Modern Modeling Methods Conference

May 21–22, 2014

UCONN
UNIVERSITY OF CONNECTICUT

2014 Modern Modeling Methods Conference

University of Connecticut

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, Susan Rasman, Danielle Bousquet, Dani Yomtov, 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 2014 Modern Modeling Methods conference attendees for coming and being a part of the fourth annual M^3 conference! I hope to see you all back in Storrs on May 19-20, 2015 for the fourth annual Modern Modeling Methods conference. Proposals for concurrent sessions will be due January 30, 2015, and can be submitted online at our website: **www.modeling.uconn.edu**

D. Betsy McCoach, Ph.D.
2014 Chair, Modern Modeling Methods Conference
Professor and Program Coordinator, Measurement, Evaluation, and Assessment Program
Educational Psychology Department, Neag School of Education, UCONN

Monday, May 19th: Pre-conference Workshop

Advances in Latent Variable Modeling Using Mplus 7.2

Bengt Muthén

University of California, Los Angeles

This 1-day course gives an overview of unique latent variable modeling opportunities in Mplus related to mediation analysis with causal effects, factor analysis, IRT, mixture modeling, multilevel modeling, and Bayesian analysis. Examples with Mplus input and output are discussed. The emphasis is on applications rather than technical aspects. After a brief overview of the many uses of finite mixture modeling, applications of new mixture modeling developments are discussed. One major development goes beyond the conventional mixture of normal distributions to allow mixtures with flexible non-normal distributions. This has interesting applications to cluster analysis, factor analysis, SEM, and growth modeling. The talk focuses on applications of Growth Mixture Modeling for continuous outcomes that are skewed. Examples are drawn from national longitudinal surveys of BMI as well as twin studies. Extensions of this modeling to the joint study of survival and non-ignorable dropout are also discussed. New features of the May 2014 Mplus Version 7.2 are demonstrated.

9:00 am - 5:30 pm in Laurel Hall 102

The following topics will be emphasized:

Causal inference in mediation analysis offers counterfactually-based causal definitions of direct and indirect effects for models with categorical and count outcomes, drawing on research by Robins, Greenland, Pearl, VanderWeele, Imai and others. Background information is provided in the following papers on the Mplus website.

Item response theory modeling in the general latent variable framework of the Mplus program offers many unique features.

Mixture modeling includes latent class analysis, item response mixture modeling, factor mixture analysis, latent transition analysis, latent class growth analysis, growth mixture modeling, and survival mixture analysis.

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!). For more information about Bengt Muthén, check out his website**: www.statmodel.com/bmuthen/**

Monday, May 20th

5:30 pm – 7:30 pm

Nathan Hale Inn, Mansfield Room

Welcome Reception

All conference participants 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.

Tuesday, May 20th

7:30 am – 8:00 am in Laurel Hall Atrium

REGISTRATION AND BREAKFAST

8:00 am – 9:30 am in Laurel Hall 102

Keynote Address:

Bengt Muthén, Ph.D.

Advances in Mixture Modeling

CONCURRENT SESSION #1 9:45 am – 11:45 am

Session 1.1: Multitrait-Multimethod Models (Laurel Hall 201)

Neglect the Structure of Multitrait-Multimethod Data at your Peril: Implications for Associations with External Variables.
Laura Castro-Schilo, Keith F. Widaman, & Kevin J. Grimm

Uncrossing the Correlated Trait-Correlated Method Model for Multitrait-Multimethod Data.
Laura Castro-Schilo, Kevin J. Grimm, & Keith F. Widaman

Bayesian Versus Frequentist Estimation of Multitrait-Multimethod Confirmatory Factor Models.
Jonathan L. Helm & Laura Castro-Schilo

Session 1.2: Individual-Participant-Data Meta-Analytic Modeling: Measurement Challenges & Network Analysis (Laurel Hall 202)

Modeling Network Individual Data Meta-Analysis.
Tania B. Huedo-Medina, Argie Veroniki, & Dylan Yaworski

Generalized Linear Modeling to Integrate Measures in Individual-Participant-Data Meta-Analysis.
Nathan Lally, Xiaoran Li, Francisca Galindo-Garre, Maria Dolores Hidalgo, & Tania B. Huedo-Medina

Response Conversion Under Item Response Theory when Conducting Individual-Participant-Data Metaanalysis.
Xiaoran Li, Nathan Lally, Francisca Galindo-Garre, Maria Dolores Hidalgo, & Tania B. Huedo-Medina

Session 1.3: Practical Issues in Missing Data Analysis (Laurel Hall 205)

You Shall Combine, But How? A Comparison of Methods to Combine Chi-Square Using Multiple Imputation.
Mauricio Garnier-Villarreal, Alexander M. Schoemann, & Todd D. Little

What to do with Incomplete Nominal Variables? A Comparison of Recommended Techniques for Creating Multiple Imputations of Unordered Factors.
Kyle M. Lang

Removing Bias Due to Practice Effects Using a Three-Form Planned Missing Data Design.
Terrence D. Jorgensen, Mijke Rhemtulla, Alexander Schoemann, Brent McPherson, Wei Wu, & Todd D. Little

A Proof-of-Concept Simulation of the Accelerated Longitudinal Planned Missing Design for Latent Panel Modeling.
Luke McCune

Session 1.4: Bayesian and Simulation Techniques (Laurel Hall 206)

Bayesian Outlier Detection in Structural Equation Models.
Davood Tofighi

Comparative Study of Two Calibration Methods for Micro-Simulation Models.
Stavroula Chrysanthopoulou

Testing Applications of Bayesian Statistics with Informed Priors in Multilevel Analysis of Small Data Sets with Complex Structures: A Monte Carlo Study.
Tyler Hicks, George MacDonald, Jeff Komrey, Eun Sook Kim, Jeanine Romano, & Sandra Archer

Session 1.5: Issues in Latent Class Analysis (Laurel Hall 301)

Identifying the Correct Number of Classes in Latent Profile Analysis: The Impact of Sample Size, Profile Distribution, and Model Specification.
Sara K. Johnson

Obtaining Meaningful Latent Class Segments with Ratings Data by Adjusting for Level, Scale and Extreme Response Styles.
Jay Magidson

Session 1.6: Causal Modeling (Laurel Hall 302)

Empirical Tests of Directional Dependence using Real Datasets.
Felix Thoemmes, Sarah Moore, & Marina Yamasaki

How do we Combine Two Treatment Arm Trials with Multiple Arms Trials in Ipd Meta-Analysis? An Illustration with College Drinking Interventions.
David Huh, Eun-Young Mun, & David C. Atkins

Adjusting for Selection Bias in Assessing the Relationship Between Sibship Size and Cognitive Performance.
Gebrenegus Ghilagaber & Linda Wänström

CONCURRENT SESSION #2 11:30 am – 12:30 pm

Session 2.1 (Laurel Hall 201)

Discrete and Continuous Time Models from the Perspective of Structural Equation Modeling.
Pascal R. Deboeck & Aaron M. Boulton

Session 2.2 (Laurel Hall 202)

Empirical SEM Confidence Intervals by Bootstrapping.
Craig Michael Krebsbach

Session 2.3 (Laurel Hall 205)

Testing the Homogeneity of Variance Assumption: An Investigation of the Performance of Ten Different Approaches in One-Factor ANOVA Models.
Harold Holmes, Patricia Rodriguez de Gil, Issac Li, Aarti Bellara, Yi-Hsin Chen, Tyler Hicks, & Eun Sook Kim

Session 2.4 (Laurel Hall 206)

Bayesian Model Averaging for Propensity Score Analysis.
David Kaplan & Jianshen Chen

Session 2.5: Advanced Longitudinal Modeling Techniques (Laurel Hall 301)

Using a Random-Effects Tobit Model to Analyze Longitudinal Data with a Large Proportion of Zeroes.
Haiyi Xie, Gary R.Bond, Robert.E. Drake, & Gregory J. McGugo

