. 
Overview of Model Estimation

• Estimating the parameters of the theoretical model in such a way that the theoretical parameter values yield a covariance matrix as close as possible to the observed covariance matrix $S$.

• Several fitting functions are available to researchers
  – e.g., ordinary least squares [OLS], generalized least squares [GLS], maximum likelihood [ML]

• For more information see:
MODEL TESTING
Overview of Model Testing

• **A two-step approach** (Anderson & Gerbing, 1988; James, Mulaik, & Brett, 1982)
  - Model testing involves the analysis of the measurement model and the structural model.

• **Step One: Measurement model testing**
  – Ensure that the chosen observed indicators for a latent construct actually measure the construct.
  – Conduct a confirmatory factor analysis (CFA)
Measurement Model CFA Results

Note. N=281. All model parameters are significant at the p > .01 level.
Example Study CFA Results

- MAKSS was significantly, positively correlated with its factor indicators:
  - MAKSS Awareness subscale ($r = .85$, $p < .01$)
  - MAKSS Knowledge subscale ($r = .80$, $p < .01$)
  - MAKSS Skill subscale ($r = .83$, $p < .01$)
- Therapeutic working alliance was significantly positively correlated with its factor indicators:
  - WAI-SF Bond subscale ($r = .86$, $p < .01$)
  - WAI-SF Task subscale ($r = .91$, $p < .01$)
  - WAI-SF Goal subscale ($r = .85$, $p < .01$)
- OQ-45.2 was significantly positively correlated with its factor indicators:
  - OQ-45.2 Symptom distress subscale ($r = .81$, $p < .01$)
  - OQ-45.2 Interpersonal relationships subscale ($r = .79$, $p < .01$)
  - OQ-45.2 Social role subscale ($r = .88$, $p < .01$)
- Model fit:
  - $\chi^2(19) = 44.80$, $p < .01$; RMSEA = .08; CFI = .97
Overview of Model Testing

• Structural model testing
  – the extent to which the model is supported by the sample data.

• Model fit:
  – fit of individual model parameters
  – overall global fit of the entire model

• For individual model parameters, consider:
  – Statistical significance of parameter
  – Direction (sign should agree with theoretical model)
  – Parameter estimates should remain within expected values
Structural Model Results

Model Fit:
\[ \chi^2(32) = 75.30, \ p < .01; \ RMSEA = .08; \ CFI = .97; \ PNFI = .67 \]

Note. \( N=281 \). All model parameters are significant at the \( p > .01 \) level.
Overview of Model Testing

• Absolute fit:
  – Concerned with the structural model’s ability to reproduce the sample covariance matrix $S$
  – $\chi^2$ test = non-significant value indicates that that theoretical model covariance matrix $\Sigma$ and the sample covariance matrix $S$ are similar.
  – RMSEA = Any value lower than 1.00 is assumed to be an adequate fit to the data, with values lower than .05 being a very good fit to the data (Steiger, 1990).
  – Goodness-of-fit index (GFI), and adjusted goodness-of-fit index (AGFI)
Overview of Model Testing

• Comparative fit:
  – Compare the model to a baseline model that is said to fit the data poorly.
    • Determine the relative position of model fit on a continuum that ranges from worst fit to perfect fit.
  – Normed fit index (NFI)
  – Comparative fit index (CFI)
  – Relative fit index (RFI)
*Values range from 0 to 1, looking for values greater than .90
Testing Alternative Models

• Structural models should also be tested against viable alternative models (Kelloway, 1998)
  – That is, two or more plausible models are compared to one another to determine which model best fits the sample data.
  – **Nested alternative models**: developed by fixing some of the free parameters in the theoretical model to 0
Alternative, Direct Model

Awareness → Counselor Multicultural Competence → Counseling Outcomes → Symptom Distress

Knowledge → Counselor Multicultural Competence → Counseling Outcomes → Interpersonal Relationships

Skill → Counselor Multicultural Competence → Counseling Outcomes → Social Role
Example Study SEM Results

Model Fit
\[ \chi^2(35) = 123.19, \ p = .64; \ \text{RMSEA} = .20; \ \text{CFI} = .75 \]

Note. \( N=281 \). All model parameters are significant at the \( p > .01 \) level.
Overview of Model Testing

• Parsimonious fit
  – Determine whether the impact of adding additional parameters to the model is worth the decrease in degrees of freedom.
  – Parsimonious normed fit index (PNFI)
  – Parsimonious goodness-of-fit index (PGFI)
    • Both indices range from 0 to 1
    • No standard cutoff point for determining a good fit
    • Best used to compare two or more models; the model having the highest PNFI or PGFI would be the most parsimonious model.
Comparing Nested Models

- Research question:
  - Is the model, where the therapeutic working alliance is a mediator variable, the best fit for the data?

