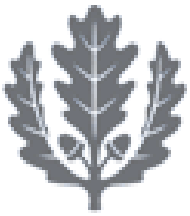
Modern Modeling Methods Conference

MAY 21–22, 2013



University of Connecticut

Neag School of Education

2013 Modern Modeling Methods Conference

Sponsored by the ***Educational Psychology Department*** and the ***Neag School of Education***, University of Connecticut

Special thanks to Dr. Del Siegle, the chair of the Educational Psychology department, and Dr. Thomas DeFranco, Dean of the Neag School of Education.

In addition, many thanks to Joanne Roberge, Cheryl Lowe, Jessica Flake, Sarah Newton, John Madura, Robbin Haboian-Demircan, Kate Copeland, and conference services for providing administrative and logistical support for the conference.

Also, thank you to all of the keynote speakers and concurrent presenters for making this wonderful program possible. A special thanks is due to Bengt for promoting the conference on the Mplus website and during his workshops this year.

Finally, thank you to all of the 2013 Modern Modeling Methods conference attendees for coming and being a part of the third annual M^3 conference! I hope to see you all back in Storrs on May 20-21, 2014 for the fourth annual Modern Modeling Methods conference. Proposals for concurrent sessions will be due January 15, 2014, and can be submitted online at our website, www.modeling.uconn.edu.

D. Betsy McCoach, Ph.D.
2013. Chair, Modern Modeling Methods Conference
Associate Professor, Measurement, Evaluation, and Assessment Program
Educational Psychology Department, Neag School of Education, UCONN

Monday, May 20th: Pre-conference Workshop

Multilevel Structural Equation Modeling with Complex Applications

Kristopher Preacher
Vanderbilt University

This workshop will introduce multilevel structural equation modeling (MSEM) as a general approach to fitting both simple and complex models to hierarchically structured data. After a brief review of structural equation modeling (SEM) and multilevel modeling (MLM), we will turn to the historical development of MSEM methods, important advantages of MSEM, and assessment of model fit in MSEM. Subsequently, the workshop will cover a variety of model specifications as special cases of MSEM. These will include: standard multilevel models, single-level structural equation models, multilevel exploratory and confirmatory factor analysis, multilevel path analysis, multilevel structural models with latent variables, multilevel mediation analysis, and multilevel reliability estimation. Models will be presented in several formats — path diagrams, equations, and software syntax. Data and Mplus syntax for all of the examples will be provided. It is not required to bring Mplus on a laptop to attend the workshop -- attendees will still benefit from a comprehensive set of slides, and syntax that they can take home and use later. Those who do bring a laptop with Mplus should be sure that they have the base package + the multilevel add-on or the base package + combination add on to be able to run the examples.

8:00 - 9:00 am	Breakfast and Registration	**Laurel Hall Atrium**
9:00 - 10:30 am	Session 1: Overview of SEM, MLM and Mplus	**Laurel Hall 101**
10:30 - 10:50 am	Break	**-**
10:50 - 12:00 pm	Session 2: MLM in Mplus and an overview of MSEM	**Laurel Hall 101**
12:00 - 1:30 pm	Lunch (on your own)	**-**
1:30 - 3:25 pm	Session 3: Specific models and Mplus syntax	**Laurel Hall 101**
3:25 - 3:40 pm	Break	**-**
3:40 - 5:00 pm	Session 4: Advanced models and topics	**Laurel Hall 101**

Kristopher Preacher, Ph.D., is Assistant Professor in the Quantitative Methods program at Vanderbilt University. His research concerns the use (and combination) of structural equation modeling and multilevel modeling to model correlational and longitudinal data. Other interests include developing techniques to test mediation and moderation hypotheses, bridging the gap between substantive theory and statistical practice, and studying model evaluation and model selection in the application of multivariate methods to social science questions. He serves on the editorial boards *of Psychological*

Methods, Journal of Counseling Psychology, Communication Methods and Measures, and *Multivariate Behavioral Research.*

Monday, May 20th

Welcome Reception

All participants in the pre-conference workshop and/or the general conference are invited to attend a welcome reception after the preconference workshop in the Mansfield room of the Nathan Hale Inn. Appetizers and beverages will be available.

5:30 - 7:30 pm	Reception	**Nathan Hale Inn, Mansfield Room**

Tuesday, May 21st

8:00 – 8:45 am	Breakfast and Registration	**Laurel Hall Atrium**
8:45 – 10:15 am	Keynote Address: **Linda Collins, Ph.D.** *Optimizing behavioral interventions: An integration of methodological perspectives from the behavioral and engineering sciences*	**Laurel Hall 102**
10:15 – 10:30 am	Break	**-**
10:30 – 12:00 pm	**CONCURRENT SESSION #1**	
	<u>Combined Session 1.1: Quantifying Uncertainty in Structural Equation Modeling</u> Accounting for Population Uncertainty in Covariance Structure Analysis by **Hao Wu, and Michael W. Browne**	
	Parameter Uncertainties in Structural Equation Modeling: Quantification and Implications by **Taehun Lee, Robert MacCallum, and Michael Browne**	**Laurel Hall 102**
	Profile Likelihood-Based Confidence Regions in Structural Equation Models by **Jolynn Pek, and Hao Wu**	
	<u>Combined Session 1.2: Longitudinal Fit and Invariance</u> Assessment of Fit Indices for the Determination of Measurement Invariance in Longitudinal Models by **Corbin T. Quick, Aaron J. Boulton, Alexander M. Schoemann, and Todd D. Little**	
	Are You Measuring Variability or Trait Change? On the Role of Measurement (Non)Invariance in Latent State-Trait Analyses by **Christian Geiser, Brian T. Keller, Ginger Lockhart, Michael Eid, and Tobias Koch**	**Laurel Hall 201**
	Models of Variability versus Models of Trait Change: How Well Do Fit Indices Distinguish Between the Two? by **Brian T. Keller and Christian Geiser**	
	<u>Combined Session 1.3: Missing Data Methods</u> A comparison of imputation strategies to missing ordinal item scores by **Wei Wu, Craig Enders, and Fan Jia**	
	Using principal component analysis (PCA) to obtain auxiliary variables for missing data estimation in large data sets by **Waylon J. Howard and Todd D. Little**	**Laurel Hall 205**
	Missing Data Methods for Confounding Variables in Marginal Structural Models by **Shu Xu and Vanessa Watorek, New York University**	

	Combined Session 1.4: Analyzing Single Subject, Diary, and Intensive Longitudinal Data Daily Diary Data: Effects of Cycles on Inferences by **Yu Liu and Stephen G. West** Modeling Cyclical Patterns in Daily College Drinking Data with Many Zeroes by **David Huh, Debra L. Kaysen, and David C. Atkins** The Misspecification of the Covariance Structures in Multilevel Models for Single-Case Data: A Monte Carlo Simulation Study by **Mariola Moeyaert, Maaike Ugille, John M. Ferron, S. Natasha Beretvas, and Wim Van den Noortgate (Presented by Patricia Rodriguez)**	**Laurel Hall 206**
	Combined Session 1.5: Profile Analysis and Non-normal SEM Modeling Configural Patterns in Latent Variable Profiles Associated with Endogenous Criteria by **Mark Davison, Ernest Davenport, and Yu-Feng Chang** Moderated Profile Analysis: Comparing Criterion-Related Predictor Variable Patterns across Populations by **Mark Davison, and Ernest Davenport** Investigation of Type I Error Rates of Three Versions of Robust Chi-Square Difference Tests by **Victoria Savalei, Jenny Chuang, and Carl Falk**	**Laurel Hall 301**
	Combined Session 1.6: Propensity Score Analysis The Impact of Measurement Error on Propensity Score Analysis: An Empirical Investigation of Fallible Covariates by **Eun Sook Kim, Patricia Rodriguez de Gil, Jeffery D. Kromrey, Aarti P. Bellara, Rheta E. Lanehart, Tyler Hicks, and Reginald S. Lee** Propensity Score Matching (with multilevel data) Using SPSS and R by **Felix Thoemmes, and Wang Liao**	**Laurel Hall 302**
12:00 – 1:30 pm	Lunch	**SU Ballroom**
1:30 – 3:00 pm	**CONCURRENT SESSION #2**	
	Session 2.1: Open-source Modern Modeling Software: The R Package Lavaan by **Yves Rosseel** This is a double session which runs through concurrent sessions 2 and 3	**Laurel Hall 102**
	Combined Session 2.2: Extensions to Growth Models	

Multiple-Indicator Latent Growth Curve Models: An Analysis of the Second-Order Growth Model and Two Less Restrictive Alternatives by **Jacob Bishop and Christian Geiser** The Autoregressive Latent Trajectories Model: An Alternative Approach to the Analysis of Panel Data by **Chris Schatschneider** Recovery of Individual Trajectories in Heterogeneous Samples Using Longitudinal Latent Profile Analysis (LLPA) by **Veronica T. Cole & Daniel Bauer**	**Laurel Hall 201**
<u>Combined Session 2.3: Stepwise Approaches to Latent Variable Modeling</u> Group means as explanatory variables in multilevel models by **Jouni Kuha, Anders Skrondal and Stephen Fisher** The bias-adjusted three-step approach to latent class modeling with external variables by **Zsuzsa Bakk, Daniel Oberski, and Jeroen K. Vermunt** Three step Latent Transition Analysis by **Bengt Muthen and Tihomir Asparouhov** Three-step estimation method for discrete micro-macro multilevel models by **Margot Bennink, Marcel A. Croon, and Jeroen K. Vermunt**	**Laurel Hall 202**
<u>Combined Session 2.4: New Developments in the Analysis of Incomplete Data Symposium</u> Utilizing Hyper Priors in Multiple Imputation for Multivariate Normal Data **by Valerie Pare** Handling Data with Three Types of Missing Values: A Simulation Study by **Jen Boyko** Approaches to Multiple Imputation in large data sets by **Chantal Larose**	**Laurel Hall 205**
<u>Combined Session 2.5: Measurement Invariance and Differential Item Functioning</u> A Simulation Study of a MIMIC-based Strategy for Detecting Items with DIF by a School Covariate by **Shonte Stephenson, and Sophia Rabe-Hesketh** Testing for Measurement Invariance with Respect to an Ordinal Variable by **Edgar Merkle, Jinyan Fan, and Achim Zeileis** Model Invariance Testing Under Different Levels of Invariance by **Holmes Finch and Brian French**	**Laurel Hall 206**
<u>Combined Session 2.6: Comparing Analyses Predicting</u> Simple Versus Residualized Change Scores	

	Toward Understanding Discrepant Results from Predicting Residualized versus Simple Change Scores by **Robert E. Larzelere, Ronald B. Cox, Jr., and Sada J. Knowles**	**Laurel Hall 202**
	Modeling Change as a Residual Difference Score or as a Simple Gain Score: Further Clarifying the Paradox by **Todd D. Little, Alexander M. Schoemann, and Matthew W. Gallagher**	
	Comparing Change Scores with Lagged Dependent Variables in Models of the Effects of Parents' Actions to Modify Children's Problem Behavior by **David Johnson**	
	<u>Combined Session 2.7: Advances in Structural Equation Modeling</u> An Efficient State Space Approach to Estimate Univariate and Multivariate Multilevel Regression Models by **Fei Gu, Kristopher J. Preacher, and Wei Wu**	**Laurel Hall 302**
	A Place for Nonlinear Structural Equation Modeling; What Insight Does It Give Us? by **Jonathan Brewster and Kathleen Buse**	
	<u>Combined Session 2.8: Extensions to Mediational Analyses</u> Flexible Mediation Analysis in the Presence of Non-linear Relations by **Beatrijs Moerkerke, Tom Loeys, Olivia De Smet, Ann Buysse, Johan Steen, and Stijn Vansteelandt**	**Laurel Hall 305**
	Modeling Indirect Effects with Multimethod Data by **Ginger Lockhart, Christian Geiser, Hui Qiao, Jacob Bishop, Martin Schultze, and Herbert Scheithauer**	
	Testing Mediation the Way it Was Meant to be: Changes Leading to Changes then to Other Changes. Dynamic Mediation Implemented with Latent Change Scores by **Emil Coman, Eugen Iordache, and Maria Coman**	
3:00 – 3:15 pm	Break	**-**

3:15 – 4:45 pm	**CONCURRENT SESSION #3**	
	Session 2.1 CONTINUED: Open-source Modern Modeling Software: The R Package Lavaan by **Yves Rosseel** This is a double session which runs through concurrent sessions 2 and 3	**Laurel Hall 102**
	Combined Session 3.2: Longitudinal Mixture Models Using Mixture Latent Markov Models for Analyzing Change with Longitudinal Data by **Jay Magidson, Statistical Innovations Inc.** A Framework for Investigating the Performance of Latent Growth Mixture Models by **Paul Dudgeon**	**Laurel Hall 201**
	Combined Session 3.3: Planned Missing Data Designs for Longitudinal Research Planned Missing Data Designs with Small Samples: How Small is Too Small by **Alexander M. Schoemann, Fan Jia, E. Whitney G. Moore, Richard Kinai, Kelly Crowe, and Todd D. Little** Assignment Methods in Three-Form Planned Missing Designs by **Terrence D. Jorgensen, Alexander M. Schoemann, Brent McPherson, Mijke Rhemtulla, Wei Wu, and Todd D. Little** Planned Missing Designs to Optimize the Efficiency of Latent Growth Parameter Estimates by **Fan Jia, Mijke Rhemtulla, Wei Wu, Todd D. Little** Two-method Planned Missing Designs for Longitudinal Research by **Mauricio Garnier-Villarreal, Mijke Rhemtulla, and Todd D. Little**	**Laurel Hall 202**
	Combined Session 3.4: Multidisciplinary Perspectives on Fit Comparing the Evidence for Categorical Versus Dimensional Representations of Psychiatric Disorders in the Presence of Noisy Observations: A Monte Carlo study of the Bayesian Information Criterion and Akaike Information Criterion in latent variable models by **Michael Hallquist, Thomas Olino, and Paul Pilkonis** The Effects of Local Item Dependence on the Sampling Distribution of the Q3 Statistic by **William P. Skorupski, and Sukkeun Im** Bootstrapping SEM Goodness of Fit and Confidence Intervals by **Craig M. Krebsbach, and Lisa L. Harlow**	**Laurel Hall 205**

	Combined Session 3.5 Multilevel Survival Analysis Modeling Microsocial Heterogeneity using multilevel survival analysis by **Mike Stoolmiller** Continuous Time Analysis of Panel Data: An Illustration of the Exact Discrete Model by **Aaron J. Boulton, and Pascal R. Deboeck**	**Laurel Hall 206**
	Combined Session 3.6: Extensions to Factor Analysis Multilevel Factor Analysis by Model Segregation: The Surprising Necessity for Robust Statistics with Normal Data by **Jonathan Schweig** The Influence of Parceling on the Implied Factor Structure of Multidimensional Data by **Brooke Magnus, and Yang Liu** To Thin or Not To Thin? The Impact of Thinning Posterior Markov Chains on Parameter Estimation in Latent Trait Models by **Jared K. Harpole, and William P. Skorupski**	**Laurel Hall 301**
	Session 3.7: Introduction to Quantile Regression for Social Science Researchers Introduction to Quantile Regression for Social Science Researchers by **Jessica Logan and Yaacov Petscher**	**Laurel Hall 302**
5:00 – 7:00 pm	**POSTER SESSION & RECEPTION** Please join us in the Student Union Ballroom to visit the 26 posters. The reception includes light appetizers and an open bar.	**SU Ballroom**

Wednesday, May 22nd

8:00 – 8:45 am	Breakfast and Registration	**Laurel Hall Atrium**
8:45 – 9:45 am	**CONCURRENT SESSION #4**	
	<u>Combined Session 4.1: Meta-analysis</u> Individual Participant Data Meta-Analytic Modeling Techniques 1993 - 2012: A Methods Review by **Samantha A. Russo, and Tania B. Huedo-Medina**	**Laurel Hall 201**
	Measurement Harmonization in Individual-Participant-Data Meta-Analytic Modeling by **Tania B. Huedo-Medina, Francisco Galindo-Garre, and Maria Dolores Hidalgo**	
	<u>Combined Session 4.2: Statistical Disclosure Limitation</u> Application of Mixture Models in Statistical Disclosure Limitation by **Anna Oganian**	**Laurel Hall 202**
	Assessing the Privacy of Randomized Vector-valued Queries to a Database Using the Area Under the Receiver-operating Characteristic Curve by **Gregory Matthews**	
	<u>Session 4.3: Bayesian Analysis of Dynamic Item Response Models in Educational Testing</u> Bayesian Analysis of Dynamic Item Response Models in Educational Testing by **Xiaojing Wang, James O. Berger, and Donald S. Burdick**	**Laurel Hall 205**
	<u>Session 4.4 Bias in Missing Data Problems due to Inclusion of Auxiliary Variables</u> Bias in Missing Data Problems due to Inclusion of Auxiliary Variables by **Felix Thoemmes**	**Laurel Hall 206**
	<u>Combined Session 4.5: New Developments in Model Selection Procedures</u> F-tests with Incomplete Data for multiple regression set-up by **Ashok Chaurasia**	**Laurel Hall 302**
	Model Selection for Correlated Predictors of Correlated Outcomes with Applications in Genetic Association Studies by **Elizabeth D. Schifano**	
	<u>Combined Session 4.6: Actor Partner Interdependence Model and Extensions</u> Modeling Psychological Subgrouping using Multiple Demographic Composition Variables by **Randi L. Garcia**	
	Interdependent Households Preferences - Case III/APIM Approach by **Adam Sagan**	**Laurel Hall 305**

