Select One:
New x Delete Course Modification ______

Type of modification:
_____ Title ____ Description _____ Credit hours
_____ Prerequisites ____ Grade Type
 _____ Number: Old ______
    New ______

Title: Advanced Research Methods

College/School: Human Sciences  Dept: HDFS

Abbreviated Title: Advanced Research Methods (30 spaces total)

<table>
<thead>
<tr>
<th>CREDIT OFFERED</th>
<th>CONTROLS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levels: (select all that apply)</td>
<td>Grading Rule:</td>
</tr>
<tr>
<td>None (Blank)</td>
<td>Undergraduate (U)</td>
</tr>
<tr>
<td>Undergraduate (U)</td>
<td>x Graduate (G)</td>
</tr>
<tr>
<td>x Graduate (G)</td>
<td></td>
</tr>
<tr>
<td>Professional (P)</td>
<td></td>
</tr>
<tr>
<td>Maximum: 1</td>
<td>Grading Type: (select one)</td>
</tr>
<tr>
<td>Minimum: 1</td>
<td>x Normal grading (Blank)</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Connector: (select one)</td>
<td></td>
</tr>
<tr>
<td>x Fixed (F)</td>
<td></td>
</tr>
<tr>
<td>___ Variable (V)</td>
<td></td>
</tr>
<tr>
<td>___ Alternate (A)</td>
<td></td>
</tr>
<tr>
<td>___ To be Arranged (T)</td>
<td></td>
</tr>
<tr>
<td>Maximum Repeat: 1</td>
<td></td>
</tr>
<tr>
<td>(Total number of credit hours that may be earned, not total number of times course may be taken)</td>
<td></td>
</tr>
<tr>
<td>Term Offered: (select one)</td>
<td></td>
</tr>
<tr>
<td>x Not Specified (Blank)</td>
<td></td>
</tr>
<tr>
<td>___ Fall Only (F)</td>
<td></td>
</tr>
<tr>
<td>___ Spring Only (S)</td>
<td></td>
</tr>
<tr>
<td>___ Summer Only (M)</td>
<td></td>
</tr>
<tr>
<td>___ Fall, Spring (FS)</td>
<td></td>
</tr>
<tr>
<td>___ Intersession (I)</td>
<td></td>
</tr>
<tr>
<td>Session Duplicate:</td>
<td></td>
</tr>
<tr>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>no</td>
<td></td>
</tr>
</tbody>
</table>

Prerequisites (course must be taken prior to this course): HDFS 7060 (Research Methods for Human Development and Family Studies II)

Corequisites (course must be taken the same term of this course): HDFS 8050 Advanced Research Methods

Prerequisite with concurrency (course may be taken prior to this course or taken during the same term):
**Brief Description for Bulletin**

HDFS 8051 ADVANCED RESEARCH METHODS LAB (1). Fr., HOPS 7060. Lab designed to enhance the application of advanced research methods and data analytic strategies used in HDFS research.

Credit will not be given for both _______ and _______.

<table>
<thead>
<tr>
<th>Activities</th>
<th>Contact Group</th>
<th>Hours Indiv</th>
<th>Credit</th>
<th>Max Enroll</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st: LAB</td>
<td>1</td>
<td></td>
<td>1</td>
<td>20</td>
</tr>
<tr>
<td>2nd:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3rd:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Justification (Indicate reason for change)**

The Department of Human Development and Family Studies has decided that students will benefit from the opportunity to apply the advanced techniques learned in HDFS 8050 to substantive research questions using actual data. Because this lab focuses on the application of techniques demonstrated in the class, and not the introduction of new material or concepts, the grade for the lab will be the same as the grade for the class.

**Additional resources or resource shifting required. If none, please explain.**

None

Attach a copy of syllabus to add a new course.

To modify an existing course, attach a copy of the old syllabus and the new syllabus.

To add an honors version of an existing course or to add a distance education version of an existing, attach the existing syllabus and the syllabus for the proposed new course.

No attachment is required to delete course.

Contact Person: Dr. Joe F. Pittman

Email: pittmjf@auburn.edu

 Revised June 2008
INSTRUCTOR: Margaret K. Keiley, Ed.D.
Annex, 1st Floor
Keilemk@auburn.edu – email instead of calling

I will be setting up an email list of all the students in the class. When you have questions for me, you must send an email to this list. I will not answer email that is sent only to me.

COURSE DESCRIPTION AND GOALS:

HDFS 8050 / 8051 will be devoted to Covariance Structure Analysis (CSA), a sophisticated statistical technique this is prominent in HDFS research. CSA is useful when the researcher has multiple indicators of many constructs and wants to test the appropriateness of complex hypothesized networks among these constructs.

CSA subsumes many familiar statistical techniques – simple and multiple regression analysis and exploratory factor analysis. However, CSA also provides powerful generalizations of these basic techniques leading to methods for multivariate regression analysis, confirmatory factor analysis, higher-order factor analysis, structural equation modeling, multi-group analysis, growth modeling, and survival analysis. CSA can incorporate many outcomes and many predictors simultaneously. Additionally, and very importantly, CSA permits measurement error variance to be separated from the observed variance so that relationships among underlying true constructs can be examined. Thus, CSA blends aspects of psychometrics with traditional statistical analysis.

CSA is a powerful, general and flexible methodology. As such, CSA will add great depth to your analytic capabilities. If you intend to be at the forefront of your field, then this is one of the methodologies that you must master.

THE FUNCTION OF HDFS 8050 / 8051:

HDFS 8050 / 8051 is an applied course. In the classes and labs you will learn by observing and engaging in the authentic activities of real data analysis. The use of new statistical techniques will be “modeled” in class, and then you will apply these new techniques yourselves to real problems using real data in out-of-class homework assignments. You will learn how to interpret the results of your data analyses in words and you will be asked to communicate these interpretations clearly and concisely in writing. Learning computer skills necessary for good data analysis will be an integral part of the course. Although each presentation in class is to some extent unique, as a rough guide the presentation of each new technique or the extension of an old technique will included at least five important components:
1. We will introduce a **research question** that requires the use of the new analytic method or the extension in question.