<table>
<thead>
<tr>
<th>Model</th>
<th>Chi-Square</th>
<th>$p$-value</th>
<th>df</th>
<th>RMSEA</th>
<th>CFI</th>
<th>PNFI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measurement Model</td>
<td>44.88</td>
<td>&lt;.01</td>
<td>19</td>
<td>0.08</td>
<td>0.97</td>
<td>0.65</td>
</tr>
<tr>
<td>Hypothesized Model</td>
<td>75.30</td>
<td>&lt;.01</td>
<td>32</td>
<td>0.08</td>
<td>0.97</td>
<td>0.67</td>
</tr>
<tr>
<td>Direct Model</td>
<td>123.19</td>
<td>&lt;.01</td>
<td>35</td>
<td>0.20</td>
<td>0.75</td>
<td>0.57</td>
</tr>
</tbody>
</table>
Findings

• The hypothesized model is a better fit to the data than the alternate, direct model.
• Therapeutic working alliance may partially mediate the relationship between counselor multicultural competence and client counseling outcomes.
• Individually complete steps 5-6
  – **STEP FIVE:** Identify and articulate one or two alternative models you can use to test against your original theoretical model.
  – **STEP SIX:** What issues do you anticipate during data collection and analysis? How might you addresses these anticipated issues?
• After completing these steps discuss them with your small group
MODEL MODIFICATION
(TRIMMING PARAMETERS)
Overview of Model Modification

• Researchers employ model modification methods in an attempt to find a model that better fits the data.
  – Specification search and examining standardized residual matrix to identify non-significant parameters

• This step is HIGHLY controversial!!!
Overview of Model Modification

• Elimination of parameters must be theoretically justified (McDonald & Ho, 2002)
  – Should not be empirically driven to simply improve model fit
  – See Olsson, Troye, & Howell (1999) for discussion on empirical vs. theoretical fit.

• Model should be discussed using tentative terms until cross-validated on an independent sample and the importance of the modifications outlined in the manuscript (McDonald & Ho, 2000)
Modified Model Results

Model Fit
\[ \chi^2(33) = 78.03, \ p < .01; \ RMSEA = .08; \ CFI = .96; \ PNFI = .71 \]

**Note.** N=281. All model parameters are significant at the \( p > .01 \) level.
# Comparing Model Fit

<table>
<thead>
<tr>
<th>Model</th>
<th>Chi-Square</th>
<th>p-Value</th>
<th>df</th>
<th>RMSEA</th>
<th>CFI</th>
<th>PNFI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measurement Model</td>
<td>44.88</td>
<td>&lt;.01</td>
<td>19</td>
<td>0.08</td>
<td>0.97</td>
<td>0.65</td>
</tr>
<tr>
<td>Theoretical Model</td>
<td>75.3</td>
<td>&lt;.01</td>
<td>32</td>
<td>0.08</td>
<td>0.97</td>
<td>0.67</td>
</tr>
<tr>
<td>Alternative Model</td>
<td>344.19</td>
<td>&lt;.01</td>
<td>35</td>
<td>0.2</td>
<td>0.75</td>
<td>0.57</td>
</tr>
<tr>
<td>Modified Model</td>
<td>78.03</td>
<td>&lt;.01</td>
<td>33</td>
<td>0.08</td>
<td>0.96</td>
<td>0.71</td>
</tr>
</tbody>
</table>
Comparing Model Fit

• Consider model parsimony
  – PNFI or PGFI: modeling having highest PNFI or PGFI is most parsimonious

• For nested models use the $\chi^2$ difference test
  • $\chi^2_{\text{diff}} = \chi^2_s - \chi^2_l$ and $df_{\text{diff}} = df_s - df_l$
  • Check values manually using a $\chi^2$ table
  • Significant value indicates the model with more parameters fits the data better

• EXAMPLE: $\chi^2_{\text{diff}} = 2.73$; $df_{\text{diff}} = 1$
• $\chi^2$ Table: 3.84 at $p \leq .05$
ADDITIONAL CONSIDERATIONS
Power and Sample Size

• Golden standard: $N \geq 200$ (Kline, 2010)

• Debate over minimum sample size:

• Power analysis
  – Estimate the statistical power needed to detect the statistical significance of individual model parameters. See Saris and Satorra (1993)
  – MacCallum, Browne, and Sugawara (1996) developed a power analysis procedure is based on RMSEA that determines the probability of detecting a reasonably correct model. See Kline (2010)
Data Screening

• Missing Data
  – Assessment of underlying issue (MCAR, MAR, MNAR)
  – Tomarken and Waller (2005) recommend using multiple imputation (MI) or full-information maximum likelihood (FIML)

• Multivariate Outliers
  – Use of Mahalanobis distance ($D^2$) for detecting multivariate outliers

• Multivariate Normality
  – SEM estimation methods ML and GLS rely on the assumption that data are multivariate normal
  – Methods for assessing multivariate normality:
    • Plotting distributions; Mardia’s (1970) multivariate skewness and kurtosis
  – If dealing with non-normal data see Tomarken and Waller (2005)
Testing Mediation Effects