	Combined Session 4.7: Multilevel Analysis Doubly-Diminishing Returns: An Empirical Investigation of the Impact of Sample Size and Predictor Prevalence on Point and Interval Estimates from Two-Level Linear Models by **Bethany A. Bell, Jason A. Schoeneberger, Elizabeth A. Leighton, Stan Haines, Mihaela Ene, Whitney Smiley, and Jeffrey D. Kromrey** Analyzing Clustered Data with Single-Level Models: Are Robust Standard Errors an Adequate Solution? by **Dan McNeish and Jeffrey Harring**	**Laurel Hall 306**
9:45 – 10:00 am	Break	**-**
10:00 – 12:00 pm	Keynote Address: **Judea Pearl, Ph.D.** *What on earth are we modeling? Data or Reality? Reflections on structural equations, external validity, heterogeneity and missing data*	**Laurel Hall 102**
12:00 – 1:30 pm	Lunch	**SU Food Court**
1:30 – 3:00 pm	**CONCURRENT SESSION #5**	
	Symposium Session 5.1: Causal Mediation Recent advances in causal mediation by **Felix Thoemmes** Misclassification of a binary mediator – effects and remedies by **Linda Valeri** What Mathematics Tells us About the Mediation Formula and Why it Matters for Policy Analysis and Scientific Understanding by **Judea Pearl**	**Laurel Hall 102**
	Combined Session 5.2: Longitudinal Models Fitting Nonlinear Latent Growth Curve Models with Individually-varying Time Points by **Sonya Sterba** Evaluating the Prediction of Growth Factors Under Misspecification of Functional Form by **Stephanie Lane and Patrick Curran**	**Laurel Hall 202**
	Combined Session 5.3: Statistics and Modeling Data Analysis Strategies for High Dimensional Social Science Data: What to do when there are more variables than subjects by **Holmes Finch, Maria Hernandez Finch, David E. McIntosh, and Lauren Moss** Extending the Robust Means Modeling Framework by **Alyssa Counsell, Matthew Sigal, Philip Chalmers, and Robert A. Cribbie**	**Laurel Hall**

205

Using Monte Carlo simulations to Understand Probabilities and Modeling: Bringing causality into the teaching of introductory statistical modeling by **Emil Coman, Maria Coman, Eugen Iordache, Lisa Dierker, and Russell Barbour**

Combined Session 5.4: Multiple Raters

Approximate Measurement Invariance in Rater-mediated Assessments: A Random Item Effects Item Response Model for Measuring Teaching Quality with Classroom Observations by **Ben Kelcey, Dan McGinn, and Heather Hill**

Laurel Hall 206

Level-Specific and Multilevel Reliability by **Joseph A. Olsen**

Reliability and Validity by the Multiple Indicator by Multiple Trait by Multiple Source by Multiple Occasion Model: Application to an ADHD Rating by **Scale by G. Leonard Burns, Mateu Servera, and Christian Geiser**

Combined Session 5.5: Educational and Health related applications

Laurel Hall 301

Realistic Models for School-based Longitudinal Sociometric Data: Multilevel Cross-classified Poisson and Negative Binomial Approaches by **Richard A. Faldowski, and Heidi Gazelle**

A Mathematical Evaluation of the Effect of Disclosure on HIV Transmission Rate in Men who Have Sex with Men by **Ann A. O'Connell, Sandra J. Reed, and Julianne M. Serovich**

Patterns of Drug use Measured using Latent Class and Latent Transition Analysis with Covariates in a representative sample of Incarcerated adults in Puerto Rico by **Rafael R. Ramirez, Carmen Rivera Medina, Jose Noel Caraballo, Jose Ruiz Valcarcel, Glorimar Caraballo Correa, and Carmen Albizu**

Ordinal Regression Analysis: Fitting Stereotype Logistic Regression Models to Educational Data by **Xing**

Combined Session 5.6: Understanding Patterns, Variations, And Trends In Adolescent And Young Adult Smoking: Opportunity For New Insights Or Source For Headaches?

Laurel Hall 305

Smoking patterns in first year college students: Trends, correlates and outcomes by **Bettina B. Hoeppner and Nancy P. Barnett**

Decisions about Covariate Specification in Time-Varying

	Effects Models: Does it Matter? by **Jennifer S. Rose, Arielle S. Selya, Lisa C. Dierker, Donald Hedeker, and Robin Mermelstein** Nicotine Dependence-Varying Effects of Smoking Events on Momentary Mood Changes among Adolescents by **Arielle Selya, Nicole Updegrove, Lisa Dierker, Jennifer Rose, Xianming Tan, Donald Hedeker, Runze Li, and Robin Mermelstein** A Bivariate Location-Scale Mixed-Effects Model with Application to Mood Variation among Youth Smokers by **Oksana Pugach, Donald Hedeker, Robin and Mermelstein**	
3:15 – 4:45 pm	Keynote Address: **Bengt Muthen, Ph.D.** *Late-breaking News: Some Exciting New Methods!*	**Laurel Hall 102**

Thursday, May 23rd: Post-conference Workshop

Advances in Modeling Using Mplus

Bengt Muthen

University of California Los Angeles

The use of latent variables is a common theme in many statistical analyses. Continuous latent variables appear not only as factors measured with errors in factor analysis, item response theory, and structural equation modeling, but also appear in the form of random effects in growth modeling, components of variation in complex survey data analysis and multilevel modeling, frailties and liabilities in survival and genetic analyses, latent response variables with missing data, priors in Bayesian analysis, and as counterfactuals and potential outcomes in causal analysis. In addition, categorical latent variables appear as latent classes in finite mixture analysis and latent transition analysis (Hidden Markov modeling), latent trajectory classes in growth mixture modeling, and latent response variables with missing data on categorical variables. All these features are covered by the general latent variable modeling framework of Mplus.

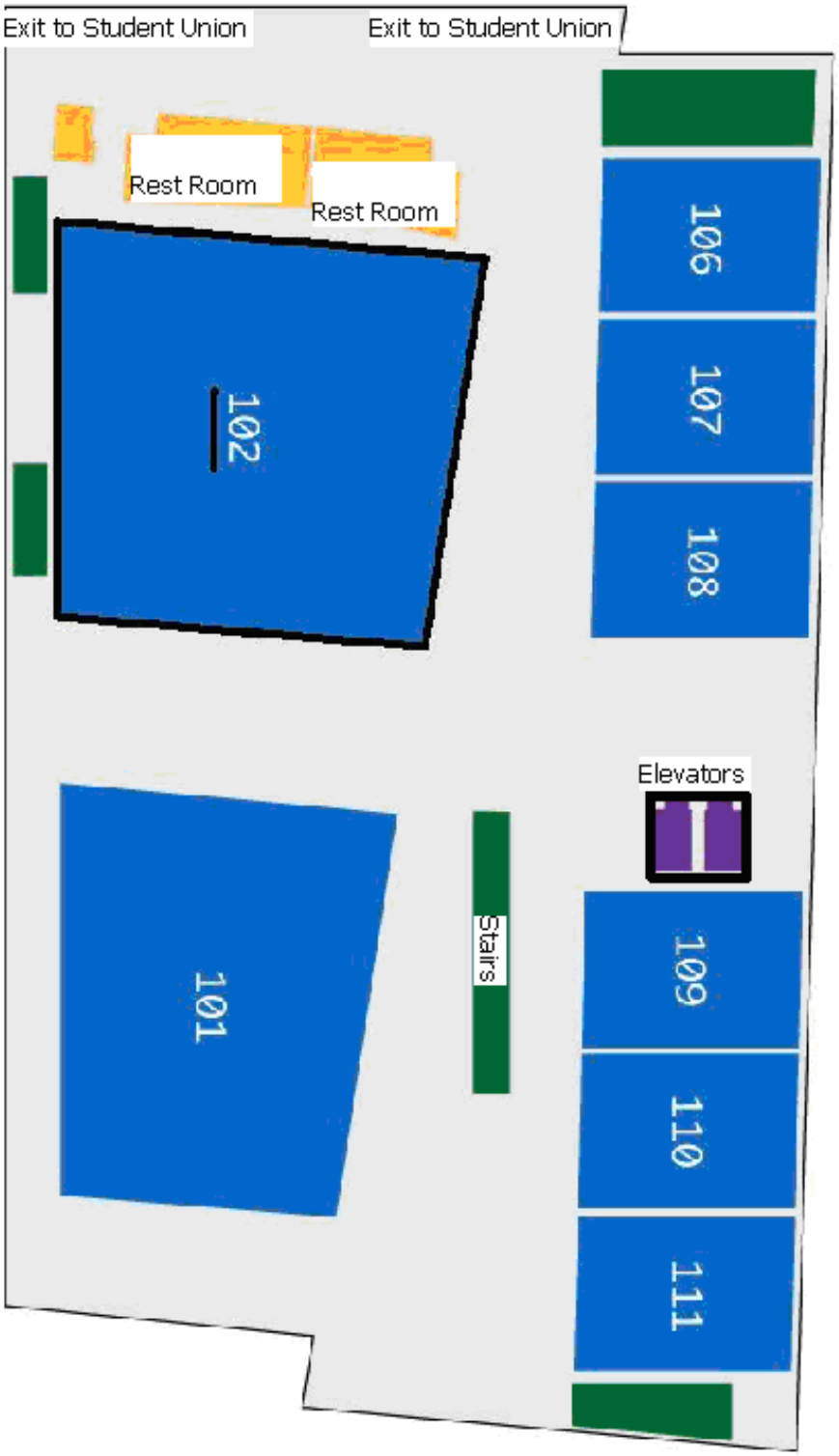
Understanding the unifying theme of latent variable modeling provides a way to break down barriers between seemingly disparate types of analyses. Researchers need to be able to move freely between analysis types to more easily answer their research questions. To provide answers to the often complex substantive questions, it is fruitful to use latent variable techniques to combine different analysis types. Mplus is unique in its ability to accomplish such combinations. This one-day workshop discusses new types of latent variable analyses added to Mplus Version 7 and 7.1. See the Mplus website for more information about Mplus 7.1.

8:00 – 8:30 am	Coffee and Registration	**Laurel Hall Atrium**
8:30 - 10:00 am	Session 1	**Laurel Hall 102**
10:00 - 10:15 am	Break	**-**
10:15 - 11:45 pm	Session 2	**Laurel Hall 102**
11:45 - 1:15 pm	Lunch (on your own)	**-**
1:15 - 2:45 pm	Session 3	**Laurel Hall 102**
2:45-3:00 pm	Break	**-**
3:00 - 4:30 pm	Session 4	**Laurel Hall 102**
4:30-4:45 pm	Break	**-**
4:45-6:00 pm	Bonus Session	**Laurel Hall 102**

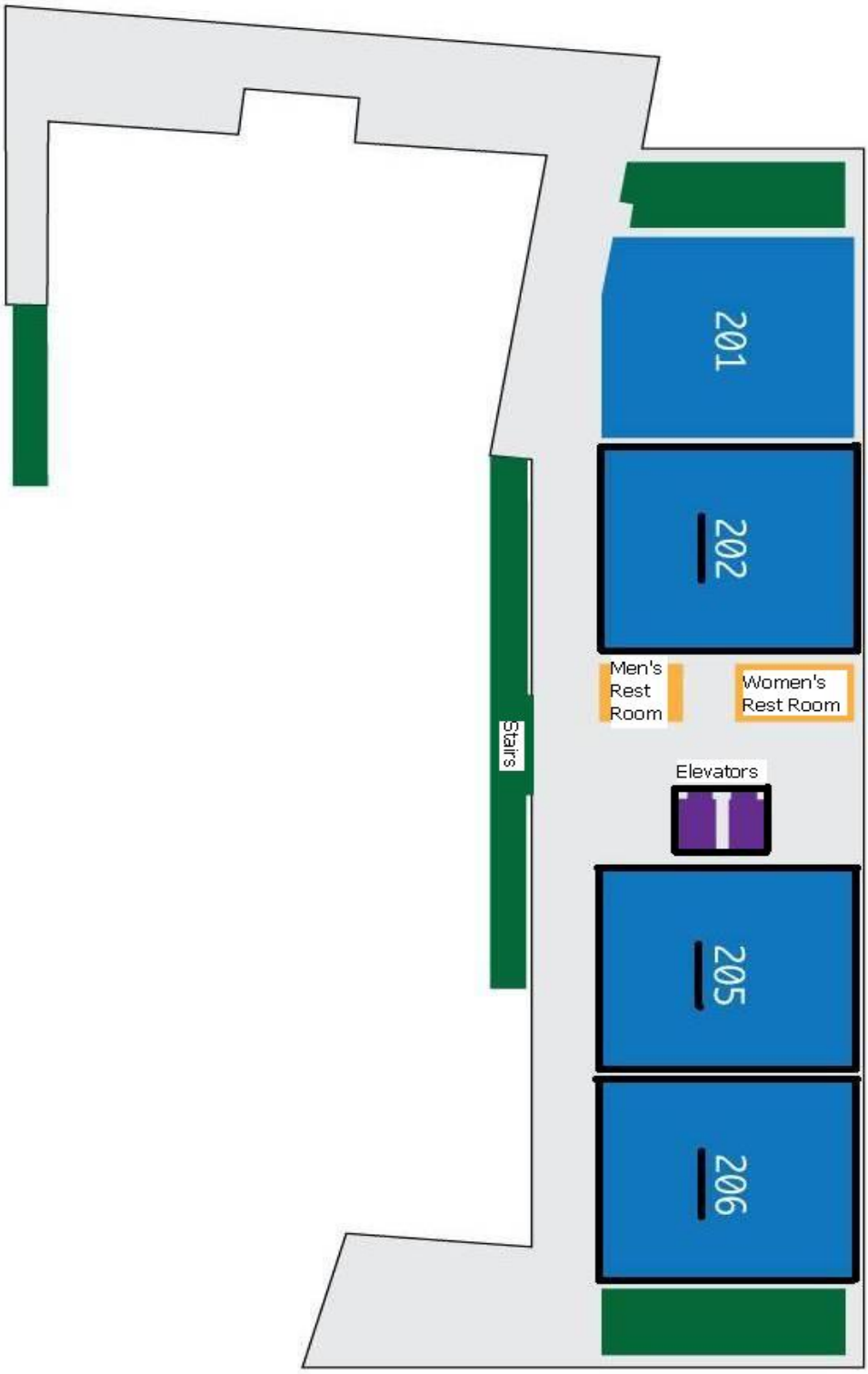
Bengt Muthén obtained his Ph.D. in Statistics at the University of Uppsala, Sweden and is Professor Emeritus at UCLA. He was the 1988-89 President of the Psychometric Society and the 2011 recipient of the Psychometric Society's Lifetime Achievement Award. He has published extensively on latent variable modeling and is one of the developers of the *Mplus* computer program, which implements many of his statistical procedures. Dr. Muthén's research interests focus on the development of applied statistical methodology in areas of education and public health. Education applications concern achievement development while public health applications involve developmental studies in epidemiology and psychology. Methodological areas include latent variable modeling, analysis of individual differences in

longitudinal data, preventive intervention studies, analysis of categorical data, multilevel modeling, and the development of statistical software (namely *Mplus*!).

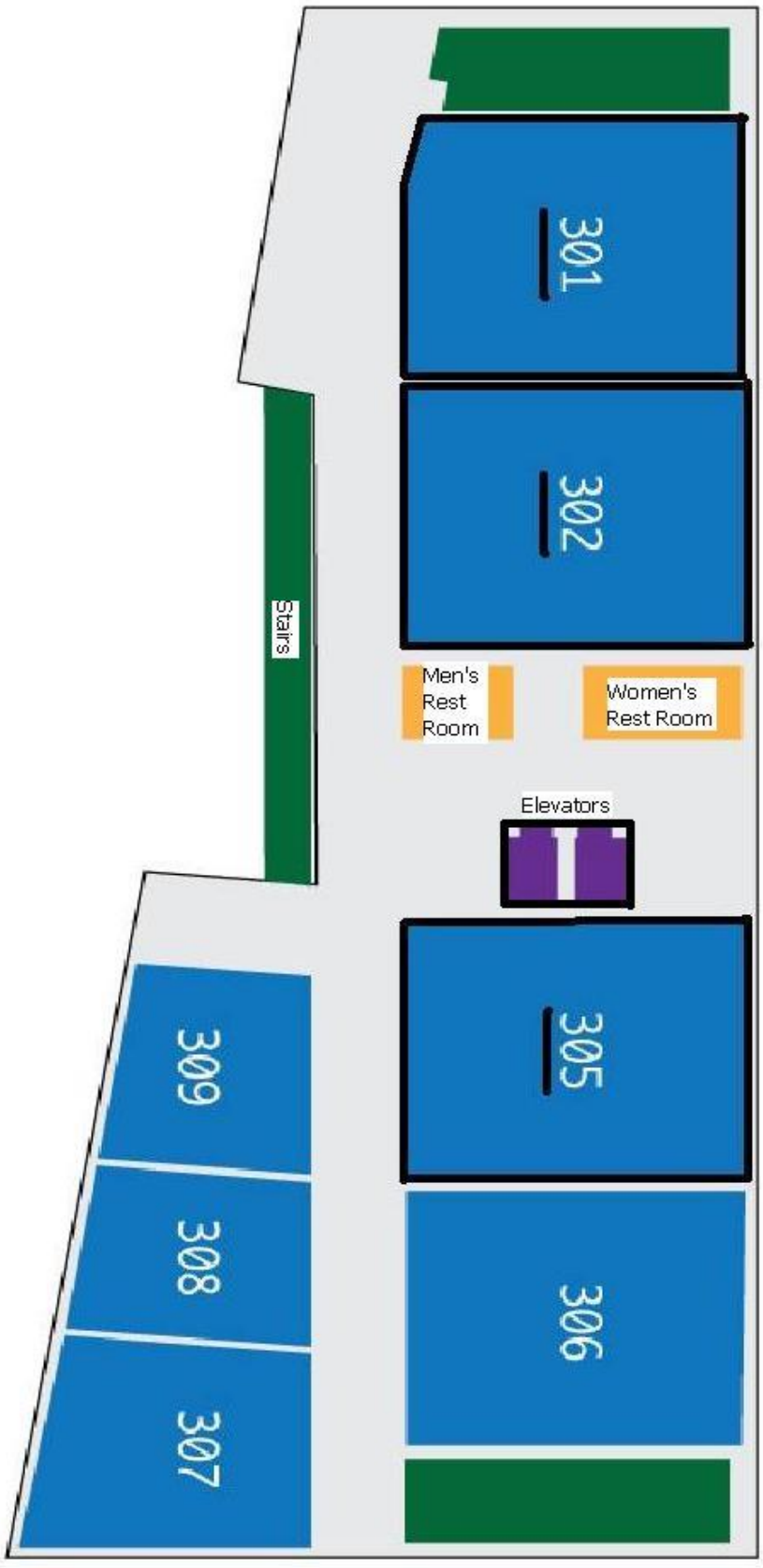
Laurel Hall (Classroom Building) Floor Plans
1st Floor



Laurel Hall (Classroom Building) Floor Plans
2nd Floor



Laurel Hall (Classroom Building) Floor Plans
3rd Floor



Tuesday, May 21st, 2013

Breakfast and Registration

Laurel Hall Atrium 8:00–8:45am

Opening Keynote Address

Laurel Hall 102 8:45am–10:15 am

Optimizing Behavioral Interventions: An Integration of Methodological Perspectives from the Behavioral and Engineering Sciences

Linda Collins, Ph.D

Linda Collins is Director of The Methodology Center and Professor of Human Development and Family Studies at The Pennsylvania State University. A Fellow of the American Psychological Association and the Association for Psychological Science, Dr. Collins has published numerous books and journal articles in her areas of research including *Latent Class and Latent Transition Analysis: With Applications in the Social, Behavioral, and Health Sciences,* published by Wiley in 2010. Her interests include experimental and non-experimental design and models for longitudinal data.