2. We will examine one or more **real data-sets** that have been used in the past to address the research question. Comments may be offered on the utility and reasonableness of the **research design** that led to the collection of the data-sets.

3. **Statistical models** suitable for representing the substantive process in question will be proposed and their parameters examined.

4. The selected models will be **fit to the data using appropriate statistical methods and computer software**.

5. The **obtained goodness-of-fit statistics, diagnostics, and parameter estimates** will be examined critically and relevant parts of the output used to **support cogent substantive interpretations of the findings**.

In addition, you will learn some of the technical details underpinning the selected analysis, the connections of the new techniques with “classical” methods, the inter-relationships among the different techniques discussed, and ways of dealing with the limitations of each technique in practice.

**CLASS FORMAT:**

The basic format for the classes will be lecture. The format for the labs will be hands-on experience with MPlus on the computer and guidance from me. You will be expected to: 1) attend all classes and labs, arriving on time; 2) complete all assigned preparatory and background reading at the appropriate times, and 3) complete all assigned homeworks by the designated deadlines.

**REQUIRED READINGS:** The Mplus programs for each example we use and all of the readings are on J drive. These are your basic texts. I will send you powerpoints each week before class and upload them to the J drive. You must bring the appropriate Mplus programs, notes on the readings, and powerpoints to class each week. These programs, powerpoints, and readings are arranged in the order that we will use them during the semester. Except for the readings in the Mplus manual. The entire manual is on the J-drive. Read the appropriate readings from that for each week it is assigned.

**GRADING POLICY:**

<table>
<thead>
<tr>
<th>Homework #</th>
<th>Points</th>
<th>Grades</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homework #1</td>
<td>100</td>
<td>1350-1500 A</td>
</tr>
<tr>
<td>Homework #2</td>
<td>100</td>
<td>1200-1349 B</td>
</tr>
<tr>
<td>Homework #3</td>
<td>100</td>
<td>1050-1199 C</td>
</tr>
</tbody>
</table>
Homeworks:

There are 14 homework assignments. The homework assignments are structured and do not require a lot of writing. They are basically very short memos, but they will require a lot of thought. Most of the homeworks will require that you ask a research question, analyze data, answer a few questions, and interpret what you find. Your work will be graded on both the statistical analysis and the writing. Papers handed in late without my permission will be docked ½ a grade for each day late. I will select model papers to put on display via email.

Participation:

Class and lab attendance is required. You may miss one class and one lab during the semester with no penalty. If you are chronically late or miss class on a regular basis, your final grade will be lowered. Participation in class is strongly suggested.

DISABILITY ACCOMMODATION:

Students who need special accommodations in class, as provided for by the American Disabilities Act, should arrange a confidential meeting with the instructor during office hours the first week of classes - or as soon as possible if accommodations are needed immediately. You must bring a copy of your Accommodation Memo and an Instructor Verification Form to the meeting. If you do not have these forms but need accommodations, make an appointment with the Program for Students with Disabilities, 1244 Haley Center, 844-2096.

PLAGIARISM:

Please make sure you understand the university guidelines about plagiarism. Any information or opinions drawn from all sources, including other students, are to be attributed specifically to these sources. Attention to these university policies is particularly important in a course like HDFS 7060 / 7061 in which collaboration with a partner is required. If, for instance, you and
your partner work closely with another group in the class during planning, execution, or interpretation of your data analysis — a process that I encourage and fully support — you should make sure that the other group’s contribution is recognized explicitly in your written account. This then avoids the natural questions that arise when textual similarities are detected during grading. Remember, given that you have the opportunity to revise and resubmit your own work, there is no penalty to making an honest mistake. The key idea in this course is that the lessons learned in doing the work yourself are more important than getting everything right the first time.

POLICY ON INCOMPLETES:

Incompletes will be considered in rare circumstances. That is, if the student has a serious personal situation (such as long-term illness) that prevents her/him from completing the course and the last day for withdrawing from the course has passed.