Investigating Differential Changes using Mixture Latent Change Scores (MLCS) Modeling.
Emil Coman, Judith Fifield, Jack J. McArdle, & Monique Davis-Smith

12:30 pm – 1:40 pm in the Student Union Ballroom

LUNCH

CONCURRENT SESSION #3 1:45 pm – 3:15 pm

Session 3.1: Explanatory Item Response Theory Models (Laurel Hall 201)

The Effects of Reader's Characteristics, Text Genre, and Comprehension Processes on Reading Comprehension.
Paulina Kulesz, David Francis, Marcia Barnes, Jack Fletcher, Amy Barth, & Mary York

Reader and Text Characteristic's Contribution to Inferential Processing in Adequate and Struggling Comprehenders.
Yusra Ahmed, David Francis, Marcia Barnes, Jack Fletcher, Amy Barth, & Mary York

A Diagnostic Cognitive Model (DCM) of Algebraic Processing.
Justin Neil Young, Tammy D. Tolar, David J. Francis, & Jeffrey J. Morgan

Session 3.2: Exploring the lavaan Ecosystem: Packages to Extend the Capability of lavaan (Laurel Hall 202)

***semTools*: Useful Tools for SEM.**
Alexander M Schoemann, Sunthud Pornprasertmanit, & Patrick J. Miller

Getting the Most out of your Family of Data with the R-package fSRM.
Lara Stas, Felix Schönbrodt, & Tom Loeys

***semPlot*: Unified Visualizations of Structural Equation Models.**
Sacha Epskamp

lavaan.survey.
Daniel Oberski

***simsem*: SIMulated Structual Equation Modeling in R.**
Sunthud Pornprasertmanit, Alexander M. Schoemann, Patrick J. Miller

Graphical Structural Equation Modeling with Onyx.
Timo von Oertzen, Andreas Brandmaier, & Siny Tsang

Session 3.3: New Developments in the Analysis of Incomplete Data (Laurel Hall 205)

A Survey of Missing Data Techniques Employed in Genome-Wide Association Studies.
Gregory J. Matthews & Ofer Harel.

Clustering Incomplete Data using Normal Mixture Models.
Chantal D. Larose, Ofer Harel, & Dipak Dey

Comparison of Transformations used in the Estimation of R^2 and Adjusted R^2 in Incomplete Data Sets using Multiple Imputation.
Valerie Pare & Ofer Harel

Session 3.4: Bayesian Modeling (Laurel Hall 206)

Bayesian SEM Perspectives (From the Hills of Asymptotia).
Albert Satorra

Bayesian Estimation for Structural Equation Models with Sample Weights.
Bengt Muthen

Bayesian Analysis of Multiple Indicator Growth Modeling using Random Measurement Parameters Varying Across Time and Person.
Tihomir Asparouhov

Session 3.5: Modeling Treatment and Causal Effects (Laurel Hall 301)

Treatment Effects in Randomized Controlled Trials with Noncompliance and Missing Data.
Shu Xu, Michael F. Lorber, Amy M. S. Slep, Richard E. Heyman, & Danielle M. Mitnick

Estimation of Heterogeneity of Treatment Effect using EM Algorithm.
Evgeniya Reshetnyak, Ying Liu, Barry Rosenfeld, & William S. Breitbart

Probing Causal Mechanisms and Strengthening Causal Inference by Means of Mixture Mediation Modeling.
Emil Coman, Judith Fifield, Suzanne Suggs, Deborah Dauser-Forrest, & Martin-Peele Melanie

Session 3.6: Modeling and Mediation (Laurel Hall 302)

Parametric Sensitivity Analysis Thresholds for Mediation Effects.
Ben Kelcey, Kenneth Frank, & Michael Seltzer

The Markov Equivalence Class of the Mediation Model.
Felix Thoemmes, Wang Liao, & Marina Yamasaki

Mediation Analysis in AB/BA Crossover Studies.
Haeike Josephy, Tom Loeys, & Stijn Vansteelandt

3:30 pm – 5:00 pm in Laurel Hall 102

Keynote Address:

James Robins, Ph.D.

Estimation of Causal Effect of Exposures that Change Over Time: Methods and Case Studies

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

POSTER SESSION AND RECEPTION

Please join us in the Student Union Ballroom to visit the 44 posters. The reception includes appetizers and an open bar.

Wednesday, May 21st

7:30 am – 8:00 am in the Laurel Hall Atrium

BREAKFAST AND REGISTRATION

8:00 am – 9:30 am in Laurel Hall 102

Keynote Address:

Sophia Rabe-Hesketh, Ph.D.

Simple Methods for Handling Non-Randomly Missing Data

CONCURRENT SESSION #4 9:45 am – 10:45am

Session 4.1: New Developments in Variable Selection for Regression Models (Laurel Hall 106)

An Introduction to Variable Selection for Regression Models.
Ofer Harel

An Empirical Bayes Approach to Variable Selection and QTL Analysis.
Haim Bar

Model Selection Through Sparse Estimation in Finite Mixture Regression Models.
Elizabeth D. Schifano, Robert L. Strawderman, & Martin T. Wells

Session 4.2 (Laurel Hall 107)

Structural Equation Modeling and Factor Indeterminacy.
Edward Rigdon

Session 4.3: Longitudinal Models (Laurel Hall 108)

Fit Criteria Performance and Parameter Estimate Bias in Growth Curve Models with Small Samples in the SEM Framework.
Dan McNeish

Issues in Latent Growth Modeling with Longitudinal Public-Release Data.
Ming Li, Jeffrey Harring, & Laura Stapleton

Session 4.4: Effect Size and Confidence Intervals (Laurel Hall 109)

An Investigation of Accuracy and Precision of the Generalized Eta-Squared Effect Size Based on Various Research Designs.
Patrice Rasmussen, Patricia Rodriguez de Gil, Anh Kellerman, Thanh Pham, Jeanine Romano, Yi-Hsin Chen, & Jeffrey Kromrey

Robust Confidence Intervals for Effects Sizes in Multiple Linear Regression.
Paul Dudgeon

Session 4.5: Latent Transition Analysis (Laurel Hall 110)

Implementing the 3 Step Latent Transition Analysis in MPLUS using a Sub-Population from a Complex Sample Design.
Rafael R Ramirez, Jose Noel Caraballo, & Carmen Rivera Medina

Session 4.6: Advanced Modeling of Healthcare Data (Laurel Hall 111)

Hierarchical Bayesian Exploratory Factor Analysis for Health Care Quality Utilization and Quality Data.
Alan M. Zaslavsky & A. James O'Malley

Modeling Multilevel Data with Cross-Classified Outcome Variables using Kronecker-Structured Covariance Matrices, with Applications to Multivariate-Outcome Random-Coefficient Models for Healthcare Quality Data.
Alan M. Zaslavsky & Laura A. Hatfield

CONCURRENT SESSION #5 11:00 am – 12:00 pm

Session 5.1: Causal Modeling in Communication (Laurel Hall 106)

Iterative Meta-Causal Analysis: Modeling the Impact of Job Loss on Communication and Personality.
Mark Hamilton

Testing the Viability of Alternative Structures with a Distributed Computing System.
James Watt & Mark Hamilton

The Influence of Synchrony and Sensonry Modality on the Person Perception Process in Computer-Mediated Groups.
Kristine Nowak & James Watt

Analyzing and Modeling Behavioral Interaction Data.
Arther Vanlear & Teharan Davis

Session 5.2 (Laurel Hall 107)

An Introduction to Integrative Data Analysis in the Behavioral and Social Sciences.
Jennifer Walsh

Session 5.3 (Laurel Hall 108)

A Demonstration of a New Linear Modeling Procedure in SPSS Statistics: Automatic Linear Modeling (*linear*).
Hongwei Yang

Ordinal Logistic Regression Models for Complex Sample Survey Data Using Stata, SPSS and SAS.
Xing Liu.