Concurrent Session 1 10:30am–12:00pm

Combined Session 1.1: Quantifying Uncertainty in Structural Equation Modeling
Location: Laurel Hall 102

Accounting for Population Uncertainty in Covariance Structure Analysis

Hao Wu, and Michael W. Browne

We present an approach to quantifying errors in covariance structures in which model error, identified as the process underlying the discrepancy between the population and the structured model, is explicitly modeled as a random effect with a distribution, and the dispersion parameter of this distribution to be estimated gives a measure of misspecification. Analytical properties of the resultant procedure are investigated and the measure of misspecification is found to be related to the RMSEA. The asymptotic sampling distributions of the estimators are also established. Simulations validate the analytical results and demonstrate the importance of accounting for the variations in the parameter estimates due to model error.

Parameter Uncertainties in Structural Equation Modeling: Quantification and Implications

Taehun Lee, Robert MacCallum, and Michael Browne

In this paper we explore and quantify the uncertainties arising from estimating unknown model parameters, a type of uncertainty inherent in the process of fitting any statistical model to empirical data. Specifically we show that there may be infinitely many parameter sets within a given model structure that are virtually equally acceptable as alternative explanations of the data. These alternative set of parameter values, termed fungible estimates due to Waller (2008), may come from very different regions in the parameter space, being very different from the optimal solution. To achieve this goal, we propose methods for identifying fungible estimates using structural equation modeling (SEM) as the context.

Profile Likelihood-Based Confidence Regions in Structural Equation Models

Jolynn Pek, and Hao Wu

Structural equation models (SEM) are often used to evaluate a network of directional and non-directional relationships between latent and measured variables. Typically, point estimates of key parameters or functions of parameters are reported and interpreted using Wald tests and their commensurate confidence intervals. In the context of finite sample size and using SEM to assess multiple parameters, we recommend the construction of likelihood-based confidence regions for k parameters instead. We present a framework for computing exact profile-likelihood based confidence regions for SEM and emphasize the value of profile likelihood-based confidence regions over conventional Wald-type confidence regions. We also highlight the added value of examining confidence regions for multiple parameters over reporting several confidence intervals. An empirical example involving a latent variable mediation model serves to illustrate the construction and use of profile-likelihood based over Wald-type confidence regions. We conclude with some recommendations for the application of profile likelihood-based confidence regions in SEM.

Combined Session 1.2: Longitudinal Fit and Invariance
Location: Laurel Hall 201

Assessment of Fit Indices for the Determination of Measurement Invariance in Longitudinal Models

Corbin T. Quick, Aaron J. Boulton, Alexander M. Schoemann, and Todd D. Little

While several previous studies have evaluated alternative fit indices for the determination of measurement invariance (MI) in multiple-group models (e.g., Cheung & Rensvold, 2002; Chen, 2007; Meade, Johnson, & Braddy, 2008), to our knowledge, none have investigated MI in longitudinal models. We discuss a number of respects in which longitudinal models differ from multiple-group models, and explain how this could influence testing MI. We discuss the results of a Monte Carlo simulation investigating MI in longitudinal models. For each of 1,920 invariant conditions and 7,776 conditions with lack-of-invariance (LOI), 200 item-level data sets were generated and fit to a null model (as per Widaman & Thompson, 2003), configural invariant model, metric invariant model, and scalar invariant model. Nested model difference scores (X^2, 2df, and AFIs) were calculated and 95th percentile values across replications within conditions computed. ?AFI cut-off criteria are provided specifically for longitudinal MI determination, and additional findings regarding longitudinal MI discussed.

Are You Measuring Variability or Trait Change? On the Role of Measurement (Non)Invariance in Latent State-Trait Analyses

Christian Geiser, Brian T. Keller, Ginger Lockhart, Michael Eid, and Tobias Koch

Latent state-trait (LST) models are designed to measure variability around a fixed set-point or trait, whereas latent growth curve (LGC) models focus on more long-lasting and potentially irreversible trait changes. Although researchers generally agree that testing measurement invariance (MI) is an important step in longitudinal analyses, such tests are rarely explicitly performed in LST applications. In this article, we examine (1) whether MI across time should routinely be established in LST models and (2) what the consequences of measurement non-invariance are for the interpretation of the results. We show that researchers testing the hypothesis that the longitudinal dynamics of a construct are best described by a variability (state-trait) process need to (a) include a mean structure and (b) establish strong factorial invariance in their LST analysis. Using simulations, we demonstrate that LST models with non-invariant parameters can mask trait changes and obscure the fact that a LGC or latent change model may be more appropriate for the data. Furthermore, we show that the inappropriate application of LST models to trait change data can lead to bias in the estimates of consistency and occasion-specificity, which are typically of key interest in LST analyses. Guidelines for testing MI in LST analyses are provided.

Concurrent Session 1 Continued 10:30am–12:00pm

Models of Variability versus Models of Trait Change: How Well Do Fit Indices Distinguish Between the Two?

Brian T. Keller and Christian Geiser

Longitudinal structural equation modeling (LSEM) is often used to examine the stability, variability, and/or change in social science constructs over time. This study presents a simulation study in which we investigated how well common model fit indices distinguish between models of trait change versus models of variability. The results suggest that popular fit statistics (chi-square test and approximate fit indices) can fail to differentiate between variability and change models under certain conditions. We discuss implications of this issue for applied longitudinal research.

Combined Session 1.3: Missing Data Methods
Location: Laurel Hall 205

A Comparison of Imputation Strategies To Missing Ordinal Item Scores

Wei Wu, Craig Enders, and Fan Jia

The current study uses Monte Carlo method to compare multiple strategies to impute missing ordinal item scores in the context where the item score will aggregated to scale scores for further analyses. The examined strategies include: a) using linear regression to impute the ordinal missing data by treating them as continuous (multivariate normal strategy), b) using methods designed specifically for categorical data to impute missing ordinal data including binomial and multinomial logistic regression, multivariate probit model, ordinal (proportional odds) regression, and discriminant analysis. The preliminary result suggested that the multivariate normal strategy seemed to outperform the other strategies in reproducing the relationship between scale scores and the reliability coefficients for the scales regardless of the number of categories in the ordinal items.

Using Principal Component Analysis (PCA) to Obtain Auxiliary Variables tor Missing Data Estimation in Large Data Sets

Waylon J. Howard and Todd D. Little

To deal with missing data, methodologists have recommended an "inclusive" strategy where a large set of auxiliary variables are used to inform the missing data handling procedure. In large-scale studies, the set of possible auxiliary variables is frequently too large. We propose using principal components analysis to reduce the number of possible auxiliary variables to a manageable number. A series of Monte Carlo simulations were used to compare the performance of the inclusive strategy with 8 auxiliary variables to the PCA strategy using just 1 principal component derived from the 8 original variables. Results indicate that the PCAAUX approach removes both linear and nonlinear parameter bias as effectively as AUX and is at least as efficient. We conclude that using PCA to reduce the number of auxiliary variables is an effective and practical way to reap the benefits of the inclusive strategy in the presence of many possible auxiliary variables.

Missing Data Methods for Confounding Variables in Marginal Structural Models

Shu Xu and Vanessa Watorek, New York University

This presentation provides an evaluation of methods for missing confounding variables in a marginal structural model. Four methods are considered (1) impute multiple complete data sets, conduct the analysis, and summarize the effects using Rubin's rule; (2) use a product weighting variable in the estimation model where the product weighting variables is estimated as the product of the inverse likelihood of being assigned into the treatment group and the inverse likelihood of observing missing values; (3) pick one out of multiple imputed data then proceed; and (4) treat the missing values in confounders with mean substitution then proceed. We also consider how proportion of missing values and missing mechanism impact the performance of each method. Example data come from Gasper (2011). We study the causal effect of drug use on high school dropout using a marginal structural model.

Concurrent Session 1 Continued 10:30am–12:00pm

Combined Session 1.4: Analyzing Single Subject, Diary, and Intensive Longitudinal Data
Location: Laurel Hall 206

The Misspecification of the Covariance Structures in Multilevel Models for Single-Case Data: A Monte Carlo Simulation Study

Mariola Moeyaert, Maaike Ugille, John M. Ferron, S. Natasha Beretvas, and Wim Van den Noortgate (Presented by Patricia Rodriguez)

The three-level model for combining data from multiple cases is an appropriate method to estimate treatment effects over cases and over studies, at least if the underlying assumptions are not violated. In this article we explore by means of a simulation study the consequences of misspecification of the covariance matrices on inferences regarding the overall treatment effects' and covariances of the treatment effects, since it is not always obvious how to specify the covariance matrix at each level. It is valuable to be aware of the robustness of the three-level model to changes in the specification of the covariance matrices. We also include an empirical example of the comparison of correctly and misspecified covariance matrices. The results indicate that the treatment effect estimates are relatively robust for the ignoring of covariance. As expected, causes the misspecification in the random part of the multilevel model biased variance estimated.

Daily Diary Data: Effects of Cycles on Inferences

Yu Liu and Stephen G. West

Daily dairies and other intensive measurement methods are increasingly used to study the relationships between two time varying variables X and Y. Analytic work and three Monte Carlo studies investigated the impact of omitting weekly cycles in daily dairy data under the longitudinal multilevel model framework. In cases where cycles existed in both the time-varying predictor series (X) and the outcome series (Y) but were ignored, the effects of the within- and between-person components of X on Y were biased, as were their corresponding standard errors. The direction and magnitude of the bias depended on the phase difference between the cycles in the two series. In cases where cycles existed in only one series but were ignored, the standard errors of the regression coefficients for the within- and between-person components of X were biased, with the direction and magnitude of bias depending on which series contained cyclical components.

Modeling Cyclical Patterns in Daily College Drinking Data with Many Zeroes

David Huh, Debra L. Kaysen, and David C. Atkins

Daily drinking data often show highly skewed distributions that are bounded at zero as well as regular patterns across days of the week. Alcohol researchers have typically relied upon dummy variables for either weekend vs. weekday or for each day of the week. The present research evaluated the use of cyclical terms (i.e., sine and cosine regressors) in a zero-altered regression as a model for daily drinking data, in comparison to two dummy variable modeling strategies. Results showed that the cyclical model provided a more parsimonious approach than multiple dummy variables. Interestingly, a single dummy variable model evidenced superior overall fit, but obscured covariate effects that were detected in the other approaches. The combination of cyclical terms with zero-altered regression represents a feasible option for evaluating longitudinal drinking with high zero counts. However, drinking patterns are not perfectly sinusoidal, so care must be taken in evaluating the fit of models incorporating cyclical terms.

Combined Session 1.5: Profile Analysis and Non-normal SEM
Location: Laurel Hall 301

Modeling Configural Patterns in Latent Variable Profiles Associated with Endogenous Criteria

Mark Davison, Ernest Davenport, and Yu-Feng Chang

A model is proposed for identifying latent predictor score patterns associated with latent outcome variables. The model employs two new devices a path coefficient vector of contrast coefficients to describe a configural pattern in a structural model, and a new type of latent variable whose values quantify the match of the person's latent predictor variable profile pattern to a theoretical pattern associated with the factor. The model is illustrated using data on perceptions and evaluations of political candidates during a debate. Findings suggest a pattern of scores on the perceptual variables associated with perceived debate success for female observers but not for male observers.

Moderated Profile Analysis: Comparing Criterion-Related Predictor Variable Patterns across Populations

Mark Davison, and Ernest Davenport

This presentation describes a technique for comparing criterion-related predictor variable pattern vectors across populations via a marriage of criterion related profile analysis and moderated regression analysis that allows for a test that the profile pattern vectors associated with high scores on a criterion variable are equal in multiple populations. Results of an example of this procedure suggest a difference in the pattern of SAT scores associated with high GPAs for STEM versus non-STEM majors. It also appears that moderated profile analysis can be used to support the discriminant validity of measures and provide evidence supporting interpretation of configural profile patterns in test batteries. Limitations of the procedure are given as well as extensions to nonlinear, multilevel, and latent variable models.

Investigation of Type I Error Rates of Three Versions of Robust Chi-Square Difference Tests

Victoria Savalei, Jenny Chuang, and Carl Falk

A Monte Carlo simulation was conducted to investigate the Type I error rates of several versions of chi-square difference tests for nonnormal data. The studied statistics included the original uncorrected difference test, obtained by taking the difference of the ML chi-squares for the respective models; the original robust difference test, due to Satorra and Bentler (2001); the recent modification to this test which ensures that the statistic remains positive (Satorra & Bentler, 2010); and finally a hybrid procedure proposed by Asparouhov & Muthen (2010). A variety of types of constraints were studied. An interesting finding was that the uncorrected test D appeared to be robust to nonnormality when the constraint was setting factor correlations to zero. The robust tests performed well and similarly to each other in many conditions. The new strictly positive test exhibited slightly inflated rejection rates in conditions that involved constraining factor loadings, while the original difference test and the hybrid procedure exhibited rejection rates slightly below nominal in conditions that involved constraining factor correlations or factor loadings. While more research is needed on the new strictly positive test, we tentatively recommend the original difference test or the hybrid procedure.

Combined Session 1.6: Propensity Score Analysis

Location: Laurel Hall 302

The Impact of Measurement Error on Propensity Score Analysis: An Empirical Investigation of Fallible Covariates

Eun Sook Kim, Patricia Rodriguez de Gil, Jeffery D. Kromrey, Aarti P. Bellara, Rheta E. Lanehart, Tyler Hicks, and Reginald S. Lee Propensity score (PS) analysis is widely used for causal inference when randomized experiments are not viable. The first step of PS analysis involves covariate selection. This study investigated the impact of covariate selection on bias reduction in PS analysis when a pool of covariates has diverse levels of measurement error. Bias, standard errors, interval coverage, and interval width for treatment effect estimates and standardized mean differences for covariate balance before and after conditioning were examined. The relations of covariate reliability with other major factors of covariate selection including correlations among covariates, relations to outcome, and relations to treatment assignment in bias reduction are discussed. Also we aim to provide practical suggestions to applied researchers about covariate selection in the presence of measurement error.

Propensity Score Matching (with multilevel data) Using SPSS and R

Felix Thoemmes, and Wang Liao

Propensity score analysis is becoming an increasingly popular tool for social scientists. However, there are some misunderstandings about the more complex issues involved in propensity score matching. Those include the treatment of missing data, estimation and matching with complex samples (clustered data), presence of treatment-covariate interactions, and multi-valued treatments. In this methodological illustration, we present the basic logic and approach to propensity score analysis, but also cover some of the more challenging issues mentioned above. We also present an SPSS custom dialog written by the authors that allows applied researchers to easily conduct propensity score matching from within SPSS using the familiar point-and-click interface. Commands are then passed onto R and tables and graphs are returned to the regular SPSS output window. The custom dialog performs propensity score analyses with simple random samples, but also with complex sampling designs (clustered data).

Tuesday, May 21st, 2013

Lunch in Student Union Ballroom

12:00–1:30pm

Concurrent Session 2

1:30pm–3:00pm

Session 2.1: Open-source Modern Modeling Software: The R Package Lavaan

** This is a double session which runs through concurrent sessions 2 and 3

Location: Laurel Hall 102

Yves Rosseel

For many years, teachers, applied researchers, and statisticians working in the field of structural equation modeling (SEM) were stuck with commercial 'black-box' software. Recently however, several free and open-source packages for structural equation modeling and related methods have been developed, and are rapidly gaining momentum. In this presentation, I will focus on the R package lavaan.

The presentation will consist of two parts. In the first part, I will explain the aims behind the development of the package, give an overview of its most important features, and provide some examples to illustrate how lavaan works in practice. In the second part, I will discuss how lavaan attempts to capture the (computational) history of SEM, and how preparations are being made to shape the future of SEM. In particular, I will demonstrate how the computational infrastructure of lavaan can be used by modelers to build, test, and validate the modern modeling methods of tomorrow.

Combined Session 2.2: Extensions to Growth Models

Location: Laurel Hall 201

Multiple-Indicator Latent Growth Curve Models: An Analysis of the Second-Order Growth Model and Two Less Restrictive Alternatives

Jacob Bishop and Christian Geiser

Latent growth curve models (LGCMs) are widely employed methods for analyzing change in psychology and the social sciences. To date, most applications use first-order (single-indicator) LGCMs. These models have several limitations that can be overcome by employing multiple-indicator LGCMs. Currently, the most commonly used multiple indicator LGCM uses a second-order parameterization. In this article, we discuss two alternative, but less well-known multiple indicator LGCMs and compare them to the common second-order LGCM. Our presentation reveals that the second-order LGCM represents a fairly restrictive way of analyzing growth with multiple indicators. We discuss in which way the alternative models overcome some of these limitations.

The Autoregressive Latent Trajectories Model: An Alternative Approach to the Analysis of Panel Data

Chris Schatschneide

In this Methodological Illustration proposal, I plan on presenting the Autoregressive Latent Trajectories (ALT) model as an alternative to the either/or approach that is typically taken by researchers who either fit the simplex model or the latent growth curve model. I will demonstrate how this model – and its variations can be fit in MPLUS, including the fitting of the nonlinear constraints on the loadings, and how the parameters estimates can be interpreted. I will end with the presentation of a monte-carlo simulation study that will vary the forms of growth over time (fanspread growth, mastery learning" growth, and uniform growth) and the average correlations over measurement time –points and will allow for a direct comparison of the ALT modeling approaches with the latent growth curve alone and the simplex alone modeling approaches.

Recovery of Individual Trajectories in Heterogeneous Samples Using Longitudinal Latent Profile Analysis (LLPA)

Veronica T. Cole & Daniel Bauer

In the current proposal we introduce a technique, longitudinal latent profile analysis (LLPA), for the exploratory modeling of group-level and individual growth. A type of mixture model, the LLPA allows complex shapes of growth to be modeled without assuming any particular functional form of change in the data, or that the same functional form applies to all individuals under study. We apply this model to a longitudinal dataset tracking social skills in a sample of boys from kindergarten through fifth grade, finding that a four-class solution provides the strongest fit to the data. From this solution we reconstruct individual trajectories for all individuals in the sample, in order to make inferences about the nature of trajectories at the individual level. We discuss how the LLPA model may be extended to include covariates, and the model's potential utility in predicting change in a person-centered framework.