**Justification for Graduate Credit:**
This class builds on the two foundational graduate research methods courses that all graduate students in HDFS will take. It is a rigorous and intensive study of basic research design and methodology. It is taught only by graduate faculty and it addresses material at an advanced level that will prepare masters/doctoral graduates to analyze, evaluate, and conduct research relevant to the range of careers they will pursue.
<table>
<thead>
<tr>
<th>Date</th>
<th>Class/Lab</th>
<th>H.O. #</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>8/19</td>
<td>Class 1</td>
<td></td>
<td>Review of Univariate and Bivariate Analyses and Introduction to CSA</td>
</tr>
<tr>
<td></td>
<td>Lab 1</td>
<td></td>
<td>Examining Data: Univariate and Bivariate Analyses</td>
</tr>
<tr>
<td>8/26</td>
<td>Class 2</td>
<td></td>
<td>Regression and Path Analysis, Part I:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Regression – Single Outcome</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Multivariate Regression Analysis</td>
</tr>
<tr>
<td></td>
<td>Lab 2</td>
<td></td>
<td>Mplus Introduction: Fitting Multiple Regression</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Fitting a Multivariate Regression Model</td>
</tr>
<tr>
<td>9/2</td>
<td>Class 3</td>
<td></td>
<td>Regression and Path Analysis, Part II:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Path Analysis</td>
</tr>
<tr>
<td></td>
<td>Lab 3</td>
<td></td>
<td>Conducting a Path Analysis</td>
</tr>
<tr>
<td>9/9</td>
<td>Class 4</td>
<td></td>
<td>Modeling Mediation</td>
</tr>
<tr>
<td></td>
<td>Lab 4</td>
<td></td>
<td>Conducting a Mediation Analysis</td>
</tr>
<tr>
<td>9/16</td>
<td>Class 5</td>
<td></td>
<td>Modeling Moderation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Modeling Linked Data</td>
</tr>
<tr>
<td></td>
<td>Lab 5</td>
<td></td>
<td>Conducting a Moderation Analysis</td>
</tr>
<tr>
<td>9/23</td>
<td>Class 6</td>
<td></td>
<td>Introducing Measurement Error</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Observed, True, and Error Variance</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Reliability</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Structural Equation Models (SEM), Part I</td>
</tr>
<tr>
<td></td>
<td>Lab 6</td>
<td></td>
<td>Fitting a Simple Structural Equation Model (SEM)</td>
</tr>
<tr>
<td>9/30</td>
<td>Class 7</td>
<td></td>
<td>Structural Equation Models (SEM), Part II</td>
</tr>
<tr>
<td></td>
<td>Lab 7</td>
<td></td>
<td>Fitting a More Complex SEM</td>
</tr>
<tr>
<td>10/7</td>
<td>Class 8</td>
<td></td>
<td>Confirmatory Factor Analysis (CFA), Part I</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>First-Order and Second-Order CFA</td>
</tr>
<tr>
<td></td>
<td>Lab 8</td>
<td></td>
<td>Fitting a CFA</td>
</tr>
<tr>
<td>10/14</td>
<td>Class 9</td>
<td></td>
<td>Between-Wave Analysis with Longitudinal Data:</td>
</tr>
<tr>
<td>Date</td>
<td>Class</td>
<td>Topic</td>
<td>Lab</td>
</tr>
<tr>
<td>--------</td>
<td>--------</td>
<td>--------------------------------------------</td>
<td>--------</td>
</tr>
<tr>
<td>10/21</td>
<td>Class 10</td>
<td>Multi-Group Analysis</td>
<td>Lab 10</td>
</tr>
<tr>
<td></td>
<td>Lab 9</td>
<td>Auto-Regressive and Cross-Lagged Models</td>
<td></td>
</tr>
<tr>
<td>10/28</td>
<td>Class 11</td>
<td>Growth Modeling, Part I</td>
<td>Lab 11</td>
</tr>
<tr>
<td></td>
<td>Lab 11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11/4</td>
<td>Class 12</td>
<td>Growth Modeling, Part II</td>
<td>Lab 12</td>
</tr>
<tr>
<td></td>
<td>Lab 12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11/11</td>
<td>Class 13</td>
<td>Growth Modeling, Part III</td>
<td>Lab 13</td>
</tr>
<tr>
<td></td>
<td>Lab 13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11/18</td>
<td>Class 14</td>
<td>Growth Modeling, Part IV</td>
<td>Lab 14</td>
</tr>
<tr>
<td></td>
<td>Lab 14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12/2</td>
<td>Class 15</td>
<td>Odds and Ends</td>
<td>Lab 15</td>
</tr>
<tr>
<td></td>
<td>Lab 15</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
In the first class I will review the importance of conducting univariate and bivariate analysis, prior to conducting any more sophisticated statistical analysis, and how to conduct these analyses. Then, I will overview the course and introduce statistical concepts important for the understanding and manipulation of covariance structures. To demonstrate the relationship between correlation and covariance, and their links to variance, we will examine data made available by Raudenbush and Chan (1992) about the tolerance of deviant behavior and exposure to deviant behavior during adolescence.

Next, using the data that were introduced last week, Raudenbush and Chan's (1992) study of the tolerance and exposure to deviant behavior during adolescence, we will fit regression models using both SAS and MPLUS, and compare the results. In the process of learning how to fit multiple regression models using MPLUS, we will examine two different methods of data input. First, we will read in raw data from an external file (TOLER_RAW.DAT). Second, we will place the estimated covariance matrix and a vector of sample means from an external file. The data set, TOLER_MC.DAT, contains these values for the three variables -- TOL_13, EXP_13, and FEMALE -- for 168 adolescents.

Readings

   This article is a fairly decent overview of covariance structure analysis, written in simple language. Kline's statistical family tree is a context for structural equation modeling with latent variables that you might find useful. He also presents some of the "mis-uses" of any CSA statistical package (LISREL, AMOS, MPLUS, CALIS). The appendix contains an annotated reading list of articles that Kline considers important for a good understanding of structural equation modeling.

   This, too, is a pretty good, although a bit more technical, overview. Francis discusses the steps that are involved in model fitting, including how to assess whether the model fits. He also presents some of the "pros" and "cons" of structural equation modeling.

   Hoyle and Smith have written a fairly decent introduction to measurement and structural equation modeling that is not too technical. They differentiate classes of models, such as observed and latent variable models, and describe how these models can be used to investigate group differences and differences over time.
HDFS 8050: HOMEWORK #1
Examples of Covariance Structure Analysis and Questions to Ponder

Task

Using PSYCH INFO, identify several published research papers in a substantive area that is of interest to you in which the authors have used some form of CSA to conduct their analysis. The paper that you select must be an account of primary research, it cannot be a review or an opinion piece, and it must be an application of covariance structure analysis to quantitative data. It should include all the accouterments of a research paper: a short literature review, a statement of the research questions, a description of the research design (sample, measures, procedures, etc.) and data analyses, a summary of the findings (with appropriate tables and figures), and a discussion. Download or find those papers and bring a copy of them to class next week.

Type or print an accurate citation for the selected article at the head of its title page, using the citation style appropriate for the journal from which it was taken. On the back of the title page, write your name. Include the photocopied paper with your memo, it will be returned to you later.