Session 5.4 (Laurel Hall 109)

DataToText: Consumer-Oriented Dyadic Data Analysis using R.
David Kenny

Session 5.5: Measurement Modeling (Laurel Hall 110)

An Investigation of the Alignment Method for Detecting Measurement Non-Invariance Across Many Groups with Dichotomous Indicators.
Jessica Kay Flake, Erin Strauts, Betsy McCoach, Jane Rogers, & Megan Welsh

Dimensionality at Multiple Levels: Examining NAEP Mathematics with An Exploratory, Multilevel Item Factor Analysis Model.
Nathan Dadey & Gregory Camilli

Session 5.6 (Laurel Hall 111)

Evaluating Measurement Equivalence and Translation Effectiveness of a Customer Engagement Instrument Across National Cultures and Types of Customers.
Dan Yu & Yongwei Yang

A Mixture Model for Nuptiality Data with Long-Term Survivors.
Paraskevi Peristera & Gebrenegus Ghilagaber

12:00 pm – 1:00 pm in the Student Union Ballroom

LUNCH

Concurrent Session #6 1:00 pm – 2:30 pm

Session 6.1 (Laurel Hall 106)

A Workshop on Bayesian Nonparametric Regression Analysis.
George Karabatsos
*Attendees are welcome to bring laptops to follow along with the computations

Session 6.2: Propensity Score Analysis: Empirical Investigations of Common Problems and their Impacts on Treatment Effect Estimates (Laurel Hall 107)

Impact of Measurement Error in Propensity Score Analysis.
Eun Sook Kim

Treatment of Missing Data in Propensity Score Analysis.
Patricia Rodriguez de Gil

Single-Level vs. Multi-level Propensity Scores with Nested Data.
Patricia Rodriguez de Gil & Jeffrey Kromrey

Covariate Balance in Propensity Score Models: Much Ado about Nothing?
Jeffrey Kromrey

Session 6.3: Missing Data (Laurel Hall 108)

A Further Look into Planned Missing Data Designs in Analysis of Change.
Fan Jia, Wei Wu, Mijke Rhemtulla, & Todd D. Little

A Latent Variable Chained Equations Approach for Multilevel Multiple Imputation.
Craig K. Enders & Brian T. Keller

Session 6.4: Item Response Theory (Laurel Hall 109)

Unipolar Item Response Models.
Joseph F. Lucke

Robustness of Mixture Item Response Models to Two Correlated Sources of Differential Item Functioning.
Erin Strauts & Jessica Kay Flake

Realistic IRT Item Parameter Generation for Monte Carlo Simulation Studies.
Ling Ning, Cindy Walker, & Bo Zhang

Session 6.5: Modeling Dyadic Data (Laurel Hall 110)

Modeling Growth in Dyads at the Group Level.
Thomas Ledermann & Siegfried Macho

A Structural Equation Model Of Dyadic Discrepancy Over Time.
Holly Laws, Aline Sayer, Paula Pietromonaco, & Sally Powers

The Actor-Partner Interdependence Model For Categorical Dyadic Data: An Introduction to GEE.
Tom Loeys, William Cook, Olivia De Smet, Anne Wietzker, & Ann Buysse

Session 6.6: Modeling Secondary Data (Laurel Hall 111)

Using the Pair-Wise Likelihood Method to Analyze Large Datasets with Discrete Responses.
Maria T. Barendse, Frans J. Oort, Marieke E. Timmerman, & Y. Rosseel

Maximum Likelihood Adjustment of Anticipatory Covariates in Analyzing Retrospective Survey Data.
Gebrenegus Ghilagaber & Rolf Larsson

2:45 pm – 4:15 pm in Laurel Hall 102

Keynote Address:

Edward Vytlacil, Ph.D.

Accounting for Individual Heterogeneity in Treatment Effect Analysis

Thursday, May 22nd: Post-conference Workshop

Bayesian Methods for the Social and Behavioral Sciences

David Kaplan

University of Wisconsin-Madison

Bayesian statistics has long been overlooked in the quantitative methods training for social and behavioral scientists. Typically, the only introduction a student might have to Bayesian ideas is a brief overview of Bayes' theorem while studying probability in an introductory statistics class. This is not surprising. First, until recently, it was not feasible to conduct statistical modeling from a Bayesian perspective because of its complexity and lack of available software. Second, Bayesian statistics represents a powerful alternative to frequentist (classical) statistics, and is therefore controversial. Recently, however, there has been great interest in the application of Bayesian statistical methods, mostly due to the availability of powerful (and free) statistical software tools that make it possible to estimate simple or complex models from a Bayesian perspective. The orientation of this workshop is to introduce practicing social and behavioral scientists to the basic elements of Bayesian statistics and to show through discussion and practice why the Bayesian perspective provides a powerful alternative to the frequentist perspective. It is assumed that workshop attendees will have a background in basic statistical methods up to, and including, regression analysis. Some exposure to multilevel modeling and factor analysis is desirable.

9:00 am – 5:00 pm in Laurel Hall 102

Morning topics include:

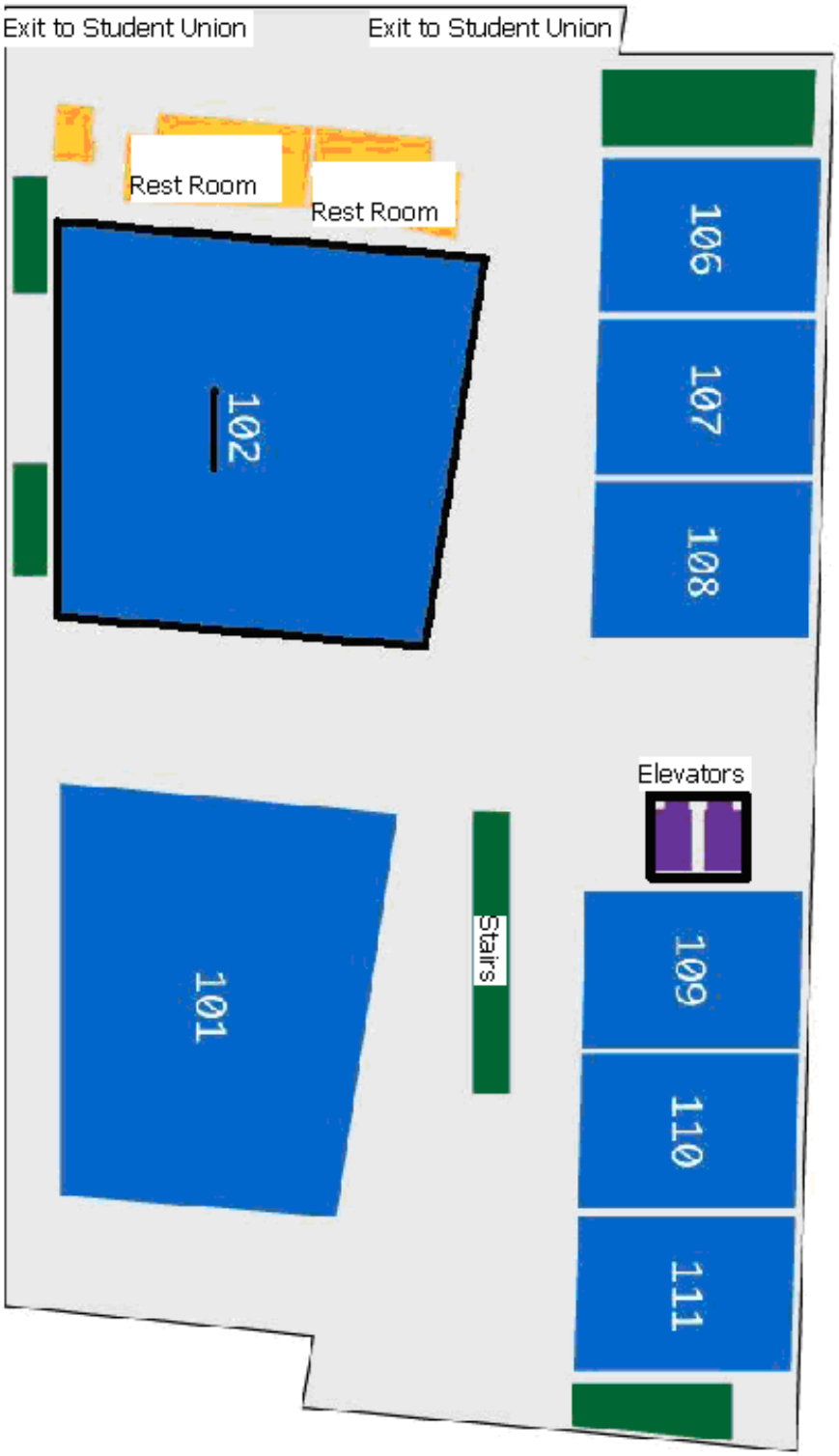
1. Major differences between Bayesian and frequentist paradigms of statistics, with particular focus on how uncertainty is characterized;
2. Bayes' theorem;
3. Bayesian model building and model evaluation;
4. Bayesian computation.