Symposium Session 2.3: Stepwise Approaches to Latent Variable Modeling
Location: Laurel Hall 202

Although simultaneously estimating the structural and measurement part of a latent variable model is known to obtain the most efficient parameter estimates, in many situation stepwise estimation procedure is preferred. The later is mostly used to reduce the parameter space of complex latent variable models.
In this symposium we introduce stepwise approaches to latent variable modeling using pseudo maximum likelihood procedures. The first presentation relates a continuous latent variable measured with error to distal outcomes in a multilevel setting using a two-step estimation approach. In the second presentation a three step approach to latent class analysis (three-step LCA) is introduced, focusing on parameter estimation and obtaining correct standard errors.
The third and fourth presentations show how to use the three- step LCA in two different settings, namely in latent transition analysis and in micro-macro analysis.

Group Means as Explanatory Variables in Multilevel Models

Jouni Kuha, Anders Skrondal and Stephen Fisher

Research questions for models for clustered data often concern the effects of cluster-level averages of individual-level variables. For example, data from a social survey might characterize neighborhoods in terms of average income, ethnic composition etc. of people within each neighborhood. Unless the true values of such averages are known from some other source, they are typically estimated by within-cluster sample estimates, using data on the subjects in the observed data. This incurs a measurement error bias when these estimates are used as explanatory variables in subsequent modeling, even if the individual observations are measured without error. The measurement error variance can, however, be estimated from within-cluster variation, using knowledge of the sampling design within each cluster, and we can then apply relatively standard measurement error methods to adjust for the error. This talk considers such estimation for generalized linear mixed models. The estimation uses a two-step approach which is an instance of pseudo maximum likelihood estimation and also mostly similar in spirit to regression calibration estimation in measurement error modeling. The methods are illustrated with models for political attitudes and behavior, using data from the 2010 British Election Study.

The Bias-adjusted Three-Step Approach to Latent Class Modeling with External Variables

Zsuzsa Bakk, Daniel Oberski, and Jeroen K. Vermunt

A popular way to connect latent class membership to external variables is to relate the external variables to the estimated scores on class membership; this approach is called three step latent class analyses (LCA). While the three step LCA is a popular approach, until recently it had the disadvantage that the parameters describing the association of latent class membership and auxiliary variables were underestimated. In the current paper we present how unbiased parameter estimates of this association can be obtained, by using the known classification error probabilities as fixed value parameters in the third step analysis. Next to correct parameter estimates we also show how correct standard error estimates can be obtained. The correction for parameter bias we propose is already implemented in Mplus (Muthen and Asparauhov, 2012), while the correction for standard errors is yet to be implemented in commercial software. We show the results of a simulation study where we test the performance of the parameter bias correction, and the SE bias correction methods.

Three Step Latent Transition Analysis

Bengt Muthen and Tihomir Asparouhov

In latent transition analysis (LTA) several latent class variables are measured at different time points and the relationship between these variables is estimated through a logistic regression. A three-step estimation can be conducted for the LTA model with Mplus where the latent class variables are estimated independently of each other and are formed purely based on the latent class indicators at the particular point in time. This estimation approach is very desirable in the LTA context because the one-step approach has the drawback where an observed measurement at one point in time affects the definition of the latent class variable at another point in time and where covariates create a more complex analysis. We illustrate the estimation with two different examples. The first example is a simple LTA model with 2 latent class variables. The second example is an LTA model with covariates and measurement invariance.

Three-step Estimation Method for Discrete Micro-Macro Multilevel Models

Margot Bennink, Marcel A. Croon, and Jeroen K. Vermunt

In 'reversed' multilevel analysis, a group-level outcome is explained by means of individual- and/or group-level predictors. Croon and van Veldhoven (2007) proposed a latent variable model to analyze these micro-macro relations when the outcome is continuous. Bennink, Croon and Vermunt (under review) extended this approach to discrete outcomes. In this latent variable approach, the scores of the individual-level units are treated as indicators of a latent variable defined at the group-level, and all individuals within the same group are exchangeable indicators for the latent group-level score (measurement part of the model). The outcome variable is not regressed on the aggregated scores of the individuals, but on the latent group-level variable (structural part of the model). The current presentation is restricted to the situation in which the latent variable is discrete.

Symposium Session 2.4: New Developments in the Analysis of Incomplete Data
Location: Laurel Hall 205

Utilizing Hyper Priors in Multiple Imputation for Multivariate Normal Data

Valerie Pare

Multiple imputation is one method commonly utilized to deal with missing data. Imputations typically require the assignment of prior distributions to unknown model parameters. However, since there is inherent uncertainty in what the hyper parameters of these prior distributions should be, not accounting for this uncertainty will inherently lead to over-confident inferences. This presentation will propose utilizing hyper priors to account for this uncertainty using the rules of two stage multiple imputation. In particular, we will examine the utility of this method when data is assumed to be multivariate normal. This is joint work with Dr. Ofer Harel.

Handling Data with Three Types of Missing Values: A Simulation Study

Jen Boyko

Incomplete data is a common obstacle to the analysis of data in a variety of fields. Values in a data set can be missing for several different reasons including failure to answer a survey question, dropout, planned missing values, intermittent missed measurements, latent variables, and equipment malfunction. In fact, many studies will have more than just one type of missing value. Appropriately handling missing values is critical in the inference for a parameter of interest. Many methods of handling missing values inappropriately fail to account for the uncertainty due to missing values. This failure to account for uncertainty can lead to biased estimates and over-confident inferences.

One area which is still unexplored is the situation where there are three different types of missing values in a study. This complication arises often in studies involved with cognitive functioning. These studies tend to have large amounts of missing values of several different types. I am proposing the development of a three stage multiple imputation approach which would be beneficial in analyzing these types of studies. Three stage multiple imputation would also extend the benefits of standard multiple imputation and two stage multiple imputation, namely the quantification of the variability attributable to each type of missing value and the flexibility for greater specificity regarding data analysis. Specifically for this talk, I will be discussing results of simulation studies which examine accuracy of this method under varying conditions of missingness. This is joint work with Dr. Ofer Harel.

Approaches to Multiple Imputation in Large Data Sets

Chantal Larose

When the number of records (N) in a dataset is less than number of parameters (P), traditional approaches to data analysis break down. The N < P data structure is becoming increasingly common, especially with the increasing availability of microarray data. We introduce a method to apply model-based soft clustering and multiple imputation to such incomplete datasets. Model-based clustering will help reduce dimensionality and create interpretable groups of parameters, while multiple imputation fills in missing values with plausible replacements so valuable records need not be discarded. We illustrate the problem with simulations and a data example. This is joint work with Dr. Ofer Harel and Dr. Dipak Dey.

Combined Session 2.5: Measurement Invariance and Differential Item Functioning
Location: Laurel Hall 206

A Simulation Study of a MIMIC-based Strategy for Detecting Items with DIF by a School Covariate

Shonte Stephenson, and Sophia Rabe-Hesketh

School-level DIF testing with a MIMIC modeling specification allows for a direct path from the school covariate to the observed item response. However, little is known in the literature about the association between school context and item functioning and there are no prior published examples of the analysis steps for testing differential item functioning (DIF) at the school-level. In this conference paper, we first present our DIF strategy, demonstrate how to apply each modeling step and then present the results from a simulation study. The simulation study investigates the performance of the iterative method for detecting items that exhibit "school-DIF" under various study conditions. Specifically, we evaluate its power and Type I error of the process for building a multilevel factor analytic model to detect DIF by a school covariate.

Testing for Measurement Invariance with Respect to an Ordinal Variable

Edgar Merkle, Jinyan Fan, and Achim Zeileis

Researchers are often interested in testing for measurement invariance with respect to an ordinal auxiliary variable such as age group, income class, or school grade. Traditional (likelihood ratio) tests neglect the fact that the auxiliary variable is ordinal, and they are also known to be overly sensitive at large sample sizes. In this paper, we propose test statistics that explicitly account for the ordinality of the auxiliary variable, resulting in higher power against ``monotonic'' violations of measurement invariance and lower power against ``non-monotonic'' ones. The statistics are derived from a family of tests based on stochastic processes that have recently received attention in the psychometric literature. The tests are illustrated via both simulation and application.

Model Invariance Testing Under Different Levels of Invariance

Holmes Finch and Brian French

Invariance testing is an important aspect of latent variable modeling in which researchers assess the extent to which model parameters are equivalent across two or more groups. Most empirical research in this area has focused on methods for measurement invariance testing (comparison of group factor loadings). In contrast, very little work has investigated the performance of methods for assessing invariance of model intercepts, unique variances, factor variances, and factor covariances. The goal of this study, therefore, was to use a Monte Carlo simulation study to ascertain how well several different testing criteria could accurately identify a lack of invariance, as well as the presence of invariance, on model parameters other than factor loadings. A variety of simulation conditions were manipulated, including level of noninvariance, sample size, degree of measurement noninvariance, number of factors, and indicators per factor. Implications of results are discussed.

Concurrent Session 2 Continued 1:30pm–3:00pm

Symposium Session 2.6: Comparing Analyses Predicting Simple Versus Residualized Change Scores

Location: Laurel Hall 301

Statistical recommendations vary about whether to analyze residualized or simple change, which sometimes yield contradictory conclusions (Lord, 1967; several recent studies). The first paper documents studies finding contradictory conclusions from the two types of change, robust results from one study, and illustrates a way to check robustness in a published study. The discrepancies reflect distinct counterfactuals from the null hypotheses. The second paper uses a simulation to show that the results of residualized analyses vary with the stability of the outcome, unlike simple-change analyses, viewed as an advantage of residualized change analyses. The third paper extends a published simulation to account for bidirectional influences between a child outcome and a parental cause of interest. Whereas the original simulation had favored analyzing simple change scores, the new simulation found biases in opposite directions for analyses of the two types of change.

Toward Understanding Discrepant Results from Predicting Residualized versus Simple Change Scores

Robert E. Larzelere, Ronald B. Cox, Jr., and Sada J. Knowles

Child development research is fundamentally about predicting change. Should longitudinal analyses predict simple or residualized change (e.g., latent growth models or cross-lagged panel analyses)? Analyses of these types of change two types of change can yield discrepant results, consistent with Lord's (1967)classic paradox about gender differences in weight gain. This paper will show how cross-lagged results can be re-analyzed to check the robustness of the result for both types of change, using an exemplary study of the apparent effect of spanking on externalizing problems (Gershoffet al., 2012). Latent growth modeling showed a reversed sign for the crucial coefficient. The pattern of results replicated when the waves were reversed, consistent with statistical artifacts, not unidirectional causation.

Modeling Change as a Residual Difference Score or as a Simple Gain Score: Further Clarifying the Paradox

Todd D. Little, Alexander M. Schoemann, and Matthew W. Gallagher

There is a long history of debate and discussion regarding whether researchers interested in modeling and predicting intraindividual changes should use residual difference scores or simple gain scores. Lord (1967) highlighted the distinction between the two approaches to model change as a paradox in that two researchers would get different answers if the same two wave data were analyzed as a difference/gain score or as a residual change score. In an attempt to resolve this paradox, we conducted a population simulation study to examine differences between these two models. The results of our simulation indicate that the difference score (simple gain score) and the residual change score both yield unbiased results, but that the two methods provide answers to very different questions.

Comparing Change Scores with Lagged Dependent Variables in Models of the Effects of Parents' Actions to Modify Children's Problem Behavior

David Johnson

There has been considerable debate in the Sociological literature over the advantages and disadvantages of controlling for prior levels of an outcome variable in a regression analysis by including the dependent variable from an earlier wave as an independent variable in the model. When examined in a fixed effects models framework, Allison (Allison 1990; Allison 1994)has convincingly demonstrated that a change score (fixed effects) approach to the analysis of longitudinal data has considerable advantages over alternative approaches such as lagging the dependent variable (also termed the residualized difference score). In a 2005 paper (Johnson 2005) I compared the two approaches for the analysis of two-wave data with a simulation study. The simulation found that the change score approach generally performed better than the lagged dependent variable models when examining the impact on the outcome of an event occurring between the two waves. In this current paper I extend the simulation model to allow for a causal effect of the child's behavior on the parent's response. The models include an effect of individual differences that impact both child behavior and parental response and allows for measurement error in the measure of child problem behavior.

Combined Session 2.7: Advances in Structural Equation Modeling
Location: Lauren Hall 302

An Efficient State Space Approach to Estimate Univariate and Multivariate Multilevel Regression Models

Fei Gu, Kristopher J. Preacher, and Wei Wu

Estimating multilevel regression models as structural equation models was thoroughly discussed by Bauer (2003) and Curran (2003). Based on the equivalence between structural equation models and state space models (e.g., Chow, Ho, Hamaker, & Dolan, 2010), the state space formulation for the multilevel regression models can be derived by a direct translation of the corresponding structural equation formulation. In this paper, instead of translating the existing structural equation formulation, we introduce a more efficient state space approach to estimating multilevel regression models. Though the state space approach has been well established for decades in the time series literature, it does not receive much attention from educational and psychological researchers. To the best of our knowledge, the state space approach to estimating multilevel regression models is barely known, and (almost) never implemented, by multilevel modelers in education and psychology. We first provide a brief outline of the state space formulation. Then, state space forms for univariate and multivariate multilevel regression models are illustrated, and the utility of the state space approach is demonstrated with both simulated and real examples. It is concluded that the results from the state space approach are essentially identical to those from specialized multilevel regression modeling and structural equation modeling software. More importantly, the state space approach is a much efficient treatment for multilevel regression models.

A Place for Nonlinear Structural Equation Modeling; What Insight Does It Give Us?

Jonathan Brewster, and Kathleen Buse

Linear regression techniques have served researchers well, however some management phenomena under investigation may not behave linearly. Structural equation modeling techniques that employ linear methods may be limited in adequately describing the relationship and variable characteristics under study. Richer insight and more accurate modeling of life may be available through non-linear techniques. We evaluated two data sets from prior research projects to illustrate the potential usefulness of a nonlinear SEM modeling technique.

Combined Session 2.8: Extensions to Mediational Analyses
Location: 305

Flexible Mediation Analysis in the Presence of Non-linear Relations

Beatrijs Moerkerke, Tom Loeys, Olivia De Smet, Ann Buysse, Johan Steen, and Stijn Vansteelandt

In social sciences, mediation analysis has been formulated in linear settings using the Baron & Kenny (1986) approach within the literature. Using natural effect models, we propose a more flexible approach which builds on the same counterfactual framework, but enables interpretable and parsimonious modeling of direct and mediated effects and facilitates tests of hypotheses that are otherwise difficult or impossible to test. We illustrate this in a study of individuals who ended a romantic relationship and explore whether the effect of attachment anxiety during the relationship on unwanted pursuit behavior after the break-up is mediated by negative affect during the break-up.

Modeling Indirect Effects with Multimethod Data

Ginger Lockhart, Christian Geiser, Hui Qiao, Jacob Bishop, Martin Schultze, and Herbert Scheithauer

The use of multiple raters or methods to assess constructs of interest is widely promoted in the social and behavioral sciences. This multimethod (MM) approach to improving measurement is attractive to many researchers because self-report alone is often not practical or feasible for vulnerable populations such as the very young or those affected by psychopathology or substance use. Additionally, because social/behavioral research questions often reflect an underlying mediating process, in which proximal variables are intermediate in a causal chain between the predictor and the outcome, several approaches have been devised to integrate MM techniques with statistical mediation analysis. One important challenge, however, is that the most commonly applied approaches to MM mediation analysis vary greatly, due in part to the dearth of guidelines for handling multimethod data for mediation models. The purpose of this talk is to provide an overview and demonstration of an innovative method, multimethod (MM) mediation analysis. MM mediation

analysis is a novel integration of two statistical methods: 1) the multitrait multimethod (MTMM) approach (Campbell & Fiske, 1959) and 2) statistical mediation analysis. The authors will discuss the application of this method for both structurally different (e.g. self, parent, and teacher report) and interchangeable (e.g. peer report) raters.

Testing Mediation the Way It Was Meant to Be: Changes Leading to Changes then to Other Changes. Dynamic Mediation Implemented with Latent Change Scores

Emil Coman, Eugen Iordache, and Maria Coman

We propose an approach for testing (true) dynamic mediation relationships between variables. We describe the current approaches to mediation testing, then demonstrate the need to reconsider changes and dynamics in investigating chained change processes. We illustrate a practical solution for testing various dynamic mediation processes using the vehicle of Latent Change Scores. We begin by introducing the LCS setup in its simplest model, the two-wave univariate LCS model, to which we add a second variable, then the third one. We describe several such models, like attitude about physical activity ->motivation to engage in physical activity -> actual physical activity behavior, and delve into the meanings of the findings of significant links between changes in attitudes leading to later changes in motivation then to changes in self-efficacy, coupled with interpreting the average changes themselves (average latent change in attitudes 1->2, average latent change in motivation 2->3, and changes in self-efficacy 3->4).

Concurrent Session 3

3:15pm–4:45pm

Session 2.1 CONTINUED: Open-source Modern Modeling Software: The R Package Lavaan
Location: Laurel Hall 102
Yves Rosseel

Combined Session 3.2: Longitudinal Mixture Models
Location: Lauren Hall 201

Using Mixture Latent Markov Models for Analyzing Change with Longitudinal Data
Jay Magidson, Statistical Innovations Inc.
Mixture latent Markov (MLM) models are latent class models containing both time-constant and time-varying discrete latent variables. They can be extremely useful in analyzing data arising from longitudinal surveys, clinical trials and related designs. Such models often fit data better than latent growth models since the autocorrelation structure in data often satisfies the Markov assumption. In this presentation we introduce and illustrate the utility of MLM models in several real world examples using the new GUI implemented in Latent GOLD® 5.0 which can easily accommodate even hundreds of time points.
The transition probability sub-model of the general MLM characterizes the latent Markov change process for each latent class. Various model modification restrictions can be applied to the parameters in this sub-model to test common MLM variations such as time homogeneous change, and mover-stayer structures where one latent class (stayer class) shows no change. In addition, we show how various 'longitudinal bivariate residuals', new in version 5.0 of Latent GOLD, can assist in model modification.