Written Memo

Based on the readings that you have done from the Class #1 reading list, classify the study you have selected as either confirmatory factor analysis, structural equation modeling, path analysis, growth modeling, survival analysis and, in one paragraph, explain why you believe your categorization to be accurate.

Questions you should be prepared to discuss next week:

1. From the articles that you have read, what do the various authors mean when they refer to covariance structure models as "causal" models? What seems to you to be an appropriate interpretation of the word "causal" in the term, "causal models"?

2. Explain how an exogenous variable differs from an endogenous variable.

3. From your understanding of the authors you have read thus far, what are a few of the advantages and disadvantages of structural equation modeling?
Before we begin today's topic, we will first review the material from the readings for Class I and consider the questions you were to be prepared to discuss.

Then, we will discuss multivariate regression -- regression with more than one outcome. To introduce these ideas we will examine data from a study that was reported by Davis and Franzoi in 1986, replicating the results of another study they reported in 1985. The data set, LONELY_F.DAT, contains the data for 161 Females from the 1986 article. The authors were conducting a multi-group analysis. We will only examine the model fit for the females. The data for the 173 males is in the data set, LONELY_M.DAT. The hypothesized model that the authors were testing related loneliness with self-disclosure to peers, mother, and father. Each form of self-disclosure was believed to be related to other factors, such as warmth, self-consciousness, and perspective-taking. We will not be fitting the model that the authors hypothesized, but a less complex one to illustrate multivariate regression with MPLUS. In our fitted model, we will predict each form of self-disclosure and then examine whether the residuals associated with each of these forms of self-disclosure are related.

Data References

1. Franzoi, S.L. & Davis, M.H. (1985). Adolescent self-disclosure and loneliness: Private self-consciousness and parental influences. Journal of Personality and Social Psychology, 48, 768-780. This article delineates the hypothesized model they replicate in the later article. You could use the data from this article to compare the model fit for the females (or males) across time. (You would have to enter the data yourself!)

2. Davis, M.H. & Franzoi, S.L. (1986). Adolescent loneliness, self-disclosure, and private self-consciousness: A longitudinal investigation. Journal of Personality and Social Psychology, 51, 595-608. Just skim this article concentrating on pages 595-602 where they discuss the replication of their original model from 1985. In the remainder of the article the authors investigate several alternative hypothesized models. You could use the data for the males to fit the same model that we fit in class and compare the results for males and females.

Readings


Note – Make sure to familiarize yourself with Chapters 14-19 that describe all of the basic Mplus commands and language. Whenever you have a question – look here!
Preliminary Reading


Task

Fit the hypothesized multivariate regression model depicted in Figure 1 on page 192. The data (means, SDs, correlations) are in the file, `MATH.DAT`, in the order of the variables listed in Table 1 on page 195. To read in the data, you must list **ALL** the variables, but in the `USEVAR` line you will only use the variables that are shown in Figure 1 on page 192. The variable you will leave out is the full scale of math self-efficacy; the first 3 variables listed are all sub-scales of this full scale. The full scale is not used in the fitted model.

When you write your program, make sure to think about what the default settings are that you might want to change for your fitted model. For example, looking at the authors’ hypothesized model, should you allow the residuals of the outcome variables to covary or not? What about the predictor variables? Should they covary with each other? If so, why? If not, why not?

After you have made the decisions about the questions in paragraph two above, fit your model and compare your $R^2$ statistics for each fitted regression model to those quoted in the article on page 195, and your parameter estimates for each fitted regression model to those shown in Tables 2 and 3, pages 195-196. Then print out your program and output. Staple them together and submit to us along with a cover sheet on which you answer the following questions:

1. How did you make the decisions you made about the questions in paragraph 2 above?

2. Does your hypothesized model fit the data? Describe the evidence you have for your answer.

3. Which paths in the hypothesized models appear to be the most crucial in the prediction of each of the two outcomes? Describe the evidence for your answer.

4. What story would you tell about your findings?
In this class, we will extend our notions of multivariate regression to the case of path analysis, in which there are multiple predictors, multiple outcomes, and outcomes as predictors of other outcomes. We will test hypotheses about complex networks of relationships among observed variables. Using the data described in Data Reference (below) about the school performance of African-American adolescents, we will test whether measures of adolescents' ability, sense of importance of school, and feelings of ethnicity predict their grades and their engagement in school.

Readings


Another fairly decent overview of CSA and SEM. In addition, Fergusson discusses the uses of CSA in studying linked data.


Data References


When reading this article concentrate on pages 21-37, including figures 1-4, but ignoring Tables 3-8. In this section of the article they report their use of path analysis, rather than multiple regression, to answer their questions. Note on page 29, near the bottom, they indicate that they included demographic characteristics in their models to control for age, SES, and sex. We cannot do that without more information, so our estimates are slightly different from theirs. We are using the data for the African-American Public School adolescents (*ACHIEV_BP.DAT, n=135*).

The data for the Catholic School African-American adolescents (*n=60*) and for 100 Catholic School White Adolescents is in *ACHIEV_MG.DAT*. We will use these data later after we learn about multi-group analysis to conduct a multi-group analysis!
HDFS 8050: HOMEWORK #3
Conducting Path Analyses

Preliminary Reading for Data Analysis


Task #1

The data are in **DRINK.DAT**, in the order in which they appear in Table 1 on page 338. The estimated means are first, the standard deviations second, and the estimated correlation matrix is last. You are to fit their hypothesized model in figure 1 (page 337). As they describe it in figure 1, their hypothesized model is indicated by the solid lines; it does not include the broken lines. Please note that in this hypothesized model and their final fitted model on page 339, there is a latent construct, positive mother-child relationship, which is measured by two observed variables, **warmth** and **positive communication**. We are asking you to fit a simpler model, one that does not include latent constructs. Therefore, think of positive mother-child relationship as just **warmth**, and drop positive communication from the path model. The data file, however, includes both of these variables, so your “variable name” statement must name 7 variables altogether, although your “use variable” statement will contain 6 variables.