Afternoon topics include:

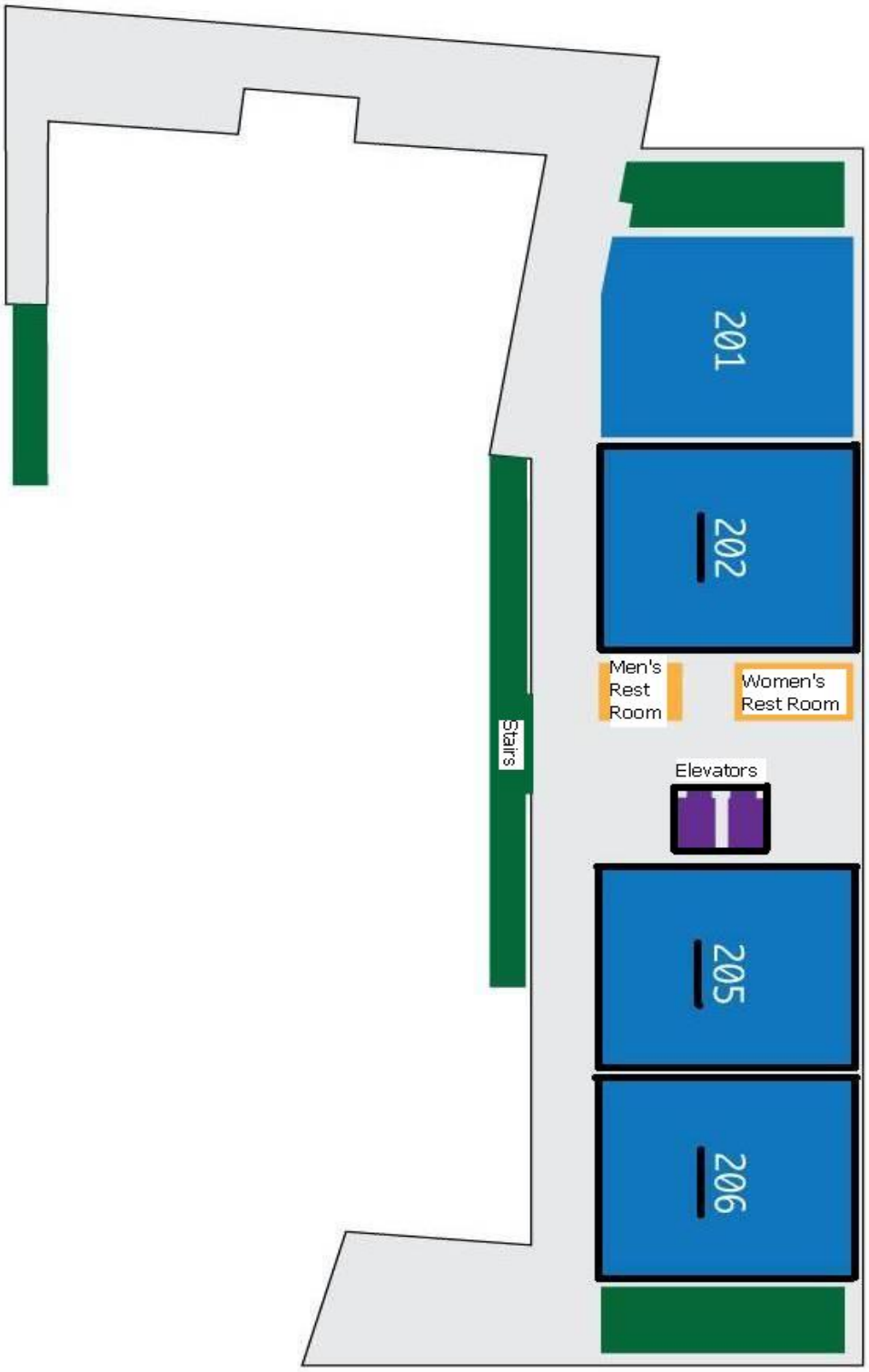
1. Bayesian analyses using R;
2. An example of Bayesian regression;
3. An example of Bayesian factor analysis;
4. An example of Bayesian hierarchical linear modeling;
5. Wrap-up: Relative advantages of the Bayesian perspective

Dr. Kaplan is Professor of Quantitative Methods and Chair of the Department of Educational Psychology at the University of Wisconsin - Madison, and holds an affiliate appointment in the Department of Population Health Sciences. Dr. Kaplan's current research focuses on the development of Bayesian methods applied to a wide range of education research settings. His specific interests include: Bayesian model averaging; objective versus subjective Bayesian modeling; Bayesian posterior predictive causal inference; and Bayesian approaches to problems in large-scale survey methodology. Dr. Kaplan's collaborative research involves applications of advanced quantitative methodologies to substantive and methodological problems in international comparative education. He is most actively involved in the OECD Program for International Student Assessment (PISA) where he has served on its Technical Advisory Group and currently serves as Chair of its Questionnaire Expert Group. He also sits on the NAEP Questionnaire Standing Committee. Dr. Kaplan is a Fellow of the American Psychological Association (Division 5) and was a Jeanne Griffith Fellow at the National Center for Educational Statistics. Dr. Kaplan received his Ph.D. in education from UCLA in 1987.