A Framework for Investigating the Performance of Latent Growth Mixture Models
Paul Dudgeon
Latent growth mixture models (LGMMs) have become extensively used in applied research settings. But the performance of these models under different conditions reflecting the kind of response process generating such data and the model specification choices used by researchers is not very well understood. This proposal contains details of general framework for undertaking a simulation study that has been extended from Skrondal's (2000) recommendations for using the benefits of experimental designs in Monte Carlo simulations. The proposed framework enables a large number of factors reflecting the response process that generates data containing latent heterogeneity to be systematically investigated to be integrated with the manipulation of model specification factors typically used by researchers in fitting LGMMs. The results of using this research design framework for the investigating both the incidence of local solutions in estimating LGMMs and the accuracy of latent class classification procedures is illustrated and detailed.

Symposium Session 3.3: Planned Missing Data Designs for Longitudinal Research
Location: Laurel Hall 202

Planned missing data designs are a way to exercise some control over the type of missingness that is bound to occur in longitudinal data collection. Though the idea of deliberately inserting missing data into a research design may sound perverse, planned missing designs allow researchers to reduce the testing burden on participants, leading to higher-quality data with less unplanned missingness and smaller fatigue and practice effects. The aim of this symposium is to investigate the implementation of planned missing data designs in longitudinal designs. This goal encompasses four presentations: (a) limitations for recovering parameters in planned missing data designs with small samples, (b) how to distribute missingness among survey or test items across time points in a planned missing data design, (c) how correlated growth curve models can be estimated with planned missing data designs and (d) how the two-method measurement design can be implemented most efficiently in longitudinal models.

Planned Missing Data Designs with Small Samples: How Small is Too Small
Alexander M. Schoemann, Fan Jia, E. Whitney G. Moore, Richard Kinai, Kelly Crowe, & Todd D. Little
Planned missing data designs enable researchers to ask participants fewer questions during the data collection process. An important question, however, is just how few participants are needed to properly implement planned missing data designs in studies. This presentation addresses that question by using simulated three-form planned missing data to assess the analytic model's convergence, and bias in point estimates and standard errors. Three models were examined: (a) a small, cross-sectional model (3 constructs), (b) a modest, two-timepoint model (3 constructs at each time point), and (c) a larger, mediation model (3 constructs over three time points). Multiple imputation (MI) and full-information maximum likelihood (FIML) were used to handle the missing data. Both estimation methods led to high convergence with modest and large models. Based upon this study's convergence, point estimate, and bias findings, we provide suggestions on determining minimum sample sizes needed for three-form planned missing design under various scenarios.

Assignment Methods in Three-Form Planned Missing Designs
Terrence D. Jorgensen, Alexander M. Schoemann, Brent McPherson, Mijke Rhemtulla, Wei Wu, & Todd D. Little
The three-form planned missing data design allows researchers to measure more items from more participants on more occasions using the same budget as a complete data design. The design also reduces burden, fatigue, and response reactivity, and because the assignment of missingness is completely random, no systematic bias in parameter estimation accompanies such a design. After randomly assigning participants to complete one of three forms on the first occasion, for all subsequent measurements researchers may then assign each participant either to the same form or a different form (randomly or systematically). This talk is an exploration of potential advantages and drawbacks of same-form, different-form, and random-form methods of assignment, including the effect on bias in parameter estimates and standard errors, 95% coverage rates, and model convergence rates.

Planned Missing Designs to Optimize the Efficiency of Latent Growth Parameter Estimates
Fan Jia, Mijke Rhemtulla, Wei Wu, Todd D. Little
We examine the performance of planned missing designs for correlated latent growth curve models. Sample data of size $N = 500$ were simulated according to a latent growth curve model that included two constructs measured at 5 timepoints, with correlated intercepts and slopes. When multi-form missingness is imposed, relative efficiency of factor loadings and residual variances can be low, but relative efficiency of structural parameters (e.g., means, variances, and covariances among latent growth constructs) tends to be very high. Wave missingness, in contrast, has a bigger effect on the efficiency of structural parameters. We discuss these results in terms of the cost effectiveness of each of these designs, and consider the additional impact of unplanned missingness (in the form of attrition).

Two-method Planned Missing Designs for Longitudinal Research
Mauricio Garnier-Villarreal, Mijke Rhemtulla, & Todd D. Little
Two-method measurement combines both measures to allow the benefits of the gold standard (i.e., validity) and the benefits of a large sample (i.e., power; Graham, Taylor, Olchowski, & Cumsille, 2006) using a structural equation model that separates construct variance from method variance (the source of systematic bias). The present study examines several extensions of the 2-method measurement design to longitudinal research. We examined the performance of models where the gold standard is measured at 1, 2, or all occasions, where the amount of bias is or is not constrained to have the same value over time, and where factorial invariance across time is or is not assumed. Our findings suggest that the two-method measurement design can be effectively implemented in multiwave studies by administering the gold-standard measure to a very small subsample at two or more waves of data collection.

Combined Session 3.4: Multidisciplinary Perspectives on Fit
Location: Laurel Hall 205

Comparing the Evidence for Categorical Versus Dimensional Representations of Psychiatric Disorders in the Presence of Noisy Observations: A Monte Carlo study of the Bayesian Information Criterion and Akaike Information Criterion in Latent Variable Models
Michael Hallquist, Thomas Olino, and Paul Pilkonis
Model selection criteria may help to resolve whether a set of psychometric indicators, such as diagnostic criteria, is best represented using a categorical or dimensional latent structure. Prior Monte Carlo studies, however, have focused mostly on multivariate normal data. In applied research, psychometric data rarely conform to this structure, and the model that generated the data is seldom among those tested. We conducted a simulation study to explore how observations inconsistent with the data-generating model influenced the relative evidence for a categorical versus dimensional latent structure. For data generated with a categorical structure, our results revealed that the Akaike Information Criterion was relatively robust, often selecting the correct model even when the data included many noise observations. The BIC, however, often preferred the dimensional model when small amounts of noise (e.g., 5%) were included. We provide recommendations for deciding on the latent structure of empirical data using model selection criteria.

The Effects of Local Item Dependence on the Sampling Distribution of the Q3 Statistic
William P. Skorupski, and Sukkeun Im
The purpose of this study is to demonstrate the behavior of the Q3 (Yen, 1984) statistic in the presence of differing magnitudes and types of local item dependence (LID). Ultimately, the goal of the study is to derive an empirical sampling distribution of Q3 for use in hypothesis testing about underspecified dimensionality and associated effect sizes. As the degree of extra-dimensional covariance among items increases, it is expected that the magnitude and distribution of Q3 will be affected. LID is simulated via the Testlet Response Theory (TRT) model (Wang, et al, 2002), so the question of interest is whether the distribution of Q3 may be mathematically described as a function of the testlet effect parameter and its variability. Results indicate that a mixture of known probability distributions may be used to model the sampling distribution, making it suitable for future hypothesis testing and effect size determination.

Bootstrapping SEM Goodness of Fit and Confidence Intervals
Craig M. Krebsbach, and Lisa L. Harlow
Structural equation modeling (SEM) fit and resulting indices have a varied set of suggested uses. The presented research address the use of bootstrapping within the SEM framework particularly with small sample designs. A 3 (size) x 3 (extension) x 3 (loading) x 5 (samples) confirmatory factor analysis model simulation was conducted with means and confidence intervals computed as well as a Bollen-Stine transformed p-value for each sample. Samples were analyzed for consistency with the initial bootstrap of equal size as well as for the extensions. The IFI and SRMR preformed the most consistent and the Bollen-Stine transform was supported as a necessity for SEM bootstrapping.

Concurrent Session 3 Continued 3:15pm–4:45pm

Combined Session 3.5 Multilevel Survival Analysis
Location: Laurel Hall 206

Modeling Microsocial Heterogeneity Using Multilevel Survival Analysis
Mike Stoolmiller
About 25 years ago, survival or hazard regression analyses were introduced to psychology (W. Gardner & W. A. Griffin, 1989; W. A. Griffin & W. Gardner, 1989) as powerful methodological tools for studying real time social interaction processes among dyads (e.g., parent-child). Almost no additional published applications have appeared, although such data are commonly collected and the applicable questions are central to many important theoretical perspectives. To demonstrate multilevel versions of these methods, the authors use an example from emotion regulation theory in which the hazard rate of angry emotions in the face of parental negative behavior (scolding) is hypothesized to be positively associated with the level of future child antisocial behavior. The authors discuss the limitations of traditional approaches to the analysis of social interaction and demonstrate improvements in the ability to model individual differences in microsocial behavior now available in Mplus.

Continuous Time Analysis of Panel Data: An Illustration of the Exact Discrete Model
Aaron J. Boulton, and Pascal R. Deboeck
Discrete time SEM models of change (e.g., a cross-lagged panel design) are popular in many fields. Unfortunately, these models produce parameter estimates that depend on the lag between observations. The purpose of this talk is to demonstrate the Exact Discrete Model (EDM), a continuous-time model of change that can be estimated within the SEM framework and produces lag-independent estimates. After the model is introduced, we will present two applied examples of the EDM. R program code will accompany the examples. We will conclude with a brief discussion of related methods and modeling extensions.

Combined Session 3.6: Extensions to Factor Analysis
Location: Laurel Hall 301

Multilevel Factor Analysis by Model Segregation: The Surprising Necessity for Robust Statistics with Normal Data
Jonathan Schweig
Measures of classroom environments have become central to policy efforts that assess school and teacher quality. This has sparked a wide interest in using multilevel factor analysis to test measurement hypotheses about classroom-level variables. One approach partitions the total covariance matrix and tests models separately on the between and within levels. This paper shows that when using this approach, inferences based on the normality assumption may not yield valid inferences about the classroom-level measurement structure even when the data is normally distributed. This paper then presents an empirical example and a simulation study to demonstrate how item intraclass correlations influence the quality of inferences based on the normality assumption. The results have implications for the study of classroom environments.

The Influence of Parceling on the Implied Factor Structure of Multidimensional Data
Brooke Magnus, and Yang Liu
The effects of two parceling methods on the implied factor structure of multidimensional data are examined. Item covariance matrices are derived from several bifactor models and then parceled; a one-factor CFA is then fit to the parcel covariance matrices to determine parcel loadings on this factor. To study the extent to which the primary factor from the generating bifactor model is recovered after parceling, factor scores from the one-factor CFA are computed and compared to the true values from the bifactor generating model. Results indicate that when the influence of the secondary factors is weak, the difference between parceling methods is small. This difference increases as the secondary influence becomes stronger. Additionally, factor score-true value correlations suggest that when the secondary influence is strong and the primary loading is weak, factor scores are more highly correlated with the secondary factors than the primary factor. Implications for applied research are discussed.

To Thin or Not To Thin? The Impact of Thinning Posterior Markov Chains on Parameter Estimation in Latent Trait Models

Jared K. Harpole, and William P. Skorupski

The practice of thinning MCMC chains for item response theory (IRT) models has mixed reviews in the literature. Recently, Link and Eaton (2012) found that thinning MCMC chains from a t-distribution produced more biased estimates versus not thinning. The purpose of the present study is to extend the work of Link and Eaton (2012) to include the impact of thinning versus not thinning on parameter estimation when fitting a 2PL IRT model. A simulation study was conducted using data based on a study from Rodebaugh et al. (2004). Fifty data sets were generated and analyzed using JAGS and R. The parameter estimates for each replication were evaluated for bias, mean square error, and credible interval width. Results indicated that thinning the MCMC chains tended to produce poorer estimates on average versus not thinning. Practitioners are advised not to utilize thinning with IRT models solely for improving parameter estimates.

Session 3.7: Introduction to Quantile Regression for Social Science Researchers
Location: Laurel Hall 302

Introduction to Quantile Regression for Social Science Researchers

Jessica Logan and Yaacov Petscher

Quantile regression (Koenker & Basset, 1978) expands on the basic framework of regression techniques. Where regression typically estimates the strength of the relation of X with Y, quantile regression estimates that same relation at multiple points in the distribution of Y. Quantile regression is a type of conditional median modeling that is common in economics, but is rarely used in educational science. However, the technique is a useful tool for addressing many analytic problems faced by education scientist. For example, it was developed to examine heteroschedasticity, and so makes no presumptions about the distribution of error terms, making it excellent for examining constructs with floor and ceiling effects. In the proposed methodological illustration, we will conceptually introduce quantile regression, describe research questions suitable for quantile regression, provide opportunities to apply the statistical technique to provided data for hands-on practice, and teach best practices in model building and publication

Tuesday, May 21st, 2013

Poster Session and Reception

Student Union Ballroom 5:00pm – 7:00pm

Poster Information

1. **Investigating School Climate Using Multilevel Latent Class Analysis**
   Diana Mindrila, Christine DiStefano, and Diane Monrad

When classifying, a crucial assumption is that the data are independent; however, much of the research conducted in education follows a hierarchical structure, where data are nested and have dependent relationships. Ignoring the nested structure when conducting analyses may lead to incorrect findings. The purpose of this study is to investigate a new method, Multilevel Latent Class Analysis (MLCA). MLCA allows for a blending of latent class analysis (LCA) to identify subgroups of similar cases coupled with multilevel strategies that recognize dependencies in the data and allow the contextual features to be incorporated. This study provides an illustration of MLCA with a school climate measure that is administered to teachers state-wide. Analyses will contrast information obtained with MLCA and Latent Profile Analysis (LPA).

2. **What's it like to be "Neurotic"? An Open Vocabulary Approach to Personality**
   Margaret L. Kern, Johannes C. Eichstaedt, H. Andrew Schwartz, Lukasz Dziurzynski, Lyle H. Ungar, and Martin E. P. Seligman

Language in social media reveals a lot about people's personality and mood as they discuss the activities and relationships that constitute their everyday lives. Schwartz et al. (2013) introduced an open language approach that identifies and visually summarizes the dominant naturally occurring words and phrases that most distinguish people as a function of characteristics such as age or gender. We applied this technique to examine personality differences expressed online. Method: Using millions of posts from 69,792 Facebook users, we examined the correlation of Big Five personality traits with online word usage. Analyses involved feature extraction, correlational analysis, and visualization. Results: The distinguishing words and phrases were face valid and provide insight into processes that underlie the Big Five traits. Discussion and significance: Open-ended data driven exploration of large datasets combined with established psychological theory and measures offers new tools to further understand the human psyche.

3. **The Mediating Effect of Sponsor Awareness and Sponsor Image on the Involvement-Purchase Intention Relationship**
   Ahlam Fakhar, Catherine Bachleda, and Zineb Elouazzani

In one of the few studies to use a multiple mediation model, this research explored the mediational role of sponsor awareness and sponsor image on the involvement-purchase intention relationship. Using the casual steps approach, the product of coefficient approach and bootstrapping the indirect and conditional indirect effect of the two mediators was estimated and the size effect calculated. Results revealed that whilst image was an effective mediator, awareness acted as a suppressor on the involvement and purchase intention relationship. Implications based on the findings highlight the importance of considering sponsor image when seeking to invest in sponsorships and using involvement as a segmentation base.

4. The effects of Marketing Capabilities on Performance Outcomes of International New Ventures: The Mediating Role of Market Information Resources

Silvia Lozano Martin, and Rajshekar Javalgi

In an emerging market context, this study examines the impact of market information resources on marketing capabilities and performance outcomes at the export venture level of international new ventures (INVs). To date, the literature on high-tech INV firms has failed to consider the importance of market information resources within the marketing capabilities interplay in approaching new markets overseas. With survey data from 260 INV firms in Mexico structural equation modeling was used to test measurement veracity and proposed relationships between the constructs constrained in the measurement model. The results indicate that market information resources play a central role in the process of attaining superior export venture performance. These findings indicate that the rapid environment that INVs find themselves requires an entrepreneurial orientation mediated with market information resources to impact on marketing capabilities from which to achieve superior export venture performance. This investigation contributes to the study of INVs by demonstrating that market information resources help INVs to focus on deploying marketing capabilities based on new product development, sales and distribution service to enhance export venture performance.

5. Changes in Risk Behavior Achieved by Activating Dynamic Coupling Processes: dynamic growth modeling of a health prevention intervention

Emil Coman, Carolyn Lin, Suzanne Suggs, Eugen Iordache, Maria Coman, and Russell Barbour

We introduce the modeling of bivariate dynamical changes as a method of assessing the intervention effects of a substance use prevention program. Applying the techniques of dynamical modeling with latent change scores (LCS) to a four-wave data set, we show how the intervention has spurred a strong coupling effect between Internal locus of Control (ILC) and Intent to Use Substances (IUS) in the intervention group, as well as new mutual dynamic links – i.e., changes in ILC causing subsequent IUS changes, and vice versa.

6. Performance of the MIMIC-Model Methods for DIF Detection

Pei-Hsuan Chiu, and H. Jane Rogers

Recently, a number of MIMIC-model methods have been proposed for DIF testing. These methods differ in the statistical criteria or the procedures that are used during the DIF detection process. In addition, a key difference of these methods is how the (DIF-free) items are selected to serve as anchors. The purpose of this study is to compare four different MIMIC-model methods for DIF testing. A series of Monte Carlo simulation studies were conducted using MPLUS with manipulation of test length, sample size and the percentage of DIF-items including in the test (including DIF-free condition). The Type I error rate of misclassifying DIF-free items and the power of identifying DIF items for each condition were assessed as the outcome variables. Finally, a couple of real datasets from large-scale student assessments were analyzed to show applications of the MIMIC-model methods that were compared in the simulation.

7. The Relationship between Middle Years Programme Student Performance and Diploma Programme Student Performance

Melissa Gordon, and Liz Bergeron

The International Baccalaureate (IB) offers three educational programmes for a worldwide community of schools, aiming to create a better, more peaceful world. The IB offers Middle Years Programme (MYP) for students aged 11 to 16 and Diploma Programme (DP) for students aged 16-19. Previous research suggests that participation in an IB preparatory program better prepares students for the rigour and challenges of the IB DP (Bland & Woodworth, 2011). The goal of the present study is to investigate if IB Diploma Programme (DP) students benefit from previously completing the IB Middle Year Programme (MYP). This is accomplished by conducting regression analyses, multilevel modeling (HLM), and an independent t test on the existing data. Findings suggest that students who perform better during MYP moderation tend to perform better on their DP subject exams.