• Test mediation effect using bias-corrected bootstrap method
  – non-parametric method based on resampling with replacement which is done many times (e.g., 5000)
• Typically computes a mean and confidence interval for the mediation effect (e.g., Lower limit and Upper limit)
  – To be significant, zero must not occur between the LL and UL
  – Example: Mean indirect effect = .68; 95% CI ranged from .55 to .84
• For information
  – See Preacher and Hayes (2004)
  – Example in AMOS
Advanced Mediation Models

- **Multiple mediation model**: Uses two or more mediators in a model
  - [Tutorial for Multiple Mediation Model Analysis](#)

- **Mediation models with longitudinal data**
  - See Selig et al. (2009)
Moderation and Resources

• Inclusion of a moderator variable
  – A third variable that affects the direction and/or strength of the relationship between exogenous and endogenous variables

• Moderated mediation (See Preacher et al., 2007)
  – Mediation involves a moderator variable

• Mediated moderation
  – Moderation involves a mediator variable
BEST PRACTICES FOR PUBLISHING
Introduction Section: Rationale and Research Questions

• Research purpose and questions dictate the use of SEM
  – Model testing
  – Confirmatory, not exploratory
  – Inclusion of latent variables

• Provide sufficient information on the theoretical framework for the model
  – Describe what led to the initial model conceptualization
  – Introduce the latent constructs and articulate the relationships among them

(Boomsma, 2000; Hoyle & Panter, 1995)
Introduction Section: Rationale and Research Questions

• Identify plausible alternative models that might compete with the theoretical model
  – Competing models that might explain the data better than the theoretical model
  – Strengthens study design as it may mitigate confirmation bias
  – Nested models
• It can be helpful to include a path diagram of your theoretical model and possibly an alternative model.

(Martens, 2005; McDonald & Ho, 2002)
Methods Section Overview

• Sample
  – Rationale for sample size
• Instrumentation
  – Justification for choice of indicators
• Data Collection Procedures
• Data Analysis

In general you want to report sufficient information to allow for evaluation and replication (Boomsma, 2000).
Results Section Overview

• Data Screening Issues
  – Discuss the extent to which it is a problem and describe the method(s) used to handle the issue (McDonald & Ho, 2002)

• Descriptive Data Results
  – Means and standard deviations of all observed variables
  – Correlation matrix

• Tips
  – Display information in a table for up to 30 variables
  – In the case of large data sets (greater than 30), McDonald and Ho (2002) recommend making the correlation matrix available to the reader if requested.
Results Section: Measurement Model

• Report the results of the Confirmatory Factor Analysis (CFA) for the measurement model.
  – Report loadings and direction of all observed variables
  – Refer to Measurement Model figures and/or tables

• Report model fit
  – Chi-square test: report degrees of freedom, $N$, chi-square value, and $p$ value
    • $\chi^2(48, N = 500) = 303.80, p < .001$
  – Additional fit indices (Schreiber et al., 2006)
    • RMSEA report value and confidence intervals
      – RMSEA = .65, 90% confidence interval (CI) [.05, .11]
    • CFI, NFI, SRMR report value
      – CFI = .90

• Summarize conclusions drawn from statistical findings
Results Section: Theoretical (Structural) Model

- Report model fit
  - Chi-square test: report degrees of freedom, $N$, chi-square value, and $p$ value
  - Additional fit indices (Schreiber et al., 2006)
- Report all parameter estimates
  - Standardized coefficients
  - Explain in non-statistical terms the conclusions drawn from these relationships.
  - Refer to the path diagram
- Describe any additional tests performed (e.g., bootstrapping to test for mediation)
Results Section: Test of Alternative Models

• Structural models should also be tested against viable alternative models (Kelloway, 1998)

• Report results of any plausible alternative models tested
  – Report model fit and parameter estimates
  – Consider model parsimony:
    • PNFI or PGFI; model having highest PNFI or PGFI is most parsimonious
  – For nested models use the $\chi^2$ difference test
Discussion

• Overview of study purpose and summary of main study findings
• Comparison of study findings to the extant literature
• Implications for research and practice
• Limitations and future directions for research
Final Thoughts and Tips

• SEM Nomenclature
  – Define for the reader or eliminate from manuscript if possible

• Abbreviations for variable names
  – Example: Counselor multicultural competence (CMC)

• Clearly define the theoretical model early in the manuscript
  – May need to define terms like mediation and moderation for the reader

• Consistency throughout the manuscript is key
  – Model names, variable names, etc.
Final Thoughts and Tips

• Tell a concise story
  – Balance statistical jargon with a narrative about the theoretical model and alternative models.
  – Reflect on what readers really need to know
  – Decide what alternative models to include in the publication
• Use tables and figures to visually display models and data
• Fancy statistical procedures do not compensate for a poor design
QUESTIONS