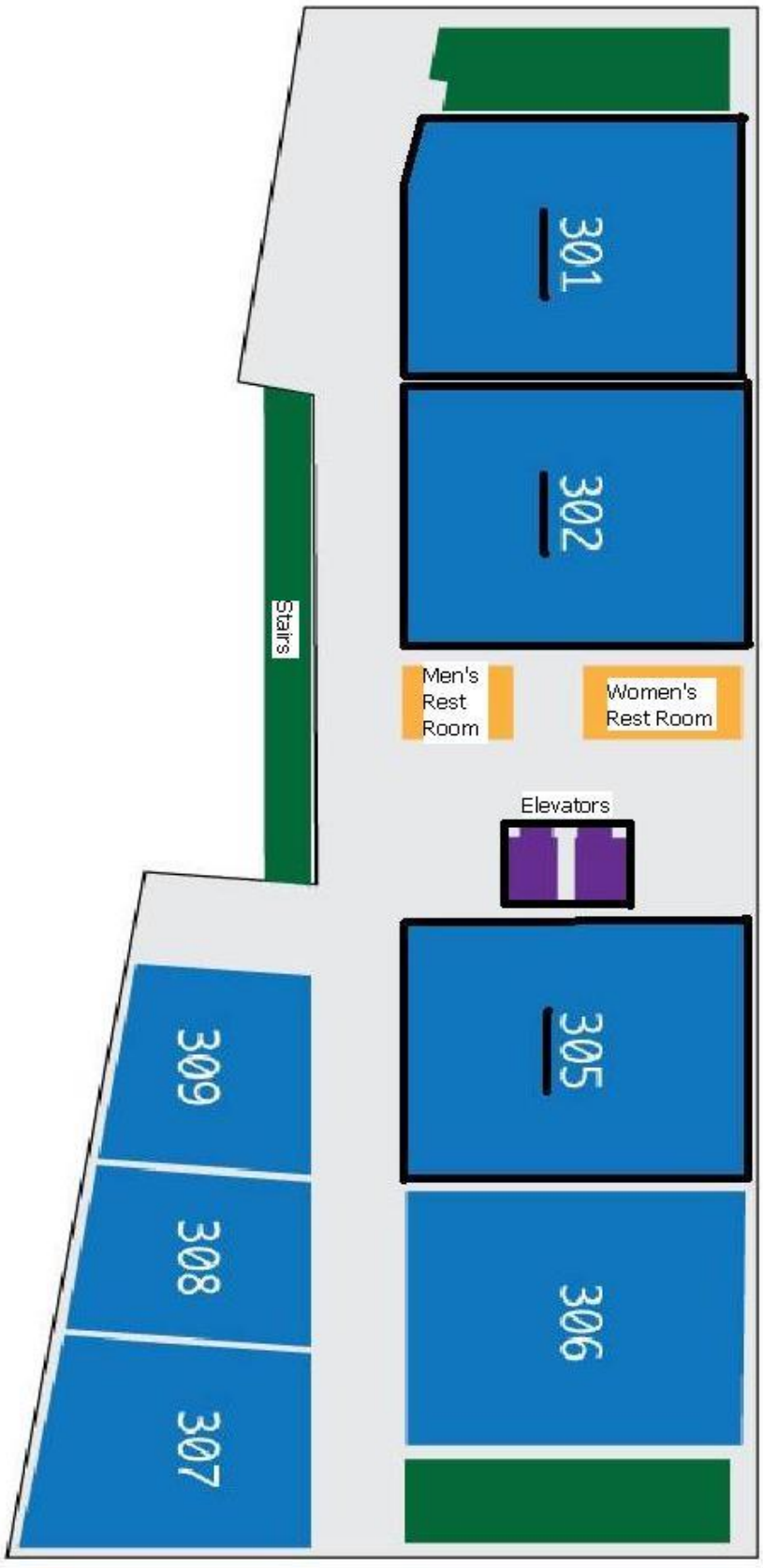
Laurel Hall (Classroom Building) Floor Plans
1st Floor



Laurel Hall (Classroom Building) Floor Plans
2nd Floor



Laurel Hall (Classroom Building) Floor Plans
3rd Floor



Tuesday May 20th - Keynote Address
8:00 am – 9:30 am
Laurel Hall Room 102

Advances in Mixture Modeling

Dr. Bengt Muthén, University of California, Los Angeles

After a brief overview of the many uses of finite mixture modeling, applications of new mixture modeling developments are discussed. One major development goes beyond the conventional mixture of normal distributions to allow mixtures with flexible non-normal distributions. This has interesting applications to cluster analysis, factor analysis, SEM, and growth modeling. The talk focuses on applications of Growth Mixture Modeling for continuous outcomes that are skewed. Examples are drawn from national longitudinal surveys of BMI as well as twin studies. Extensions of this modeling to the joint study of survival and non-ignorable dropout are also discussed.

Tuesday May 20th - Concurrent Session 1
9:45 am – 11:15 am (90 minutes)

Session 1.1: Multi-trait Multi-method Models (Laurel Hall 201)

Neglect the Structure of Multitrait-Multimethod Data at your Peril: Implications for Associations with External Variables.

Laura Castro-Schilo, Keith F. Widaman, & Kevin J. Grimm

In 1959, Campbell and Fiske introduced the use of multitrait–multimethod (MTMM) matrices in psychology, and for the past 4 decades confirmatory factor analysis (CFA) has commonly been used to analyze MTMM data. However, researchers do not always fit CFA models when MTMM data are available; when CFA modeling is used, multiple models are available that have attendant strengths and weaknesses. In this article, we used a Monte Carlo simulation to investigate the drawbacks of either using CFA models that fail to match the data-generating model or completely ignore the MTMM structure of data when the research goal is to uncover associations between trait constructs and external variables. We then used data from the National Institute of Child Health and Human Development Study of Early Child Care and Youth Development to illustrate the substantive implications of fitting models that partially or completely ignore MTMM data structures. Results from analyses of both simulated and empirical data show noticeable biases when the MTMM data structure is partially or completely neglected. Thus, both studies suggest that the collection of MTMM data is crucial to capture accurately the associations of trait constructs with external variables. Researchers should be discouraged from creating averages of observed variables when MTMM data are available. Instead, alternative MTMM CFA models can be fit, and, when substantive implications differ from model to model, these should be reported. Furthermore, we encourage researchers to think about the validity and structure of their data. Our results can aide in the understanding of empirical findings depending on whether researchers foresee positive or differential associations between trait and method factors and external variables.

Uncrossing the Correlated Trait-Correlated Method Model for Multitrait-Multimethod Data.

Laura Castro-Schilo, Kevin J. Grimm, & Keith F. Widaman

In the article that introduced the multitrait-multimethod matrix as an important way to evaluate construct validity of measures, Campbell and Fiske (1959) characterized manifest variables as "trait-method" units. Later, with the emergence of structural equation modeling (SEM), confirmatory factor analysis multitrait-multimethod (CFA MTMM) models were developed. A major benefit of such models was the ability to isolate

trait and method variance in each manifest variable. One CFA MTMM model, the Correlated Trait-Correlated Method (CT-CM; Jöreskog, 1971; Kenny, 1976; Widaman, 1985) model, became very popular, largely because it corresponds directly to Campbell and Fiske's conceptualization of MTMM data. However, this model is prone to high rates of convergence problems and improper solutions. The profuseness of analytic problems encountered with the CT-CM model is not coincidental because much work has shown its empirical underidentification in many situations (Grayson & Marsh, 1994; Kenny & Kashy, 1992). Thus, despite appealing features of the CT-CM model, researchers have opted to introduce new models that alter the mathematically represented trait-method structure of observed variables as a means to improve identification of CFA MTMM models. These alternative models include the Correlated Trait-Correlated Uniqueness (CT-CU) model (Kenny, 1976; Kenny & Kashy, 1992; Marsh, 1989) and the Correlated Trait-Correlated (Method – 1), or CT-C(M-1), model (Eid, 2000). In this paper, we introduce an alternative approach for ensuring theoretical and empirical identification of the CT-CM model under a large variety of conditions. This model is referred to as the Uncrossed Correlated Trait-Correlated Method (UCT-CM) model because it is based on systematically uncrossing the MTMM matrix put forth by Campbell and Fiske. We show results from a Monte Carlo simulation study in which data characteristics lead to an empirically underidentified CT-CM model, but a well-identified UCT-CM model. This improved identification occurs even for a model in which equality constraints are imposed on loadings on each trait factor and loadings on each method factor. Grayson and Marsh (1994) demonstrated that, with such constraints, the CT-CM model is not identified mathematically. However, the UCT-CM handles these constraints with ease. In addition, we present evidence that the CT-CU and CT-C(M-1) models can lead to serious bias if the CT-CM model was the data-generating model, and the model proposed in this presentation – the UCT-CM model –outperforms alternative models with regard to recovering population parameters.

Bayesian Versus Frequentist Estimation of Multitrait-Multimethod Confirmatory Factor Models.