8. Housing Profiles and Changes in Low-Income Children's Academic, Emotional, and Behavioral Functioning

Rebekah Levine Coley, Melissa Kull, Tama Leventhal, and Alicia Doyle Lynch

The present study sought to comprehensively assess the "housing bundle" of a representative sample of low-income families using a latent class analysis (LCA). Data were drawn from the longitudinal Three City Study of families and children living in low-income neighborhoods in Boston, Chicago and San Antonio. Profiles identified in the LCA were subsequently used as predictors in individual fixed effects models predicting child functioning. Results revealed that shifts between housing profiles was associated with changes in children's academic, emotional, and behavioral functioning. This study makes a unique contribution to the literature linking families' housing contexts and children's functioning.

9. Model Specification for Interaction and Quadratic Effects Between Formative Latent Variables

Shu-Ping Chen, and Chung-Ping Cheng

Estimating the interaction and quadratic effects of latent variables is an important issue in the social and behavioral sciences. A variety of approaches have recently been developed for the estimation of nonlinear structural equation modeling. To our knowledge most approaches have predominantly been used to estimate latent nonlinear effects between reflective exogenous variables. While Chen and Cheng (accepted) generalize the constrained approach to a matrix form that encompasses the latent nonlinear effects between exogenous and/or endogenous variables, their framework only accommodates reflective variables. In the current research, we extend Chen and Cheng's original framework to encompass the nonlinear effects between formative latent variables which can be exogenous and/or endogenous variables. Constraints are specified in matrix form and the matrices involved in model specification are partitioned to fit into our nonlinear model framework. The usage and validity of the procedure is demonstrated with a simulated dataset example using the OpenMx package.

10. Longitudinal Influence of Marital Satisfaction on Harmonious Family Interaction: Using SEM to Estimate the APIM with Distinguishable Dyads

Yeonsoo Yoo, Rachel B. Tambling, and JoAnn L. Robinson

This poster presents a longitudinal investigation of interparental relationship, parental sensitivity, and family interaction quality. These findings suggest that maternal and paternal perceptions on marital satisfaction appear to reciprocally affect each other before but not after child age 3 years old. Only mothers' marital satisfaction influenced their own sensitive behaviors toward their children and further, harmonious family interactions at dinner. This study has important implications for intervention with couples and families at critical time points.

11. Mediational Analysis of Intervention Effects on Dietary Fat Reduction Using Parallel Process Latent Growth Modeling

Leslie Brick, Si Yang, and Lisa L. Harlow

Data from three randomized trials of Transtheoretical Model (TTM) tailored multiple behavior interventions (N=9461) were pooled to examine the meditational relationship of three TTM measures (Pros, Cons, Temptations for Dietary Fat) with treatment group and stage of change. The focus of the current study was on energy balance behavior, represented as having a diet low in fat (i.e. less than 30% daily consumption) with measures at three time points: baseline, 12-months, and 24 months. Parallel processes latent growth curve modeling was used to assess mediation of measures. In this study, mediation was supported when the treatment group significantly affected the trajectory of the TTM measure, which then affected the trajectory of stage membership. The Cons scale was the only measure with a significant mediation, suggesting that intervention treatment groups resulted in an increase in the growth rate of Cons, which decreased over time, and an increase in the growth rate of stage, which increased over time.

Poster Session and Reception Continued

Student Union Ballroom 5:00pm – 7:00pm

12. Associations between Perceived Racial Discrimination, Alcohol Use, and Drinking Problems: A Structural Equation Model with Moderated Mediation

Ross E. O'Hara, Marcella H. Boynton, Denise Scott, Stephen Armeli, Howard Tennen, and Jonathan Covault

The effects of racial discrimination on alcohol use and drinking problems were examined using structural equation modeling with moderated mediation. This study extended prior knowledge by examining whether discrimination effects were mediated by discrimination-specific or generalized anger; how these mediation patterns related to drinking problems; and whether these processes differed by gender. Survey data came from 741 African-American students at a historically Black university (53% female). Discrimination predicted alcohol use only among men, and mediated fully by generalized anger. Similarly, discrimination predicted men's drinking problems through generalized anger and depressive symptoms. For women, only a risky family environment predicted alcohol outcomes. These results replicate and extend prior research by showing that African-American men exposed to discrimination may become more prone to anger, in general, prompting increased alcohol use. Additionally, indirect effects of discrimination on drinking problems via changes in affect may help explain male African-Americans' disproportionate risk for these outcomes.

13. Using Structural Equation Modeling to Assess The Effect Affective Variables and Socio-Economic Status on Mathematics Achievement at Lower Secondary Level

Khemduth Singh Angateeah , Dr Kaviraj Sharma Sukon, Dr Preethee Nunkoo-Gonpot

The purpose of this study was to examine the effect of socio-economic status and three school related construct – Attitude, Motivation and Perceived usefulness of Mathematics – on 8th grade (13 years old) achievement in mathematics in Mauritius. Though cognitive abilities of students of students and their family background are important predictors of achievement, there is an increasing amount of research evidence on the salient role of affective variables on mathematics achievement. A sample of 491 grade 8 students from 14 secondary schools was involved in 2011. Structural Equation models (LISREL 9.0) was used to measure the influence of affective variables and Socio-economic status on mathematics achievement. Socio-economic status and attitude had greatest influence on mathematics achievement. Perceived usefulness of mathematics has positive influence on motivation and attitude but negative direct influence on mathematics performance.

14. Using the Rasch Model to Assess the Interest for Shakespeare and Illustrate Rating Scale Diagnostics

Gloria Yeomans-Maldonado

An illustration of how the Rasch Model was used to investigate the psychometric properties of a 35-item survey developed to gather student attitudes regarding Shakespeare will be presented. Similar to many surveys, it employed an ordinal level of measurement. However, one of the sections utilized a "Don't know" [DK] response choice nested between the "Yes" and "No" responses choices. The test developer's intention was that this option might act as a middle point. Standard practice is to set such values to missing because the scale point does not fall along the continuum of the other anchor points. The third section of the survey instrument employed a standard three-point ordinal scale. A common practice is to treat such scales as interval and to create composite scores via averaging. However, this practice is only valid if equidistance holds between adjacent scale points. Rasch measurement diagnostics (logit scores, threshold estimates, probability curves, and category fit) were used to determine both whether DK could be treated as a middle point as well as whether the equidistance assumption was met.

15. Examination of Growth Models within Ohio Vendor Assessments

Lauren M. Porter, and Susan Mauck

Beginning in 2008, Ohio has used value-added measures as an accountability component in district and school report cards. Under the new Ohio Teacher Evaluation System (OTES) and the Ohio Principal Evaluation System (OPES) value-added measures will become an integral component of teacher and principal evaluations. Given that current statewide assessments do not cover all subjects or grades, the use of non-state assessments has been approved for evaluation purposes in select cases. This study seeks to examine the methodology behind and definitions of growth as applied to Ohio Department of Education designated Vendor Assessments. Within this study we seek to review and compare differences in the definition of growth, methodologies used to determine growth and reliability and validity measures for the 16 assessments identified in January 2013 as Vendor Assessments. Results will be used to inform policymakers, including the Ohio Department of Education, on potential changes for Vendor Assessment selection criteria.

16. Learning Using Hidden Information Applied to Inference of Motor Neuron Function from Differential Gene Expression

Therese M. Smith, and Greg Johnson

Recent work by Dr. Pamela Shaw and colleagues, obtaining differential gene expression in spared and vulnerable motor neurons in Amyotrophic Lateral Sclerosis(Unravelling the enigma of selective vulnerability in neurodegeneration: motor neurons resistant to degeneration in ALS show distinct gene expression characteristics and decreased susceptibility to excitotoxicity, Acta Neuropathologica 2013) has made use of DAVID (http://david.abcc.ncifcrf.gov/) to infer the functions these differently expressed genes might support. We undertook to review the methods of inferring functions from gene expression in the multiple tools used by DAVID and at Gene Ontology, to see whether the work of Dr. Vladimir Vapnik, described as learning using hidden information, might offer additional benefits.

17. Under Detection Problem of Cytokines: an AFT Model Application

Yu-Bo Wang, James Grady, Rong Wu, and Zhao (Helen) Wu

The roll of proinflammatory cytokines as a biomarker for human diseases has become increasingly investigated in clinical studies. Among them, IL-6, IL-10, IL-1ra, and TNF are of most interest. However, cytokine data present two typical analysis challenges including 1) left censoring due to in- struments that cannot detect very low levels of cytokines, 2) right-skewed distribution. To identify better methods to appropriately analyze this type of data, we utilize cytokine data from an NIH sponsored longitudinal study between 2006-2011, which examines cytokines as predictors of stress and drug use in a cohort of women. We will compare several approaches that have been suggested in the literature, including log transformation, mixture models and the Tobit model to a parametric accelerated failure time model borrowed from the survival analysis field. Simulation results will be presented to compare the performances of those methods.

18. Using HLM to Understand Alliances in High Conflict Divorce Treatment

Scott C Huff, Shayne R. Anderson, and Rachel B. Tambling

This poster will present an example of using Hierarchical Linear Modeling with longitudinal dyadic data. Our study compared the nature of the therapeutic alliance between high conflict divorce couples and traditional couples. Using data from 140 clinical couples, we modeled growth curves on several measures of the therapeutic alliance comparing divorced couples to traditional couples. Analysis showed that on each measure the two groups had different trajectories. Implications will be discussed, especially the value of HLM for marriage and family therapy research. Our study is an example of how HLM answers many of the common critiques of typical marriage and family therapy research.

19. Methods for Intuitive Presentations of Hierarchical Categorical Models: An Example Predicting Futile Treatment in Critical Care

Joshua F. Wiley, Thanh N. Huynh, MD, and Neil S. Wenger, MD

Categorical outcomes and hierarchical data structures are common in medical research. Generalized linear mixed models can be difficult to interpret; however, predicted probabilities and marginal effects on the probability metric can be used to provide meaningful, interpretable results for researchers and clinicians. In this poster, we demonstrate an application of the method with a study of predictors of receiving futile treatment in critical care units using a cross classified random effects Bayesian ordered probit model and the newly developed R package, postMCMCglmm to assess the accuracy of the model as number correctly classified, understand random effects in terms of standard deviations of probabilities, compute average marginal predicted probabilities to assess overall effects of predictors, and create graphs of predicted probabilities with 95% highest posterior density intervals.

20. Day-to-Day Stress-Pain Relationships among Fibromyalgia Patients: Does Social Support Moderation Change the Daily Relationship over Time?

Deborah A. Forrest

Participants in the study were a sub-sample of the national sample of women diagnosed with FM referred by Rheumatology practices in the United States. Participants (N = 210) completed a telephone interview and then completed mailed daily diaries in a Measurement Burst design with seven daily diary measurements in three bursts across three years on their daily pain, stressors and negative emotions. Social support was assessed from a baseline interview as it was found to be unchanging over the course of the study. The results indicated that individuals with FM experienced increased pain on days they experienced higher than average levels of daily stressors, and this relationship increased in magnitude over time. In addition, there was support for the notion that perceived availability of social support moderated the relationship over time. For those with low social support, the relationship between stress and pain intensified over time but in those with high social support, this relationship did not change over time. These findings have implications for FM patient education and cognitive behavioral therapies capitalizing on reduced stress, particularly for those with low social support.

21. Investigating the Directionality and Pattern of Mutual Changes of Health Outcomes: Adding Dynamic Perspectives to Static Longitudinal Analyses

Emil Coman, Marco Bardus, Suzanne Suggs, Eugen Iordache, Maria Coman, and Holly Blake

We walk the audience through stepwise analytical procedure meant to analyze complex mutual dynamic relationships between health outcomes. We introduce first simple difference scores for a set of 6 outcomes measured every six months over five time points and describe their relations, then explain the latent (true) change scores (LCS) specification. We detail alternative specifications of the LCS scores, and how they might impact the fit and the estimates of the models (e.g. latent change scores defined as a 1 indicator factor, by the prior X1 or the subsequent wave variable X2). We report on models linking the key variables from the Theory of Planned Behavior (TPB) based health intervention aimed at improving physical activity (PA), examine combinations of sets of outcomes and their mutual changes, and conclude with implications for methodologists and practitioners of such dynamic change modeling.

22. An Investigation of a Multidimensional Model of School Readiness

Jill Pentimonti, Kimberly Murphy, Laura Justice, Jessica Logan, and Joan Kaderavek

The present study aimed to examine the construct of school readiness and the extent to which it captures three dimensions of readiness skills (academic, social, and behavioral), as well as how these dimensions relate to children's end-of-kindergarten literacy skills. Moreover, this study addresses these issues among children with language impairment (LI), a population of children known to be at risk for reading difficulties. Participants were 136 preschool-aged children with LI drawn from 83 special education classrooms. Children were assessed on measures of academic, social, and behavioral skills. Confirmatory Factor Analyses indicated that school readiness for this sample of children is best characterized as two separate dimensions (academic, socioemotional). Of these components, academic readiness was found to be predictive of children's later performance on measures of literacy. The results of the study further our theoretical understanding of the dimensions of school readiness by empirically testing the nature of school readiness models.

23. Leveraging Both Between- and Within-Person Variability in Predicting Alcohol Use: Theory and Application of Centering Approaches in a Multilevel Modeling Context

Marcella H Boynton, Ross E. O'Hara, and Howard Tennen

Multilevel modeling (MLM) is increasingly becoming the analysis of choice when examining data with a nested structure. In MLM variable centering greatly facilitates model convergence and interpretation of the variances of the intercept and slope. Despite these benefits, many users of MLM are still hesitant to employ centering, in part because there are multiple centering strategies with distinct statistical and theoretical implications. The focus of this talk is to provide an overview of centering strategies within MLM as well as a discussion of how these strategies impact interpretation of MLM results. Results from two daily diary studies of alcohol use will be used to demonstrate the application and theoretical interpretation of different MLM centering strategies. The statistical arguments both in favor of and opposed to different centering techniques, in general, and within daily diary study analyses, specifically, will be considered.

24. Predicting Condom Use Using the Information-Motivation-Behavioral Skills (IMB) Model: A Multivariate Latent Growth Curve Analysis

Jennifer L. Walsh, Theresa E. Senn, Lori A. J. Scott-Sheldon, Peter A. Vanable, and Michael P. Carey

The Information-Motivation-Behavioral Skills (IMB) model often guides sexual risk reduction programs even though no studies have examined covariation in the theory's constructs in a dynamic fashion with longitudinal data. Using new developments in latent growth modeling, we explore how changes in information, motivation, and behavioral skills over 9 months relate to changes in condom use among STD clinic patients. Participants (N=1281, 50% female, 66% African American) completed measures of IMB constructs at three time points. We used parallel process latent growth modeling to examine associations among intercepts and slopes of IMB constructs. Initial levels of motivation, behavioral skills, and condom use were positively associated, with behavioral skills partially mediating associations between motivation and condom use. Changes over time in behavioral skills positively related to changes in condom use. Results support the key role of behavioral skills in sexual risk reduction, suggesting skills should be targeted in HIV prevention interventions.

25. The Application of Latent Transition Analysis to Large Scale Disaster Data: Modeling PTSD in a Population of Disaster Workers

Katarzyna Wyka, Jay Verkuilen, Judith Cukor, Brittany Mello, Megan Olden, and JoAnn Difede

Sophisticated statistical methodologies are needed in order to analyze large, population-based datasets, such as screening projects, following disasters. Currently, the most common methodology applied to disaster research uses marginal or population-averaged models. However, mixture models with latent variables have versatile applications that may provide additional insight into the psychiatric outcomes following disasters and capture population heterogeneity that is usually overlooked.

The goal of this paper was to assess the utility and feasibility of applying latent transition analysis (LTA) in large scale disaster research, specifically within a study of the longitudinal course of posttraumatic stress symptoms in WTC disaster workers. The model yielded practical findings in the field of disaster research. These findings have the potential to offer insight into the mechanisms involved in the maintenance and remission of posttraumatic stress symptoms. Challenges associated with the analysis of complex longitudinal data from large screening databases and future directions are discussed.

26. Empathic Accuracy in Daily Interactions between Romantic Partners: The Effect of Communal Behavior on the Perception of The Partner's Negative Affect

Gentiana Sadikaj, Debbie S. Moskowitz, and David C. Zuroff

The present study examined situational cues that influence the perception of the partner's negative affect. Using an event-contingent recording methodology, partners in 93 cohabiting couples recruited from the community reported on their feeling of unhappiness, interpersonal behavior, and their perceptions of the partner's unhappiness in interactions with each-other during a 20-day period. The truth and bias model of judgment (West & Kenny, 2011) was used as a conceptual and statistical framework for the data analysis. Findings suggest the presence of mean level bias (i.e., overestimation), tracking accuracy, and assumed similarity. They indicate that partner's communal but not agentic behavior affected perception accuracy. Results extend prior research by demonstrating that the partner's communal behavior serves as a cue that affects the perception of the partner's affect. This study demonstrates the use of novel methodological advancements in statistical modeling and measurement in examining questions pertaining to social perception.

Concurrent Session 4 Continued 8:45am–9:45am

Combined Session 4.1: Meta-analysis
Location: Laurel Hall 201

Individual Participant Data Meta-Analytic Modeling Techniques 1993 - 2012: A Methods Review
Samantha A. Russo, and Tania B. Huedo-Medina
Over the past 30 years, research syntheses exploring the links between medical and behavioral factors have focused on the combination of studies published summary statistics. Aggregated data meta-analysis (AD-MA) introduces potential ecological or aggregation bias. Therefore, individual-participant data meta-analysis (IPD-MA) has been recognized as the gold standard for combining evidence from existing studies. The methods involved in IPD-MA have improved since the first pooled data analyzed in 1994, but it is not known to what extent IPD-MA exhibits high quality among all of the steps and variety of statistical complexities that can still be applied. IPD-MA provide thoughtful insights into behavioral and medical relationships but to date none have applied sufficient advanced statistical model to capture the richness of IPD-MA. Our results offer some guidance as how to conduct IPD-MA and which new statistical models and techniques applied to IPD-MA are most needed to address medical and behavioral factors simultaneously.

Measurement Harmonization in Individual-Participant-Data Meta-Analytic Modeling
Tania B. Huedo-Medina, Francisco Galindo-Garre, and Maria Dolores Hidalgo
The main goal of this first project is to evaluate the methodological approaches to combine different measuers of the same construct, including advantages and disadvantages of the techniques under different circumstances to, finally, compare one of the most common ones, standardization with a moderated nonlinear factor analysis..
The reanalysis of studies' individual-level data has been recognized as the gold standard for combining evidence from existing studies. However, precise metric transformations among factors that presume to measure the same variable are needed.We have reviewed approaches to statistical harmonization which could be used when individual participant data from different sources are integrated using standardized or non self-reports or other strategies and we compared them. We found the moderate nonlinear factor analysis more generalizable harmonization technique when individual data needs to be integrated than just standardizing the metrics.