If you include an “Standardized” in your output statement in your program, you will obtain a fully standardized solution. You can then check your standardized estimates against the ones that they report in their final model on page 339.

When you are satisfied with your program and output, print them out. Staple them together, and submit them to us, along with a cover sheet on which you ponder the following questions:

1. Note that their final fitted model (page 339) has 2 fewer paths than their hypothesized model -- the one that you fit (page 337). How do you think they decided on this model as their final, “best-fitting” model?

2. Can you think of a way in which you might be able to test if one model is a “better fitting” model than another? Try doing this by fitting one addition model and report the results of your test.

3. What story would you tell about your final fitted model and what it means?

Task #2: Read the readings for next week and write your answers for these questions in the MacKinnon chapters:

**Chapter 1: Questions:** 1.7, 1.8, 1.11, 1.12

**Chapter 2: Questions:** 2.6 (bring in a copy of the article)
HDFS 8050: CLASS #4
Mediation

We will spend some time talking about the readings for today and the questions that you answered for Task #2 in the homework. This discussion will help us to understand the importance of considering tests of mediation in your own work and just how those tests should be conducted. Mediation analysis is the tool that will help us understand the mechanisms involved in developmental and intervention/treatment processes.

The data that we will use to examine mediation Courtney Pearce collected for her master’s thesis. She was interested in the attachment relationship of females to their fathers. She collected data from 172 female undergraduate students. We are going to ask a different research question than she did for her thesis.

Readings

A classic and important article. Must be read.


You will fit and interpret a mediation path analysis in which there are three variables. The outcome is social competence in adolescence (SocComp) and the two predictors are quality of attachment early in life (Attach) and affect regulation in later childhood (AffReg). We only have a small sample of 50 adolescent males. The data are in the file “mediation data.dat” on the J drive in the data folder. The variables are in this order in the file: Attach AffReg SocComp.

The theoretical model is that affect regulation during childhood mediates the quality of early attachment in its effect on later social competence in adolescence. You will fit the appropriate series of models to determine if the proposed model is substantiated. That is, you will fit 4 models, one for each step of the mediation model, plus one more to test if mediation exists in the final model. In this last model, you will constrain one path to “0” and conduct a delta chi-square test.

Your analysis should be similar to what we did in class. Please submit a memo that includes:

1. A figure of your final fitted model that includes the parameter estimates (with estimated correlations in parentheses), the $R^2$ statistic, a legend of $p$-values, and an appropriate title.

2. A table illustrating delta chi-square test that you conducted.

3. A brief substantive interpretation of the final fitted model explaining what it actually means.
We will first review the ideas related to moderation and how it is tested by creating interaction terms that are then included in a series of models. Then to illustrate moderation analyses, we will use secondary data from the Urban Poverty and Family Life Survey of Chicago (UPFLS; Wilson, 1987) that Cassandra Kirkland downloaded and used for her master’s thesis. We will be answering a different question than she did. The sample for this survey was drawn from a stratified probability sample of census tracts that consisted of individuals from impoverished areas in the city of Chicago. At least 20% of the residents reported yearly household incomes that fell below the federal poverty line. The analytic sample that Cassandra used included a total of 738 African-American respondents of whom 33% were male and 67% female. Ages ranged from 18 to 43 years old at the time of interview.

In addition, I will present an example from the Glanton House data that will illustrate how to manage having respondents who are linked to each other. One of the assumptions for the statistical models that we fit is that the data are independent. When you are examining people who are linked (e.g., couples, family members, children in classrooms), you must model the relationships among your variables with the linked respondents in the same statistical model. Another name for these kind of models are hierarchical models. That is, at the “individual” level, the relationships among the variables for each person are modeled, but at the “group” level, these relationships among variables are modeled for each person WITHIN THE SAME model.

Reading:

This is a fairly complex article about predicting multiple outcomes where observations are not independent. While you skim this reading, focus on pages 33-34, 36-38, 40-41, and 46 to the end.

For this memo, you will conduct a moderation path analysis using the data collected by Ashley Anders for her master’s thesis. Ashley was interested in a very understudied construct, sexual communication anxiety; that is, a person's level of fear in discussing her sexual relationship with a significant other. The sample for her study consisted of 123 undergraduate women who completed several scales, but we will only be using two of them. The data are in “Anders_Thesis.dat” and the order of the variables is: n_part, in_nd_ex, and phys_sat

OUTCOME: IN_ND_EX: Inhibition of Need Expression, Sexual Communication Scale (Davis, et al., 2006) was used to measure sexual communication anxiety. A series of 18 statements were answered on a 5-point Likert scale with responses ranging from "not at all like me" to "completely like me." Participants responded to statements such as "If my partner is unhappy with our sex life, I would rather not know about it," "If we're having problems with sex, I tend to let them build up for a long time before I say anything," and "If I want to have sex, I'm more likely to ‘hint’ around that I'm interested instead of just asking outright." Cronbach's alpha for this scale was a=.83 in this sample.

PREDICTOR: PHYS_SAT: Physical Satisfaction Scale (Davis, et al., 2006) was used to measure sexual satisfaction. Cronbach's alpha was a=.83 and participants rated the degree to which they agreed with 11 statements presented in the scale such as "I am satisfied with the quality of sex in my relationship," "I am usually able to satisfy my sexual needs in my relationship," and "I would like to be able to get more physical satisfaction out of sex." Statements were answered on a 5-point Likert scale.

PREDICTOR: N_PART: Number of sexual partners which ranged from none to 20.

The proposed theoretical model for this analysis is that inhibition of need expression is related to the level of physical satisfaction, but that may vary depending on the level of partners a young woman has had.