Jonathan L. Helm & Laura Castro-Schilo

Campbell and Fiske's (1959) separation of trait, method, and unique variance across a set of multitrait-multimethod (MTMM) manifest variables directly translates to a confirmatory factor model, and several reports support this approach for partitioning variance (Cole, 1987; Widaman, 1985; Schmitt & Stults, 1986). However, researchers selecting this approach often encounter estimation problems (i.e., failed convergence or solutions with out-of-bounds estimates; Widaman, 1985). Mathematical investigations identified several potential sources of these problems (Kenny & Kashy, 1992; Grayson & Marsh, 1994), one of which indicates a lack of identification when loadings for a given factor are equal. Hence, applied researchers collecting MTMM data face a conundrum when performing analysis. The advent of Bayesian estimation for structural models offers many new opportunities, including the ability to fit models that would fail to converge when estimated within a frequentist framework (Scheines, Hoijtink, & Boomsma, 1999; Asparouhov & Muthén, 2010). Bayesian approaches sample parameter estimates directly from the posterior distribution that arises from a set of prior distributions. Based on the non-identification of the MTMM model when loading are equal (e.g., Kenny & Kashy, 1992; Grayson & Marsh, 1994), the posterior distribution of the latent variable variance-covariance matrix would be flat over a given region, thereby preventing frequentist estimation algorithms to converge on a unique solution. However, Bayesian approaches aggregate across iterative replications taken from the posterior distribution, and should not fail to converge if the flatness is relatively local across the parameter space (Scheines et al., 1999). In this case, Bayesian approaches may estimate models that are otherwise not estimable by frequestist approaches. The current paper examines this flexibility when applied to one confirmatory factor MTMM model. Specifically, we compared convergence rates and parameter bias from the MTMM model fitted by frequentist versus Bayesian estimation procedures. Results from the comparison show that Bayesian estimation yields higher convergence rates than frequentist estimation, but had comparable bias in parameter estimates.

Tuesday May 20th - Concurrent Session 1
9:45 am – 11:15 am (90 minutes)

Session 1.2: Individual-Participant-Data Meta-Analytic Modeling: Measurement Challenges & Network Analysis (Laurel Hall 202)

Modeling Network Individual Data Meta-Analysis.

Tania B. Huedo-Medina, Argie Veroniki, & Dylan Yaworski

It is often the case that clinicians and decision-makers need to explore the efficacy of multiple interventions that are informed by different sets of studies. For this purpose, network meta-analysis (NMA) has been introduced as a new synthesis tool to infer for the relative efficacy or safety of multiple treatments. NMA synthesises simultaneously direct and indirect comparative data across a network of studies. The indirect comparison is a well-established methodology that provides inference on the effectiveness of two interventions via a common treatment comparator. According to this method we can compare treatment B with treatment C via A by quantitatively synthesizing the evidence from A vs. B and A vs. C studies. NMA has been widely used during the last decade providing not only more precise estimates than pairwise meta-analysis and indirect comparison, but also the ability to compare interventions not directly studied in any clinical trial before. Our data set includes evidence on multiple interventions, and the need to compare and rank these treatments suggests the use of NMA. However, the underestimated ESs and the presence of heterogeneity in NMA analysis should be investigated. We therefore explore any possible reasons for its presence by employing NMA multilevel model with study as the cluster effect. Since the study and the individual effect are not separated in a regular meta-analysis and in an IPD-MA not all possible comparisons are analyzed simultaneously, there is a strong underestimated ES. It is very likely in the frequentist setting to find a small and very homogeneous ES. In the Bayesian setting though, we 'correct' for this artifact by adjusting towards the global mean. In the Bayesian NMA RE multilevel model, we assume a fixed coefficient (β) for all treatment comparisons and we assign to it an uninformative prior. The method is more powerful than carrying out several independent pairwise multilevel models. We therefore conclude that Bayesian NMA RE multilevel model is the most appropriate method to meta-analyze these data.

Generalized Linear Modeling to Integrate Measures in Individual-Participant-Data Meta-Analysis.

Nathan Lally, Xiaoran Li, Francisca Galindo-Garre, Maria Dolores Hidalgo, & Tania B. Huedo-Medina

Some statistical harmonization approaches have been already developed to integrate different measures but they have been not compared under a meta-analytic context. Nor has any compromised solution or measurement transformation been developed that could be used when complex constructs on health behavior are used. Therefore we have compared here two particular approaches, t-score transformations versus GLM approach with and without Bayesian estimations, and tested them using a real individual-participant database that integrates studies measures of adherence to HIV/AIDS medication regimes. It is very common to use a t-score standardization to harmonized measures and obtain an effect size (ES). T-score standardization assumes different measures have an underlying normal distribution. This approach is used to create norms and compare each individual study on a common scale. Here a generalized linear models (GLM) with dummy coded indicator variables representing the type of intervention group, i.e., control versus experimental, was used with both classical and Bayesian parameter estimation.

Response Conversion Under Item Response Theory when Conducting Individual-Participant-Data Meta-Analysis.

Xiaoran Li, Nathan Lally, Francisca Galindo-Garre, Maria Dolores Hidalgo, & Tania B. Huedo-Medina

This study compares response conversion and t-score standardization on harmonized individual data. Those two approaches can be applied to the dichotomous, polytomous, and sometimes continuous data that we can have on the behavior change outcomes of HIV prevention interventions. Simulated data and real archival data were used in this study.

Tuesday May 20th - Concurrent Session 1
9:45 am – 11:15 am (90 minutes)

Session 1.3: Practical Issues in Missing Data Analysis (Laurel Hall 205)

You Shall Combine, but How? A Comparison of Methods to Combine Chi-Square Using Multiple Imputation.

Mauricio Garnier-Villarreal, Alexander M. Schoemann, & Todd D. Little

In the social and behavioral sciences, missing data are a problem that almost all researchers must address. This has led to the development of modern methods to handle missing data such as Full Information Maximum Likelihood (FIML) and Multiple Imputation (MI). Most research on modern methods of missing data has investigated how using FIML and/or MI affects parameter estimates and standard errors (Graham, 2003; Zhang & Wang, 2013). Relatively little research on modern missing data methods, however, has focused on the performance of model fit information produced by FIML and MI (Enders, 2010; Savalei, 2008). This study focuses on model fit assessment with FIML and MI in the context of Structural Equation Modeling (SEM). Within SEM, the $\chi 2$ is one of the most relevant fit indices because it gives a statistical test of how much the hypothesized model deviates from the observed variance/covariance matrix. In this paper we investigate the performance of different methods to combine the $\chi 2$ across multiple imputations and compare these methods to $\chi 2$ statistics computed using FIML and complete data. Such an examination of the different methods for combining the $\chi 2$ replicates is imminently relevant for those who wish to implement SEM with multiply imputed data. These users need to be aware of the benefits and caveats of each method, so they can utilize the method that is optimal for the circumstances of their own research. We compare the various methods based on their convergence rate, type I error rate, and power to detect misspecification for each combined $\chi 2$. The results indicate that the method most recommended in literature (MR, Meng & Rubin, 1992) is computationally intensive, and there are circumstances where it does not converge, but a simpler method does converge (LMRR, Li, Meng, Raghunathan & Rubin, 1991). The MR method shows lower type I error rates (always close to the 5%) than the LMRR method, but the LMRR method has higher power to detect misspecification of the hypothesized model (closer to the FIML and complete data cases) than the MR method.

What to do with Incomplete Nominal Variables? A Comparison Of Recommended Techniques For Creating Multiple Imputations of Unordered Factors.

Kyle M. Lang

In this study, five currently recommended approaches for creating multiple imputations of nominal variables are compared via a Monte Carlo simulation. Specifically, I use multiple imputation with chained equations (MICE) to implement three different elementary imputation models: multinomial logistic regression (MNOM), classification and regression trees (CART), and predictive mean matching (PMM). I also consider two joint modeling approaches: the general location model (Little & Schluchter, 1985) and a technique based on rounding normaltheory imputations of the dummy codes that was originally described by Allison (2002) and referred to here as "naive ranking." To judge the performance of these five techniques, each one was included in the imputation phase of a standard multiple imputation analysis. They were then judged by their ability to recover the population-level regression coefficients of the multinomial logistic regression model used to generate the data. The results indicate that MNOM consistently produced unbiased estimates of the population-level regression coefficients, except when sample sizes were very small (i.e., N=100). CART and the general location model also performed reasonably well. Although, for moderate to small sample sizes (i.e., $N < 500$), the performance of CART and the general location model deteriorated with higher proportions of missing data (i.e., $PM > .40$). PMM and naive ranking performed poorly in nearly all conditions, except for those with very low proportions of missing data (i.e., $PM < .10$). Of all the techniques tested, PMM was the worst performer, and its use is not recommended except as a last resort.