Combined Session 4.2: Statistical Disclosure Limitation
Location: Laurel Hall 202

Application of Mixture Models in Statistical Disclosure Limitation
Anna Oganian
In this paper, we explore an application of several types of mixture models in the context of Statistical Disclosure Limitation (SDL). SDL is an interdisciplinary area at the intersection of statistics, mathematics and computer science aiming to provide a safe release of statistical individual data to the public. Before releasing such information, statistical agencies have to apply SDL methods to the data. SDL methods often consist of masking the original data records (by either perturbing them or reducing their detail) in such a way to minimize the risk of disclosure of the confidential information, and at the same time, to provide legitimate data users with accurate information about the population of interest. In this paper, we propose a new scheme for disclosure limitation, which is based on mixture models.We show that our method is superior to the others because it has higher data utility and low disclosure risk.

Assessing the Privacy of Randomized Vector-valued Queries to a Database Using the Area Under the Receiver-operating Characteristic Curve

Gregory Matthews

As the amount of data generated continues to increase, consideration of individuals' privacy is a growing concern. As a result, there has been a vast quantity of research done on methods of statistical disclosure control (SDC). Some of these methods propose to release a randomized version of the data rather than the actual data. While methods of this type certainly offer some layer of protection since no actual data is released, there is still the potential for private information to be disclosed. Quantifying the level of privacy provided by these methods is often difficult. In the past, a method for assessing privacy using the receiver-operator characteristic (ROC) curve based on ideas related to differential privacy has been proposed. However, the method was only demonstrated for univariate randomized releases. Here, the ROC based privacy measure is extended to the release of randomized vectors.

Session 4.3: Bayesian Analysis of Dynamic Item Response Models in Educational Testing
Location: Laurel Hall 205

Bayesian Analysis of Dynamic Item Response Models in Educational Testing

Xiaojing Wang, James O. Berger, and Donald S. Burdick

Item response theory (IRT) models have been widely used in educational measurement testing. When there are repeated observations available for individuals through time, a dynamic structure for the latent trait of ability needs to be incorporated into the model, to accommodate changes in ability. Other complications that often arise in such settings include a violation of the common assumption that test results are conditionally independent, given ability and item dif?culty, and that test item dif?culties may be partially speci?ed, but subject to uncertainty. Focusing on time series dichotomous response data, a new class of state space models, called Dynamic Item Response (DIR) models, is proposed. The models can be applied either retrospectively to the full data or on-line, in cases where real-time prediction is needed. The models are studied through simulated examples and applied to a large collection of reading test data obtained from MetaMetrics,Inc.

Session 4.4 Bias in Missing Data Problems due to Inclusion of Auxiliary Variables
Location: Laurel Hall 206

Bias in Missing Data Problems due to Inclusion of Auxiliary Variables

Felix Thoemmes

The treatment of missing data in the social sciences has changed tremendously during the last decade. Modern missing data techniques such as multiple imputation and full-information maximum likelihood are used much more frequently. These methods assume that data are missing at random. One common approach to increase the likelihood that missing at random is achieved, consists of including many covariates as so-called auxiliary variables. These variables are either included based on data considerations or in an inclusive fashion (taking all available auxiliary variables). However, neither approach accounts for the fact that under a wide range of circumstances there is a class of variables that, when used as auxiliary variables, will always increase bias in the estimation of parameters from data with missing values. I show that this bias exists, quantify it in a simulation study, and discuss possible ways how one can avoid selecting bias-inducing covariates as auxiliary variables.

Concurrent Session 4 Continued 8:45am–9:45am

Symposium Session 4.5: New Developments in Model Selection Procedures
Location: Laurel Hall 302

F-tests with Incomplete Data for Multiple Regression Set-Up

Ashok Chaurasia

One of the most fundamental model selection procedures is the use of a partial F-test. In addition, tests for regression coefficients such as the partial F-test are common in applied research. When dealing with incomplete data, the task of conducting F-tests remains elusive. In this paper we propose a method based on the coefficient of determination to perform partial F-tests with multiply imputed data. Our proposed method can be applied for conducting the "global" F-test (test for all regression coefficients equal to zero), partial F-test (for one or more coefficients, but not all, equal zero), or for equality of regression coefficients. The proposed method is evaluated using simulated data and applied to the HIV Treatment Adherence Intervention Data. This is joint work with Dr. Ofer Harel.

Model Selection for Correlated Predictors of Correlated Outcomes with Applications in Genetic Association Studies

Elizabeth D. Schifano

There is increasing interest in the joint analysis of multiple outcomes in the social sciences and in genome-wide association studies (GWAS), especially for analysis of multiple secondary outcomes in case-control studies. By taking advantage of correlation, both across outcomes and across genetic markers (predictors), one could potentially gain statistical power. We propose novel statistical testing and variable selection procedures to identify sets of correlated genetic markers, as well as individual genetic markers within the set, associated with multiple outcomes measuring the same underlying trait. Both procedures allow for adjustment of covariate effects and are robust to misspecification of the true correlation across outcomes. For multiple secondary outcomes, we account for case-control sampling by appropriately weighting the observations, and additionally propose a weighted Bayesian Information Criterion for variable selection selection. We demonstrate the effectiveness of our approach through theoretical and empirical analysis, as well as in application to investigate genetic associations with smoking behavior, measured using multiple secondary smoking outcomes, in a lung cancer case-control GWAS. This is joint work with Tamar Sofer, David Christiani, and Xihong Lin at Harvard School of Public Health.

Combined Session 4.6: Actor Partner Interdependence Model and Extensions
Location: Laurel Hall 305

Modeling Psychological Subgrouping using Multiple Demographic Composition Variables

Randi L. Garcia

This research project tested multiple models for incorporating two composition variables including the Group Actor Partner Interdependence Model (GAPIM, Kenny & Garcia, 2012) and Faultline models (Lau & Murnighan, 1998). Two studies were conducted, one involving 732 persons in 122 mock juries, a second involving 116 persons who thought they would be interacting in 5-person groups. The analyses modeling nonindependence in groups used the Social Relations Model. There was a consistent positive effect of dyadic similarity in both race and gender on perceptions of similarity. Interactions of race and gender composition variables were found lending some support for non-additive models involving product terms. Furthermore, Study 1 indicated that only the composition of the dyad has effects on similarity, whereas Study 2 found effects involving the others' races. The implications of these results for group research in general as well as the potential benefits of future longitudinal research are discussed.

Interdependent Households Preferences - Case III/APIM Approach

Adam Sagan

In economics of households and consumer behavior, the measurement of interdependent preferences is one of the core area of interest. The present proposal adopt the integrated latent variable APIM models based on Thurstonian measurement models of income allocation preferences using paired comparison/rating scale. Preferences of the four allocation strategies (sequential, two-lane, forked and parallel) are identified with paired-comparison and ranking scales. The APIM model was used to estimate the sizes of actor and partner effects among the family members. The buying roles of the family members are used as predictor of preference interdependencies and the stage of the family life cycle is used as a contextual group-level moderator for the diagnosed effects.

Combined Session 4.7: Multilevel Analysis

Location: Laurel Hall 306

Doubly-Diminishing Returns: An Empirical Investigation of the Impact of Sample Size and Predictor Prevalence on Point and Interval Estimates from Two-Level Linear Models

Bethany A. Bell, Jason A. Schoeneberger, Elizabeth A. Leighton, Stan Haines, Mihaela Ene, Whitney Smiley, and Jeffrey D. Kromrey

Although the multilevel modeling literature provides a few sample size guidelines (e.g., the 30/30 rule), one important limitation is that the findings and recommendations are primarily from studies that examined the statistical functioning of multilevel models with normally distributed continuous predictor variables. Yet, most researchers include both binary and continuous variables in their multilevel models. Thus, little is known about the impact that binary predictor prevalence might have on the functioning of two-level linear models. By examining a variety of sample size combinations and binary variable distributions in two-level linear models, this study adds information about the accuracy and precision of fixed effect point and interval estimates for binary predictors and contributes to our understanding of the behavior of multilevel models under various design conditions.

Analyzing Clustered Data with Single-Level Models: Are Robust Standard Errors an Adequate Solution?

Dan McNeish and Jeffrey Harring

Frequently in applied settings, data are clustered. However, practitioners are not always aware of the necessity of using multilevel models or may not realize their sample came from a hierarchical structure. Mundfrom and Schultz (2002) addressed bias of regression coefficients when multilevel data are analyzed with a single-level model but did not address whether the standard errors exhibited bias as well. The objective of this proposal is to demonstrate how model estimates, including standard errors, for clustered data are affected when hierarchical data are treated in a non-hierarchical fashion and if an intermediate step such as sandwich estimators may alleviate concerns for bias when the clustering effect is low but non-zero. Clustered data will be simulated with varying ICC levels and analyzed with three statistical models ranging in their viability for analyzing multilevel data. The bias of the regression coefficients and their standard errors will be investigated.

Wednesday, May 22nd, 2013

Keynote Address

Laurel Hall 102 10:00am–12:00pm

What on earth are we modeling? Data or Reality? Reflections on structural equations, external validity, heterogeneity and missing data

Judea Pearl, Ph.D

Judea Pearl is a professor of computer science and statistics at UCLA, and distinguished visiting professor at the Technion, Israel Institute of Technology. He joined the faculty of UCLA in 1970, where he currently directs the Cognitive Systems Laboratory and conducts research in artificial intelligence, human reasoning and philosophy of science. Pearl has authored several hundred research papers and three books: Heuristics (1984), Probabilistic Reasoning (1988), and Causality (2000;2009), He is a member of the National Academy of Engineering, the American Academy of Arts and Sciences, and a Fellow of the IEEE, AAAI and the Cognitive Science Society. Pearl received the 2008 Benjamin Franklin Medal for Computer and Cognitive Science and the 2011 David Rumelhart Prize from the Cognitive Science Society. In 2012, he received the Technion's Harvey Prize and the ACM A.M. Turing Award. for the development of a calculus for probabilistic and causal reasoning.

Lunch

Student Union Ballroom 12:00pm–1:30pm

Concurrent Session 5 1:30pm–3:00pm

Symposium Session 5.1: Causal Mediation

Location: Laurel Hall 102

Mediation analysis has been used by psychologists for several decades. Research efforts have been spent extensively to figure out how to parametrically estimate indirect effects and their standard errors. However only recently was there some interest in the causal interpretation of direct and indirect effects. This symposium consists of three speakers that touch upon different aspects of causal mediation. The first talk (Felix Thoemmes, Cornell University, "Methodological Illustration") will give a general introduction to the topic and highlight some recent advances. The second talk (Linda Valeri, Harvard University, "Methodological Innovation") will present a novel approach to estimate causal mediation effects in the presence of misclassification (measurement error) of a binary intermediate variable. The last talk (Judea Pearl, UCLA, "Methodological Evaluation") discusses what kind of causal effects mediation analysis yields and their importance for policy and scientific understanding.

Recent advances in Causal Mediation

Felix Thoemmes

Researchers often use mediation analysis to learn about mechanisms of observed effects in the context of randomized or non-randomized studies. In the past, sophisticated methods have been developed that allow estimation of direct and indirect effects and their associated standard errors. Less attention has been paid to the conditions under which these effects are causally interpretable. The amount of bias can be computed using tracing rules in graphical models and can be shown to be dependent on the inter-correlation of pre-test measures of M and Y and other background variables that may have an effect on mediators and outcomes. A small simulation study along with analytic derivations is presented that shows the amount of bias that can be induced in the direct and indirect effect.

Misclassification of a Binary Mediator – Effects and Remedies

Linda Valeri

Mediation analysis serves to quantify the effect of an exposure on an outcome mediated by a certain intermediate, and to quantify the extent to which the effect is direct. When the mediator is measured with error the validity of mediation analysis can be severely undermined. Hoyle and Kenny (1999) examined the problem of measurement error in mediation analysis for the simple and linear mediation model. VanderWeele et al. (2012) and Valeri et al. (2012) studied the impact of measurement error on a continuous mediator variable when direct and indirect causal effects are estimated using generalized linear models in the presence of exposure-mediator interaction. Misclassification of a binary mediator has been considered by Ogburn et al. (2012) in a non-parametric setting. The contribution of the present work is to study the effects of non-differential misclassification of a binary mediator in the estimation of direct and indirect causal effects when mediation analysis is carried out using a parametric approach and exposure-mediator interaction can be present, and to allow for the correction of misclassification.

What Mathematics Tells us About the Mediation Formula and Why it Matters for Policy Analysis and Scientific Understanding

Judea Pearl

The recent surge in the causal mediation literature has brought along several definitions of causal direct and indirect effect that are slowly being adapted by applied researchers. Among these definitions are the notion of a natural direct effect, a controlled direct effect, and a natural indirect effect. These quantities can be identified under specific assumptions. The proposed talk will explain and contrast the different notions of effects and highlight their importance for policy analysis and answering scientific questions. In addition, the talk will cover identification conditions and show that currently adapted conditions are too stringent and can be considerably relaxed.

Combined Session 5.2: Longitudinal Models

Location: Laurel Hall 202

Fitting Nonlinear Latent Growth Curve Models with Individually-varying Time Points

Sonya Sterba

Individual growth trajectories are often theorized to be nonlinear. Additionally, individuals' measurement schedules may be unique. In a structural equation framework, latent growth model (LGM) applications typically have either (a) modeled nonlinearity assuming some degree of balance in measurement schedules, or (b) accommodated truly individually-varying time points, assuming linearity. Sterba (in press) describes how to fit four popular nonlinear LGMs with truly individually-varying time points, via a definition variable approach. The extension is straightforward for certain nonlinear LGMs (polynomial and structured latent curve), but for shape-factor LGMs requires a re-expression of the model, and for piecewise LGMs requires introduction of a general framework for imparting piecewise structure, along with tools for its automation. This talk describes these four nonlinear LGMs with individually-varying time scores, and applies each to an empirical example. Syntax will be available online. Discussion highlights advantages of modeling nonlinear growth within a SEM framework when time scores individually vary.

Evaluating the Prediction of Growth Factors Under Misspecification of Functional Form

Stephanie Lane and Patrick Curran

Previous research has suggested that, within the context of latent curve modeling (LCM), the ability of a time-invariant covariate (TIC) to predict growth factors in the presence of a misspecified functional form is relatively unaffected, though theory would predict otherwise. The current study sought to evaluate this claim using simulation. We began by generating longitudinal data corresponding to a quadratic functional form. Our simulation varied the covariance between the time-invariant covariate and the quadratic growth factor (positive covariance, no covariance, negative covariance). For each condition, data for R = 500 replications of five waves of N = 500 cases were generated. Our analyses assessed the ability of a time-invariant covariate to predict growth factors when a conditional linear growth model was fitted to the data. Across all conditions, results demonstrated considerable issues with both coverage and bias in the prediction of both the intercept and linear slope factors.

Nonlinear Latent Curve, ALT, and LCARMA Models as Rival Longitudinal Hypotheses: A Monte Carlo Discriminant Study

Stephen A. Sivo and Parul Acharya

Voelkle (2008) found that Bollen and Curran's (2004) Autoregressive Latent Trajectory (ALT) model fits quadratic latent curve data well enough to be deceiving and therefore cautioned against its use. This study examined whether the findings of Voelkle (2008) generalize to the Latent Curve Autoregressive Moving Average (LCARMA) model, and whether information gleaned from cross fitting models to respective data types could aid in distinguishing these models. Initial findings suggest that all models can be distinguished when correctly specified for a suitable data set so long as the underlying process is not weak. Furthermore, Voelkle's (2008) caution against the ALT model is limited to the situation in which the mean of the quadratic process is weak relative to the linear process with which it is associated.

Combined Session 5.3: Statistics and Modeling
Location: Laurel Hall 205

Data Analysis Strategies for High Dimensional Social Science Data: What to do when there are more variables than subjects

Holmes Finch, Maria Hernandez Finch, David E. McIntosh, and Lauren Moss

Social scientists are sometimes faced with small samples on which many measurements are made, particularly when they are working with small, difficult to access populations. Standard statistical methods such as ordinary least squares regression quite often do not work under such conditions. A variety of methods designed for such high dimensional data have been described in the quantitative biology literature, and are applied in two different high dimensional social science studies in the current research. The presentation includes a report of full analysis of the data, as well as a demonstration of the methods using the R software package.

Extending the Robust Means Modeling Framework

Alyssa Counsell, Matthew Sigal, Philip Chalmers, and Robert A. Cribbie

Analysis of variance (ANOVA) is a commonly used test in the social sciences. For ANOVA to be properly implemented, two important assumptions must be met: normality and homogeneity of variance. These assumptions, however, are rarely satisfied in social science data. Many alternatives to the traditional ANOVA exist (e.g, Welch, 1951), but a promising alternative called robust means modeling (RMM) has recently been investigated (Fan & Hancock, 2012). RMM is built upon a structural equation modeling framework, and incorporates procedures to control for non-normality and heterogeneity of variance. This paper extends previous work on RMM using new conditions and estimators to compare RMM's Type I error rates and power to traditional and modified ANOVA-based statistics. We also created RMM pairwise comparison tests and provided an R script for these procedures. Results suggest that RMM methods often perform better than traditional methods, but have liberal error rates with small sample sizes.

Using Monte Carlo simulations to Understand Probabilities and Modeling: Bringing causality into the teaching of introductory statistical modeling

Emil Coman, Maria Coman, Eugen Iordache, Lisa Dierker, and Russell Barbour

We demonstrate an innovative use of the MC simulations for basic applications of causal models with direct utility in teaching introductory statistics. We show how MC simulations can be used in the introductory stages of explaining basic statistical concepts. We specifically introduce an alternative mode of testing SEM models, which makes explicit the assumptions and inner workings of SEM model testing: we use a 'causally blind' Data Generating Process (DGP) which is the saturated model, i.e. each variable assumed to correlate to all other variables, with covariances equal to the sample statistics. We begin this methodological illustration with describing the longitudinal data from a health intervention aimed at improving physical activity (PA) using new technologies (email and SMS). We first test a simple regression and a mediation model, then we setup Monte Carlo models with the population model being simply the covariances and means of the sample data, and the models to be tested the initial ones we estimated. We compare the outputs of the two, and conclude with practical implications for incorporating Monte Carlo simulation techniques in teaching introductory statistics from a built-in causal perspective, using Mplus.