Conduct this moderation path analysis following the example we did in class and prepare a short memo in which you include:

1. A fitted path diagram that represents your final fitted model on which you present your estimated parameters, the R² statistic, the fitted equation, a p-value legend, and an appropriate title.

2. Create a fitted plot to illustrate your findings.

3. A brief substantive interpretation of the final fitted model explaining the moderating effect.
Part of what we would like to be able to do in our data analyses is to distinguish between observed and "true" variables (constructs). So far, in this course, we have been examining hypothesized relationships only among observed variables. In this class, we begin to incorporate the idea that there are "true" variables (latent variables or constructs), that underlie these observed variables.

We will begin by re-examining the Raudenbush and Chan data on tolerance and exposure to deviant behavior in adolescence, but this time we will fit a latent variable model, rather than an observed variable model. To do this, we must incorporate information about measurement error and reliability into the analysis.

Readings

A very useful article about the difficulties that often arise in fitting SEMs, what the causes of these difficulties might be, and what practical things you might do to remedy the situations. He includes problems such as non-convergence, matrices that are not positive definite, under-identification, negative variances, and improper correlations, etc.

Dare I say, a classic! And easy to read.
HDFS 8050/8051: CSA  Dr. Margaret K. Keiley  

HDFS 8050: HOMEWORK #6  
Dealing with Measurement Error: Using Reliability to Disattenuate the Findings

Reading


Task #1

When we fit the hypothesized model shown in **solid lines** in figure 1 (page 337) for Homework #3, we assumed that these observed variables were infallible, measured without measurement error. In reality, all measured variables contain random error and we can account for that in our analyses if we possess appropriate additional information. With this in mind, we are asking you to fit two further models using the Dumka and Roosa data and then compare them. Refit the model from Homework #3, but distinguish observed variables (M_DRINK, F_DRINK, FAMSTRES, M_WARM, M_MARADJ, and M_PERADJ) from latent variables (M_Drink, F_Drink, Famstres, M_Warm, M_MarAdj, and M_PerAdj) by following the procedure we used in class in “8050 Class 6 Measurement Error in SEM.ppt.” Also, for this model in Task #1, set all of the error variances to zero. The fit of this model should correspond to the model you fit for homework last week.

Task #2

Using reliability information from the following table, refit the model in Task #1 substituting the appropriate proportion of observed variance in each measure that is error variance, just as we did in class in Handout V.1b. Assume that FAMSTRES is measured without error.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td>M_Drink</td>
<td>.71</td>
</tr>
<tr>
<td>F_Drink</td>
<td>.85</td>
</tr>
<tr>
<td>M_Warm</td>
<td>.76</td>
</tr>
<tr>
<td>M_MARADJ</td>
<td>.92</td>
</tr>
<tr>
<td>M_PERADJ</td>
<td>.91</td>
</tr>
</tbody>
</table>
Task #3

Submit your output from the first two tasks, along with a cover sheet on which you present a table illustrating how you determined the error variance settings in Task #2 (include reliability, observed variance, and error variance computed). Also, comment on any changes that you noticed between the two models. If there were many changes, focus on only one of them.

1 As in Homework #3, please note that in this hypothesized model and their final fitted model on page 339, there is a latent construct, positive mother-child relationship, which is measured by two observed variables, warmth and positive communication. We are asking you to fit a simpler model, one that does not include latent constructs just yet. Therefore, think of positive mother-child relationship as just warmth, and drop positive communication from the path model. The data file, however, includes both of these variables, so your “variable name” statement must name 7 variables altogether, although your “use variable” statement will contain 6 variables.
In this class, to introduce the notion of using multiple indicators of a construct (rather than estimates of reliability) as a way of getting a handle on measurement error, we will examine data about parental support and the psychological adjustment of young adults presented in an article by Holahan, Valentiner, and Moos (Data Reference #1 below).

Next, using data from a study about young adolescent girls reported in Data Reference #2, we will begin to distinguish the structural and measurement components of latent variable models. We will fit a taxonomy of nested models and conduct \( \chi^2 \) tests to assess improvement in fit and to test equality constraints across paths.

Readings


   Concentrate on the beginning (490-494) and end (501-503) of this article where the authors caution about modifying hypothesized models.


   Breckler discusses some of the flaws that appear in published applications of SEM. He particularly focuses on the problem of "equivalent" models, models that account for the same observed covariances equally well. Can we really say a model is confirmed, if this is the case? In an appendix, he lists many articles that report on SEM analyses.


Data References


   We will not be fitting their entire model, just pieces of the “time-lag” model. The data are contained in SUPPORT.DAT.

2. Conger, R.D., Conger, K.J., Elder, G.H., Lorenz, F.O., Simons, R.L., and Whitbeck,

This week we will be fitting the model on page 214. Later we will return to fit some of the more complex models. The data for this week are contained in STRESS.DAT.
HDFS 8050: HOMEWORK #7
Fitting a Structural Equation Model

Preliminary Reading


Task #1

The data for this homework are in BERNDT.DAT which contains the estimated intercorrelations among a subset of the variables listed in Table 2 (INV, SVAL, SCHL, SABL, FABL, GRDS, TRK), as well as their estimated standard deviations and means, in that order. Fit the structural model displayed in figure 2 on page 324. For this first model, set the metric of the latent constructs by setting one factor loading equal to 1 for each factor. Obtain the non-standardized solution and a standardized solution for the fitted model.

Task #2

Submit your programs and output to us along with a memo in which you answer the following questions:

1. Does the model fit? Summarize, in one paragraph, the various pieces of evidence that you used to make your determination.

2. In the unstandardized solution, what is the estimated variance of the endogenous construct, Achievement?

3. What is the estimated intercorrelation of the exogenous constructs?

4. Briefly describe the meaning of the final fitted model. What is the story?
HDFS 8050: CLASS #8
Confirmatory Factor Analysis

In this class, we will separate out the measurement model from the structural model for special attention and we will explore the measurement properties of the indicators using confirmatory factor analysis (CFA).