Tuesday May 20th - Concurrent Session 1
9:45 am – 11:15 am (90 minutes)

Removing Bias Due to Practice Effects Using a Three-Form Planned Missing Data Design.

Terrence D. Jorgensen, Mijke Rhemtulla, Alexander Schoemann, Brent McPherson, Wei Wu, & Todd D. Little

When studying development of a skill set or perceptual ability, practice accounts for at least some of the change in performance between measurement occasions. Because establishing strong measurement invariance is required before drawing conclusions about changes in the mean of a latent variable (Little 2013), any practice-related changes in the mean of a specific indicator may bias estimates of latent mean-differences. Using a three-form planned missing design to study change in an SEM framework, assigning participants to different planned missing forms over time would decrease the number of variables that participants see on consecutive occasions, thus decreasing the effect of repeated measurement. Using dummy variables to indicate whether participants see items on consecutive occasions can predict practice effects, effectively removing their influence on estimates of latent mean-changes. In the case of the three-form design, this approach would lead to three dummy codes. However, these three dummy variables are linearly dependent, so they cannot be simultaneously included in the same model, but using only two of the dummy variables still leaves practice effects uncontrolled in the third group. Adding a fourth missing data pattern (MDP) to use as a reference group—in which no practice effects manifest because these subjects would see only one item set on each measurement occasion—allows control of bias in all groups. To demonstrate this, we conducted a Monte Carlo simulation study in which we examine three measurement occasions of a six-indicator construct, assigning two indicators to each item set. The results indicate that assigning different forms over time prevents bias to some degree, but simultaneous implementation of both the extra-coefficients and extra-participants methods, in addition to assigning different forms over time, is the only approach that produces unbiased latent means. These results should generalize to other models of mean-change over time, such as latent growth curves or latent difference-score models (see Little, 2013), but only when the construct of interest is measured by at least three indicators (e.g., a curve-of-factors model;Preacher, Wichman, MacCallum, & Briggs, 2008).

A Proof-of-Concept Simulation of the Accelerated Longitudinal Planned Missing Design for Latent Panel Modeling.

Luke McCune

Longitudinal planned missing, represented in the literature by the time-lag model (McArdle & Woodcock, 1997) and the cohort sequential design (Nesselroade & Baltes, 1979), has been thus far restricted to growth modeling and often does not fully utilize the benefits of the planned missingness by estimating a full-longitudinal model. The accelerated longitudinal design may serve as a more flexible and powerful alternative. This study presents a test of the accelerated longitudinal design in a simulated latent panel modeling framework to examine the method's appropriateness for contexts untestable using traditional longitudinal planned missing designs. Three-, four-, and five-cohort models are tested, using a continuum of sample sizes and cohort effect sizes. Results indicate that variance and covariance relationships are replicated well, while characteristics and relationships of the means show low efficiency relative to the full sample case.

Session 1.4: Bayesian and Simulation Techniques (Laurel Hall 206)

Bayesian Outlier Detection in Structural Equation Models

Davood Tofighi

The goal of the current manuscript is to present a Bayesian framework to detect outliers using Mahalanobis-based, case-discrepancy measures for a single-level structural equation model (SEM). We formalize a hypothesis-testing framework and propose two case-discrepancy tests: (a) case- included and (b) case-deleted test quantity. For the two case-discrepancy tests, we use Bayesian estimation to obtain the posterior predictive p value of a case-discrepancy measure that takes into account the model structure, the uncertainty associated with the parameter estimates and the data to derive the reference distribution for

the discrepancy measures. Using these quantities, we can test whether an individual observation is an outlier.

Comparative Study of Two Calibration Methods for Micro-Simulation Models

Stavroula Chrysanthopoulou

The purpose of this paper is to provide a comparative analysis between a Bayesian and an Empirical method for calibrating microsimulation models MSMs). The Bayesian calibration adopts methodology presented in Rutter et al (2009). The second method combines broadly applied practices for the empirical calibration of MSMs. Both methods have been applied to calibrate a streamlined MSM, describing the natural history of lung cancer. The comparative analysis comprises graphical, quantitative and qualitative assessment of the results from a simulation study. Although the empirical method proves more practical (similar results in a shorter time), the Bayesian method seems to perform better when the calibration targets include rare events, while possessing a sounder theoretical background to interpret the calibration outcomes.

Testing Applications of Bayesian Statistics with Informed Priors in Multilevel Analysis of Small Data Sets with Complex Structures: A Monte Carlo Study.

Tyler Hicks, George MacDonald, Jeff Komrey, Eun Sook Kim, Jeanine Romano, & Sandra Archer

This study evaluates the behavior of Bayesian estimators for multilevel models in the context of partially nested data structures, with a primary focus upon estimation of variance-components. Its primary purpose is to investigate the behavior of a Bayesian estimator compared with a classical estimator, given problematic conditions in a multilevel modeling context, such as unbalanced designs and small sample sizes. The study begins with an examination of several proposed methods to set priors in such situations and then summarizes the results of our study. Findings from this study evidences some advantages of using Bayesian estimators with informed priors over traditional approaches in certain multilevel modeling contexts**.**

Session 1.5: Issues in Latent Class Analysis (Laurel Hall 301)

Identifying the Correct Number of Classes in Latent Profile Analysis: The Impact of Sample Size, Profile Distribution, and Model Specification.

Sara K. Johnson

Latent profile analysis (LPA) has become a popular technique among applied researchers who aim to investigate possible population heterogeneity on groups of continuous items, but there has been little empirically-based advice on how to use statistical fit indices to choose the most appropriate number of classes. The purpose of this study was to evaluate the performance of various fit indices in identifying the correct number of LPA profiles depending on several experimental conditions: 5 sample size conditions (ranging from 300 to 3000), 3 profile distribution conditions (equal proportions; one large class; one very small class), and 3 model specifications (conditional independence; equal variances and covariances across profiles; variances and covariances freely estimated within each profile).

Obtaining Meaningful Latent Class Segments with Ratings Data by Adjusting for Level, Scale and Extreme Response Styles.

Jay Magidson

Latent class segments obtained from analysis of ratings data may differ primarily in their rating response styles rather than their preferences. This paper introduces extended latent class (LC) segmentation models, estimated with the syntax module of Latent GOLD 5.0, which classify individuals while adjusting simultaneously for confounds due to level, scale and extreme response styles. A simple illustrative example is provided with ratings data from a taste-testing experiment. Model comparisons suggest that all three types of response styles are present. The scale and extreme response styles turn out to be similar in this application, with the latter dominating the former. We show that adjusting for response styles lead to a

more meaningful segmentation than that obtained from traditional LC models which ignore response style, and new informative graphics. The extended models also provide significantly better prediction (R^2 = .41 vs. .26) and narrower confidence limits for class size.

Session 1.6: Causal Aspects for Modeling (Laurel Hall 302)

Empirical Tests of Directional Dependence using Real Datasets.

Felix Thoemmes, Sarah Moore, & Marina Yamasaki

Dodge & Rousson (2001) introduced the concept of directional dependence as a technique to infer the causal direction of a pair of variables. Recently, case studies by von Eye & DeShon (2012) and simulation studies by Pornprasertmanit and Little (2012) provided some favorable, yet qualified, evidence of the usefulness of this test. We evaluate directional dependence on a database of 79 variable pairs with known causal ordering. We find that only 20% of all cases are correctly classified, a performance slightly worse than chance. Focusing on a subset of variables in which the testable assumptions of directional dependence are fulfilled, only increases correct classification to 23%. We conclude that tests of directional dependence should be used with extreme caution.

How do we Combine Two Treatment Arm Trials with Multiple Arms Trials in Ipd Meta-Analysis? An Illustration with College Drinking Interventions.