Combined Session 5.4: Multiple Raters
Location: Laurel Hall 206

Approximate Measurement Invariance in Rater-mediated Assessments: A Random Item Effects Item Response Model for Measuring Teaching Quality with Classroom Observations
Ben Kelcey, Dan McGinn, and Heather Hill
Rater-mediated assessments use raters to evaluate targeted proficiencies in order to provide scores for participants. An important assumption underlying meaningful comparisons of scores in these types of assessments is that measurement is invariant across raters. Measurement invariance across raters suggests that raters use indicators in similar ways so that participants can be placed on a common scale. However, research has shown that even after extensive training there are important differences in how raters interpret evidence and that these differences potentially introduce variability in the structure of the scale established by the guiding rubric/instrument. This proposal draws on the concept of approximate measurement invariance to develop a random item effects item response model for measuring latent traits using rater-mediated assessments. We apply the method to the measurement of teachers' quality through classroom observations and investigate the extent to which it improves model fit and the predictive validity of the observation scores.

Level-Specific and Multilevel Reliability
Joseph A. Olsen
The motivating psychometric objective of this paper is to examine how to effectively assess the reliability of multiple item measures gathered with multilevel research designs. It is somewhat ironic that group differences are typically ignored when attempting to assess the reliability of measures explicitly designed to assess group-level properties. We then reviewed approaches to assessing the reliability of multilevel measurement using hierarchical linear modeling, generalizability theory, and multilevel structural equation modeling. We develop methods for estimating the reliability of multi-item measures of multilevel constructs using multilevel structural equation modeling, and present level-specific and overall multilevel reliability coefficients for the cases of both fixed and random items. In applying these methods to a specific research problem, we found that models based on generalizability theory and multilevel structural equation modeling fit the data better than those based on hierarchical linear modeling.

Reliability and Validity by the Multiple Indicator by Multiple Trait by Multiple Source by Multiple Occasion Model: Application to an ADHD Rating
Scale by G. Leonard Burns, Mateu Servera, and Christian Geiser
Confirmatory factor analysis (CFA) was applied to a multiple indicator (individual item ratings) by multiple trait (ADHD-inattention, ADHD-hyperactivity/impulsivity, and Academic Impairment) by multiple source (mothers and fathers) by multiple occasion (test and retest) model to obtain a comprehensive evaluation of the reliability and validity of an attention-deficit/hyperactivity disorder (ADHD) rating scale. The participants were the mothers and fathers of 802 Spanish first grade children. Support was found for the invariance of the model parameters between sources and across occasions. Convergent and discriminant validity also occurred for the ADHD-IN, ADHD-HI and Academic Impairment factors between sources and across occasions. The application of latent variable procedures to this model allows for comprehensive evaluation of the reliability and validity of rating scales.

Combined Session 5.5: Educational and Health Related Applications
Location: Laurel Hall 301

Realistic Models for School-based Longitudinal Sociometric Data: Multilevel Cross-classified Poisson and Negative Binomial Approaches

Richard A. Faldowski and Heidi Gazelle

Longitudinal sociometric nomination data, which is based on counts of the number of classroom peers who think a particular characteristic applies to their classmate, often have distributional properties that violate standard assumptions of linear multilevel longitudinal statistical models. In this presentation, we propose multilevel longitudinal cross-classified Poisson and negative binomial models as more appropriate alternatives that mitigate the distributional irregularities. By misapplying linear multilevel longitudinal analysis methods, as is commonly found in the sociometric literature, and comparing the results to ones obtained under multilevel longitudinal cross-classified Poisson and negative binomial models of the same data, we show how substantively different conclusions result. The discrepancies are so great that they raise serious questions about the validity of conclusions drawn from studies that have employed standard linear multilevel longitudinal methods.

A Mathematical Evaluation of the Effect of Disclosure on HIV Transmission Rate in Men who Have Sex with Men

Ann A. O'Connell, Sandra J. Reed, and Julianne M. Serovich

Interventions to assist HIV+ persons in disclosing their status to potential sexual partners can play an important role in curbing rates of HIV transmission, particularly among men who have sex with men, a disproportionately represented group in national HIV/AIDS statistics. Using baseline data from a randomized clinical trial of an HIV disclosure intervention, we present estimates of reduction in HIV transmission rate due to disclosure. Based on the methods of Pinkerton & Galletly (2007), we develop a mathematical probability model for evaluating effectiveness of the intervention through comparison of reduction in HIV transmission rate due to disclosure relative to the disclosure-related HIV transmission risk in a standard-of-care (case-management) control condition. We then present adjustments based on partner serostatus. Features of the disclosure intervention are discussed, and implications for design and evaluation of disclosure interventions are highlighted.

Patterns of Drug use Measured using Latent Class and Latent Transition Analysis with Covariates in a representative sample of Incarcerated adults in Puerto Rico

Rafael R. Ramirez, Carmen Rivera Medina, Jose Noel Caraballo, Jose Ruiz Valcarcel, Glorimar Caraballo Correa, and Carmen Albizu

This paper presents the results of using LCA and LTA to identify latent classes of drug use patterns in a large representative sample(n =1,179) of incarcerated adults in Puerto Rico. The principal goal of the paper was to compare 3 different methods of estimating the multinomial regression coefficients of latent classes on study covariates. A latent class model was estimated for drug use patterns before incarceration and a separate model was estimated for drug use patterns while incarcerated. Results indicated that a 1 step joint model estimation and the new 3 step procedure available in MPLUS version 7.0 provided nearly identical results for the LCA models in our data. However the pseudo class method produced significant attenuation even at entropy values as high as .7 Only when entropy reached .8 for the LCA of drug use during incarceration were the results of the pseudo class method comparable to the other two. For those reserchers who for theoretical/methodological reasons don't want covariates to affect the measurement model of LCA and/or LTA, the new 3 step method available in MPlus appears very promising at the entropy values present in our data.

Concurrent Session 5 Continued 1:30pm-3:00pm

Ordinal Regression Analysis: Fitting Stereotype Logistic Regression Models to Educational Data
Xing Liu
The stereotype logistic (SL) model is an alternative to the proportional odds (PO) model when the proportional odds assumption is untenable. This model seems to be underutilized. One major reason is the constraint of current statistical software packages. SPSS cannot perform the SL regression analysis, and SAS does not have the procedure developed to directly estimate the model. Furthermore, researchers may see different equations for the SL model are used in literature, which results in confusion and misunderstanding in application. The purpose of our proposed study is to illustrate the stereotype logistic (SL) regression model, apply it to estimate mathematics proficiency level of high school students with Stata and R, and compare the results of fitting the SL model between these two packages.

Symposium Session 5.6: Understanding Patterns, Variations, And Trends In Adolescent And Young Adult Smoking: Opportunity For New Insights Or Source For Headaches?
Location: Laurel Hall 305

Increasingly, advances in technology and research methodology have enabled researchers to examine naturalistic smoking behaviors, their development and correlates over time. At the same time, these data represent a challenge, as they require more complex statistical modeling than applied researchers are ordinarily trained to use. In this symposium we use data from large-scale, multi-site, naturalistic, longitudinal studies on adolescent and young adult smoking to illustrate how new insights can be gained using modern modeling techniques. Presentations will address: Does the strength of smoking predictors change over the course of nicotine dependence? How does smoking level affect variations in mood? Do subgroups exist whose smoking patterns evolve differently over time? Answers to these questions inform both tailoring and targeting of intervention efforts. Together, this symposium will showcase the use of advanced longitudinal modeling to capture key characteristics of adolescent and young adult smoking, applied to both long-term (spanning years) and fine-grained (day-to-day, within-day) longitudinal processes.

Smoking Patterns in First Year College Students: Trends, Correlates and Outcomes
Bettina B. Hoeppner and Nancy P.Barnett
Unlike older smokers, young adult smokers frequently engage in light and intermittent smoking. It remains unclear how stable such smoking patterns are over time, as substantial variability exists between these smokers. This study identifies subgroups that may exist among college student smokers based on the trajectory of their smoking during the first year of college. Subgroups are contextualized through an examination of pre-college characteristics and external cues (i.e., weekday periodicity, same-day alcohol use) of daily smoking. While there was a diversity of smoking patterns, smoking behaviors remained relatively stable during the first year of college. Our findings further suggest that less frequent smoking was tied more closely to external stimuli and motives, underscoring the importance of tailoring smoking interventions for these smokers.

Decisions about Covariate Specification in Time-Varying Effects Models: Does it Matter?

Jennifer S. Rose, Arielle S. Selya, Lisa C. Dierker, Donald Hedeker, Robin Mermelstein

The relative roles of smoking quantity and nicotine dependence (ND) in the development of regular smoking are difficult to disentangle, as effects of these measures on smoking regularity vary across time. Recent methodological advances in time-varying effect models (TVEMs) allowing effects of predictors on outcomes to vary over time permit a more direct examination of how smoking quantity and ND influence smoking regularity. In addition to typical longitudinal mixed model considerations, however, TVEMs require researchers to make decisions about using covariates with fixed values or time-varying values to estimate effects that are averaged over time or vary over time. Different specifications of covariates and their effects may produce different results. The current research examines how different specifications affect results of TVEMs investigating the association between smoking quantity and smoking regularity (# days smoked) controlling for time-varying ND effects. Overall, TVEM covariate specification is an important consideration as conclusions about associations between quantity and smoking regularity differ depending on specification.

Nicotine Dependence-Varying Effects of Smoking Events on Momentary Mood Changes among Adolescents

Arielle Selya, Nicole Updegrove, Lisa Dierker, Jennifer Rose, Xianming Tan, Donald Hedeker, Runze Li, Robin Mermelstein

Theoretical work on nicotine addiction has postulated that positive reinforcement plays an initial role in the development of regular smoking behavior, while negative reinforcement maintains smoking behavior at later stages. With ecological momentary assessment (EMA), it is now possible to test these theories by examining smoking-related mood changes (i.e. the difference in positive affect (PA) and negative affect (NA) after smoking relative to before smoking) as a result of individual smoking events. The current study uses a sample of light adolescent smokers to examine variation in smoking-related mood changes across participants with varying levels of nicotine dependence (ND). Our findings, to some extent, support theoretical work suggesting that positive reinforcement plays a role in early stages of dependent smoking. However, the role of negative reinforcement beyond early stages of dependent smoking is not supported. Future research should examine other contributing factors to smoking behavior among highly nicotine dependent smokers such as preemptive smoking to prevent withdrawal symptoms.

A Bivariate Location-Scale Mixed-Effects Model with Application to Mood Variation among Youth Smokers

Oksana Pugach, Donald Hedeker, and Robin Mermelstein

Ecological momentary assessments (EMA) are useful in understanding both between and within subject dynamic changes in smoking and mood. Modeling two moods (positive and negative affect; PA and NA) simultaneously will better enable researchers to explore directly the association of mood variables and what influences them at both the momentary and subject level. A total of 461 adolescents completed the baseline EMA wave resulting in 14,105 completed random prompts. Results suggest that non-smokers had more consistent positive and negative moods compared to 1-cigarette smokers. Among those who smoked, as the number of smoking episodes increased, both PA and NA became more consistent. The present paper illustrated an application of the bivariate location-scale mixed-effects model allowing for separate estimation of the between and within subject variances, as well as the joint modeling of two mood constructs.

Closing Keynote Address

Laurel Hall 102 3:15pm–4:45pm

Late-breaking news: Some exciting new methods!

Bengt Muthen, Ph.D

Bengt Muthén obtained his Ph.D. in Statistics at the University of Uppsala, Sweden and is Professor Emeritus at UCLA. He was the 1988-89 President of the Psychometric Society and the 2011 recipient of the Psychometric Society's Lifetime Achievement Award. He has published extensively on latent variable modeling and is one of the developers of the *Mplus* computer program, which implements many of his statistical procedures.

Dr. Muthén's research interests focus on the development of applied statistical methodology in areas of education and public health. Education applications concern achievement development while public health applications involve developmental studies in epidemiology and psychology. Methodological areas include latent variable modeling, analysis of individual differences in longitudinal data, preventive intervention studies, analysis of categorical data, multilevel modeling, and the development of statistical software (namely Mplus!).

Modern Modeling Methods Conference
May 20-21, 2014
University of Connecticut

The Modern Modeling Methods (M_3) conference is an interdisciplinary conference designed to showcase the latest modeling methods and to present research related to these methodologies.

Call for Papers- Proposals due January 15, 2014

The Modern Modeling Methods Conference (M3) will be held at the University of Connecticut on May 20-21, 2014. We are currently soliciting both methodological research papers and papers that illustrate methodological techniques in the area of modeling, broadly defined. Papers related to multilevel modeling, structural equation modeling, mixture modeling, and longitudinal modeling are especially encouraged. The deadline for submissions is January 15, 2014. Methodological research proposals should be no longer than 1000 words and should include purpose, background, methods, results, discussion, and significance. Methodological illustration papers should be no longer than 1000 words and should include a description of the methodology to be illustrated as well as an outline of the paper/talk. Please include a 100-150 word abstract to be used for the conference program.

For more information about the conference, or to submit a proposal, go to www.modeling.uconn.edu. If you have questions, please email D. Betsy McCoach at betsy.mccoach@uconn.edu.

MEA Program

more information:
http://neag.uconn.edu/mea/

contact:
Dr. D. Betsy McCoach
betsy.mccoach@uconn.edu

GAANN Fellowships are available for PhD students in Measurement, Evaluation, and Assessment

The Measurement, Evaluation, and Assessment (MEA) program in the Educational Psychology Department at University of Connecticut is seeking outstanding students with an interest in measurement, assessment, evaluation, educational research and statistics, or quantitative methodology for our graduate programs. We offer a MA and a PhD in Educational Psychology, with an area of concentration in Measurement, Evaluation, and Assessment. Our students typically pursue careers in academia, in state or federal agencies, in testing companies (such as ETS, Pearson, Riverside) or in research and development firms (such as MDRC, AIR, Education Development Center). Our students come from a variety of undergraduate majors including Psychology, Statistics, Mathematics, Education, Economics, and Sociology, just to name a few.

We are excited to announce that we have been awarded a GAANN Grant from the US Department of Education. This competitive grant helps to provide us with significant additional financial resources to support PhD students in our MEA program who:

- Demonstrate academic merit and maintain satis factory academic progress toward a PhD degree;
- Demonstrate financial need. The amount of the fellowship award is directly tied to a student's demonstrated need. Need is determined by the University of Connecticut's financial aid office and is based on student responses to the Free Application for Federal Student Aid (FAFSA)
- Are US citizens or permanent residents.

The GAANN fellowship covers tuition, fees, and health insurance, as well as a stipend of up to $30,000 per year (the actual amount is determined by the financial aid office, and is based on the student's determined level of need). We also offer graduate assistantships on a competitive basis to incoming students, and these are awarded based on merit, not need. We are currently actively seeking and recruiting bright, motivated, quantitatively oriented students who are interested in pursuing a PhD in Measurement, Evaluation, and Assessment, starting in the fall of 2013. All graduate assistantships and GAANN awards are competitive awards.

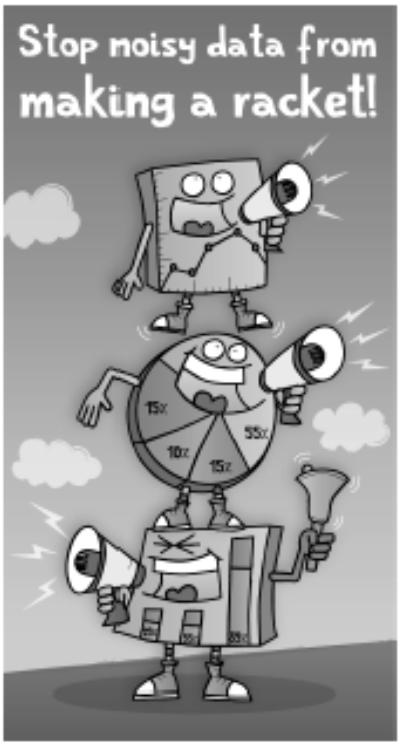
Statistical Innovations

Thinking outside the brackets!™

statisticalinnovations.com

Uncover Meaningful Segments with Latent GOLD® and LG Choice 5.0

The #1 program for latent class modeling



The key to getting clear, noise-free segments is to use the most appropriate latent class models.

Latent GOLD® 5.0

- **Cluster.** Outperforms K-means and hierarchical clustering methods.
- **Factor.** Extends traditional factor analysis in many ways.
- **Key Driver Regression.** Different effects for each latent segment.
- **Latent (Hidden) Markov and Growth Modeling.** Identify different patterns of change over time for each segment.

Latent GOLD® Choice 5.0

- **Latent Class Model for Discrete Choice Designs**
- **MaxDiff Models.** Accepts Sawtooth and SPSS data formats.
- **Conjoint Models.** Different utilities for different segments.
- **Models for Allocation/Constant-Sum Data.** Uses replication weights to handle different points allocated to different options.

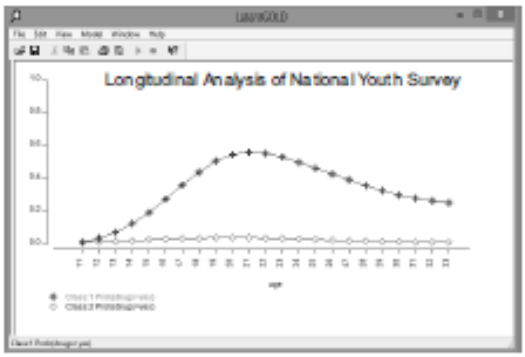
Segmentation with Longitudinal Data has Never Been Easier!

New GUI Module for Mixture Latent Markov / Transition Modeling

Point-and-Click

The user-friendly latent Markov module in Latent GOLD® lets you develop meaningful models *without* complicated syntax. And create powerful interactive longitudinal plots with ease!

Be sure to attend Dr. Jay Magidson's presentation on mixture latent Markov models Tuesday at 3:15pm.



SI Software
Latent Class Modeling
Regularized Regression
Tree Based Analysis

Training
Software Tutorials
Online Courses
Onsite Training

Consulting
Project Advice
Model Development
Study Design

375 Concord Ave., Suite 007
Belmont, MA 02478
Phone 617.489.4490
Web: statisticalinnovations.com
Email: wil@statisticalinnovations.com