Then, using Benson and Bandalos’s work on a new scale, Reactions to Tests (Data Reference #1), we will fit a second-order factor analysis.

Readings


   The authors clearly delineate 1st and 2nd order measurement models, then report an unsuccessful confirmatory factor analysis.


Data Reference


   These authors reserved half their sample in order to cross-validate their final fitted model. Cross-validation is always a good practice. We will be using the data from their calibration sample, TESTS_C.DAT.
Preliminary Reading

Task #1
Fit the hypothesized model in figure 1, page 238, using the dataset BELONG.DAT, which contains the estimated means, standard deviations, and correlations in the order they are presented in that figure. Remember, Mplus will create the covariance matrix. The authors only used the correlation matrix, and possibly the means, to fit their models, therefore your parameter estimates based on a covariance matrix will be different from theirs, although your chi-square statistics should be quite similar. It is always best to analyze the covariance matrix, when possible, but you can analyze a correlation matrix.

Also note, that the path diagram presented by these authors is a bit misleading. They have indicated that residuals exist for the two latent constructs. A rather poor graphic display! So just ignore those numbers and those arrows. There are no residuals for this model as those two latent constructs (Social Connectedness, Social Assurance) are NOT being predicted by anything!

Task #2
Fit a second model in which the two latent factors are uncorrelated and compare this to the first model that you fit. Conduct the appropriate delta chi-square test to determine which model is your “best” model.

Task #3
Submit your programs and output to us along with a memo that summarizes the following:

1. The goodness-of-fit evidence for each model.
2. The delta-chi-square test results.
3. The estimated reliabilities, factor-indicator correlations, and factor-factor correlations (if appropriate) in table form, for what you consider to be the best fitted model. Comment on the items in terms of their estimated reliabilities and correlations with the latent factors.
4. Comment on the correlation of the two factors.
5. What would you say that this measure measures?
In this class we will use covariance structure analysis to examine longitudinal data in order to answer questions about the “stability” of an outcome variable over time or about cascading influences of predictors on outcomes over time. The models we will fit in this class are often called simplex models, longitudinal panel models, or cross-lagged longitudinal models.

Often, as in this case, such questions are addressed by simultaneously predicting status at time 2 by status at time 1, and status at time 3 by status at time 2, and so on. While such analyses do not truly capitalize on the full richness of longitudinal data, they do incorporate all the data simultaneously. Using raw data from an article by Farrell about adolescent alcohol use (Data Reference #1), we will fit a simple structural equation model illustrating this methodology.

Readings
   You could conduct a path analysis and/or fit a structural equation model with these data, which are in TURN.DAT. The authors fit an SEM by following a procedure which they outline in the first paragraph on page 858. One wonders if they were confirming or exploring models, since they seem to fit a great many models to the same data, and they make many modifications to each of these models.

   As you read through this article, take note of the kinds of constraints that the authors impose on their hypothesized model to reflect what they think happens over time. The data are in HEALTH.DAT


Data Reference
   This week we will use these data to examine stability over time, later we could use them in multi-group analyses, or perhaps growth modeling. The raw data for the males are in FAR_BOYS.DAT, for the females the data are in FAR_GIRL.DAT. The raw data for males and females are in FAR_BOTH.DAT
HDFS 8050: HOMEWORK #9
Fitting a Structural Equation Model with Longitudinal Data

Preliminary Reading


Task #1

Reproduce figure 1, page 290, for either mothers or fathers, and on it display appropriate notation. Using Mplus, fit the hypothesized model. The data for the mothers is in PARENT_M.DAT, and for the fathers is in PARENT_F.DAT. You can choose which data set to use in your work. These data files contain the estimated means, standard deviations, and correlations (in that order) of the variables listed in Table 3 (in that order). The sample size is 359.

Please note: Since only one measure loads on the Difference In Sibling Delinquency latent construct, you will need to set the error variance of this measure to zero. Of course, you will have to do this twice, because the same is true for the later year.

Task #2

Discuss the goodness-of-fit of your final model and present your findings on a path diagram and interpret the findings. What is the story? Make sure to submit your output and program along with your memo.
In this class, we will examine differences in covariance structures across groups. Using data about life stressors, personal and social resources, and depression (Data Reference #1), we will investigate how individuals who experience high stress versus those who experience low stress differ when we fit the hypothesized model for each of these groups, simultaneously, using the multi-group method. First, after noting invariance across the measurement models, we will fit a taxonomy of nested models to examine differences in the structural model.

Readings


   Although this article is a bit technical, it is certainly comprehensive. You will find an overview of CFA (1st and 2nd order), then some discussion of multi-group analysis of CFAs (pages 46-47 and 54-61).


   This is a fairly easy multi-group path analysis. The data for the males are in ABUSE_M.DAT, for the females, ABUSE_F.DAT.


Data Reference

HDFS 8050: HOMEWORK #10
Multi-Group Analysis

Data References


Take a look at our final program for the path analysis in this article (Class 3 powerpoint and Mplus programs). We will be fitting the same model but across two groups – 100 European-American Catholic School Adolescent and 60 African-American Catholic School Adolescents. Revise the program we used to make it a program to examine multiple groups. The data are in ACHIV_MG.DAT but in a different order – discrim ability imp_sch ethnic grades engage. The first group in the data file is the EA adolescents and the second group is the AA ones.

Task #1
Reproduce the hypothesized model we used as the final one (Slide 14) in your program for the multiple group analysis and fit the models. We are not very concerned about model fit; more important for this homework to focus on the differences across the two groups. But, comment on model fit.

Task #2
Fit a series of nested multi-group models (see below), beginning with a model in which there are no constraints across groups, except that the model is the same. At each step, constrain the regression parameters one at a time, holding those equal that have already been tested and found to be invariant as you test the next parameter.

Task #3
Report the results of your multi-group analysis by commenting on model fit, then presenting your final fitted model in two path diagrams, one for EA Catholic Adolescents and one for AA Catholic Adolescents.

Task #4
What is the story that your final fitted multiple group model tells?

HDFS 8050: CLASS #11
Introducing Individual Growth Modeling
In this class, we will learn about the use of individual growth modeling to represent change over time. We will distinguish between intra-individual change -- how each individual is changing -- and inter-individual differences in change -- how the change over time differs across individuals. We will examine how to summarize each person’s growth trajectory by a small number of individual growth parameters and then how to examine the relationships between these individual growth parameters and possible predictors of them.

We will return to an example similar to the one with which we began this course. We have data on 168 adolescents’ tolerance of deviant behavior at five different time points -- age 11, 12, 13, 14, 15. We will fit a model that describes intra- and inter-individual differences in the change in the tolerance of deviant behavior over time. The data are contained in GROWTH.DAT.

**Reading**


**HDFS 8050: Homework #11**

**Fitting an Individual Growth Model**

**Preliminary Re-Reading**

You will examine adolescents’ use of alcohol over three time periods, at the beginning of grade 7, and at the end of grade 7 and grade 8. You might want to think about these time periods as being, Time 0, Time .75, and Time 1.75.

The raw data for these 1122 adolescents is contained in FARRELL.DAT, which includes information on 34 variables in fixed format so you must use the format statement in your program (FORMAT IS 1x,F1.0,1x,33F1.0;)

You will only need 3 variables for this analysis, T1_Drunk, T2_Drunk, and T3_Drunk, but you must first read in all the data. The data to be read are in this order (belongs in your NAMES ARE statement):

FEMALE T1_ANG1 T1_ANG2 T1_ANG3 T1_OFDRK T1_PRDRK T1_FRDRK T1_PDRK T1_BEER T1_WINE T1_DRUNK T1_LIQOR T2_ANG1 T2_ANG2 T2_ANG3 T2_OFDRK T2_PRDRK T2_FRDRK T2_PDRK T2_BEER T2_WINE T2_DRUNK T2_LIQOR T3_ANG1 T3_ANG2 T3_ANG3 T3_OFDRK T3_PRDRK T3_FRDRK T3_PDRK T3_BEER T3_WINE T3_DRUNK T3_LIQOR;

These three variables assess drinking on a 6-point scale at the three time periods. They are described on pages 478-479 of the article.

The tasks for this homework are to fit a linear growth model and determine:

1. Does the model fit? What evidence do you have for this?

2. Does change exist in the frequency of being drunk over 7th and 8th grades?

3. What is the fitted average trajectory for the change?

4. Is there inter-individual variance in the intercepts and slopes that could be predicted by another variable?

5. Show the prototypical fitted plot for the average adolescent in this population and relate the story it tells.
In this class, we will return to an example similar to the one with which we began this course. We have data on 168 adolescents’ tolerance of deviant behavior at five different time points -- age 11, 12, 13, 14, 15. We will fit models that describe how interindividual differences in the change in the tolerance of deviant behavior over time are related to gender and exposure to deviant behavior at age 11. The data are contained in GROWTH.DAT.

Readings


HDFS 8050: HOMEWORK #12
Adding Predictors to Single Domain Growth Model

This homework is a continuation of the one you did last week. Using the observed variables, T1_DRUNK, T2_DRUNK, T3_DRUNK, that assess drinking on a 6-point scale at the three time periods, at the beginning of grade 7, and at the end of grade 7 and grade 8, you will now determine if the interindividual differences in initial status and rate of change in drinking behavior are predicted by the gender of the adolescent (FEMALE; 1=female, 0=male) and the level of angry threat at the beginning of grade 7 (T1_ANG1). These variables, along with the others in the study are described on pages 478-479 of the article.

Task #1

Fit your growth model from last week’s homework, again, but add predictors that will determine if individual growth depends on gender and pressure from peers to drink at the beginning of 7th grade.

Task #2

Discuss your findings in a brief summary paragraph, including evidence of model fit and whether gender and angry threats at the beginning of 7th grade do predict initial status and rate of change. In order to do this, you will probably need to conduct one or two \( \Delta \chi^2 \) tests. Indicate on your output the variance of the rate of change, the variance of initial status, and their correlation.

Task #3

Construct pseudo-\( R^2 \) statistics for the parameter estimates of true intercept and true slope and prototypical plots of the average growth trajectories for males and females at high and low levels of angry threats at the beginning of 7th grade.

Task #4

From this analysis, what is the story you would tell?
HDFS 8050: CLASS #13
Multiple Domain Growth Modeling

In this class we will examine growth in multiple domains simultaneously. Being able to do this allows the researcher to examine several different types of questions. We could study how change in the domains of externalizing and internalizing behavior are related to each other for children from kindergarten to high school and what predicts that. In addition, we can examine how trajectories of marital adjustment of married partners relate to each other during therapy. The example we will look at today is the latter. We will be using the MFT data from Glanton House to examine how change in one partner’s dyadic adjustment are related to his/her partner’s dyadic adjustment.

Readings:


Continuing with the Farrell data, choose two domains of interest to you to model change over 7th and 8th grade. Look at the article to determine what each scale means, then fashion a research question that you can answer by using multiple domain growth modeling.
In this class we fill in gaps that we have discovered as we moved through this semester. One of the odds and ends that we will definitely cover will be the controversy over goodness-of-fit in structural equation modeling. The other odds and ends that we cover will be up to you!

**Reading**


HDFS 8050: HOMEWORK #14
Multiple Domain Growth Modeling

Add predictors to your multiple domain model from last week.