David Huh, Eun-Young Mun, & David C. Atkins

In this methodological illustration, we review challenges in individual participant-level data (IPD) meta-analysis with a focus on how to appropriately combine studies with varying numbers of treatments, a prevalent but under-addressed issue in meta-analysis. In the context of college drinking interventions, alcohol outcome data are particularly challenging due to highly-skewed distributions with many zeroes, a characteristic ignored in meta-analysis using summary statistics. We present a Bayesian multilevel modeling approach for combining two treatment arm and multiple arm trials in a distribution-appropriate IPD analysis. Illustrative data from Project INTEGRATE, an IPD study of brief motivational interventions designed to reduce excessive alcohol use and related harm among college students. The innovative analytical approach we present provides a practical method for estimating study-specific and overall treatment effects without the need to collapse intervention conditions within multi-intervention studies while accommodating non-normal distributions and other common characteristics of clinical trial data.

Adjusting for Selection Bias in Assessing the Relationship Between Sibship Size and Cognitive Performance.

Gebrenegus Ghilagaber & Linda Wänström

The aim of this study is to demonstrate how selection bias in studies of sibship size effects on cognitive performance can be adjusted for. We extend existing knowledge in two aspects: (1) as factors affecting decisions to increase family size may vary across the number and composition of current family size, we propose a sequential probit model (as opposed to binary or ordered models) for the propensity to increase family size; (2) in order to disentangle selection and causality we propose multilevel multiprocess modelling where a continuous model for performance is estimated jointly with a sequential probit model for family size decisions. This allows estimation and adjustment for the correlation between unmeasured heterogeneity affecting both family size decisions and child cognitive performance. The issues are illustrated through analyses of scores on PIAT tests among children of the NLSY79. We find substantial between-family heterogeneity in the propensity to increase family size. Ignoring such selection led to underestimation of the negative effects of sibship size on cognitive performance but our multiprocess modelling could mitigate the biasing effects of selection.

Tuesday May 20th - Concurrent Session 2
11:30 am – 12:30 pm (60 minutes)

Session 2.1 (Laurel Hall 201)

Discrete and Continuous Time Models from the Perspective of Structural Equation Modeling.

Pascal R. Deboeck & Aaron M. Boulton

While they have an extensive history, and are in some cases more theoretically justifiable, the adoption of continuous time models in the social sciences has been limited. This may be due to unfamiliar concepts and difficulty in the application of continuous time models. This presentation will examine continuous time modeling, and specifically stochastic differential equation models, from the familiar perspective of structural equation modeling (SEM) diagrams. Differences between discrete and continuous time models will be introduced using SEM diagrams. Subsequently a new approach to fitting continuous time models using SEM software will be introduced; this method is a variation of Latent Difference Scores (McArdle 2009; McArdle & Hamagami 2001) and the Oversampling Approach (Singer, 2012). Results from two simulations will be presented to compare this SEM-based approached with an analytic solution for fitting a continuous time model (the Exact Discrete Model).

Session 2.2 (Laurel Hall 202)

Empirical SEM Confidence Intervals by Bootstrapping.

Craig Michael Krebsbach

Confidence intervals (CIs) in research have become commonplace and recommended, however the field of structural equation modeling (SEM) has not followed suit to date. Introduced is a novel and relatively straight-forward empirically based bootstrap procedure for CIs in the SEM framework. The Bollen-Stine bootstrap in the lavaan package of R is the main focus of this introduction to SEM CIs for any fit index. Emphasis on CI necessity is also included.

Session 2.3 (Laurel Hall 205)

Testing the Homogeneity of Variance Assumption: An Investigation of the Performance of Ten Different Approaches in One-Factor ANOVA Models.

Harold Holmes, Patricia Rodriguez de Gil, Issac Li, Aarti Bellara, Yi-Hsin Chen, Tyler Hicks, & Eun Sook Kim

Variance homogeneity is a critical assumption that should be met when conducting an ANOVA, as violations of this assumption may lead to perturbations in Type I error rates. However, previous empirical research suggests minimal consensus among studies as to which test is appropriate for a particular analysis. Using simulation methods, this study investigated the performance of ten different approaches for testing the homogeneity of variance assumption in one-way ANOVA models in terms of their Type I error rates and statistical Power. In addition, this study describes the rationale associated with examining the variance assumption in ANOVA and whether the results of each test investigated could inform decisions regarding the selection of a valid test for mean differences.

Session 2.4 (Laurel Hall 206)

Bayesian Model Averaging for Propensity Score Analysis.

David Kaplan & Jianshen Chen

This paper considers Bayesian model averaging as a means of addressing uncertainty in the selection of variables in the propensity score equation. We investigate an approximate Bayesian model averaging approach based on the model-averaged propensity score estimates produced by the R package BMA, but which ignores uncertainty in the propensity score. Therefore, we also provide a fully Bayesian model averaging approach via MCMC to account for uncertainty in both parameters and models. A detailed study of our approach examines the differences in the causal estimate when incorporating non-informative versus

informative priors in the model averaging stage. We examine these approaches under common methods of propensity score implementation. In addition, we evaluate the impact of changing the size of Occam's window used to narrow down the range of possible models. Two comprehensive simulation studies and one case study are conducted. Overall, results of show that both Bayesian model averaging propensity score approaches recover the treatment effect estimates well and generally provide larger uncertainty estimates, as expected. Covariate balance checks for the case study show that both Bayesian model averaging approaches offer good balance. The fully Bayesian model averaging approach also provides posterior probability intervals of the balance indices.

Session 2.5: Advanced Longitudinal Modeling Techniques (Laurel Hall 301)

Using a Random-Effects Tobit Model to Analyze Longitudinal Data with a Large Proportion of Zeroes.

Haiyi Xie, Gary R.Bond, Robert.E. Drake, & Gregory J. McGugo

Data with a high percentage of zero values frequently arise in health care research. For zero-inflated continuous (or semi-continuous) data, the two-part model has been proposed and is widely used. An alternative approach, the Tobit model, though popular among economists, is less known and infrequently used in health care research. The Tobit model is designed to analyze censored outcome variable (i.e., outcomes with floor or ceiling effects). In this paper, we illustrate the extended Tobit model for longitudinal data - random-effects Tobit model by analyzing the impact of employment status (steady competitive work, steady noncompetitive work, minimal work, and no work) on length of hospital stays for participants with severe mental illness from a randomized two-year follow-up supported employment study. Employment status and time were included as predictors, and pre-randomization working status and several clinical and demographics variables were included as covariates. The results showed that clients with employment in any kind had fewer days of hospitalization, relative to those who were unemployed; the difference between steady competitive workers and non-workers was stronger than that between non-competitive workers and the unemployed.

Investigating Differential Changes Using Mixture Latent Change Scores (MLCS) Modeling.

Emil Coman, Judith Fifield, Jack J. McArdle, & Monique Davis-Smith

We introduce the Mixture Latent Change Scores (MLCS) models of multi-wave (longitudinal) data for investigating differential changes in unobserved (latent) groups of patients (or cases in general). We present a description of the LCS tool, its utility in replicating known models of change, like RANOVA and LGM, and the potential advantages over the mixture LGM approach to infer differential complex trajectories of changes by unknown latent classes. We present an example using a four wave data from an intervention meant to prevent diabetes in prediabetic patients, using CDC's DPP (Diabetes Prevention Program). The MLCS model seems able to extract more meaningful latent classes than the GMM model, and in our example it extracted classes that appeared to be more balanced, and had some nuanced trajectories that are meaningful, like initial increase, followed by a plateau, then decrease, and initial decrease, plateau, then increase. Limitations and suggestions for extensions are finally provided.

Lunch, 12:30 pm –1:45 pm
In the Student Union Ballroom

Tuesday May 20th - Concurrent Session 3
1:45 pm – 3:15 pm (90 minutes)

Session 3.1: Explanatory Item Response Theory Models (Laurel Hall 201)

The Effects of Reader's Characteristics, Text Genre, and Comprehension Processes on Reading Comprehension.

Paulina Kulesz, David Francis, Marcia Barnes, Jack Fletcher, Amy Barth, & Mary York.

When the psychological processes underlying test performance are understood, psychometric theory dictates that item difficulty can be explained through these processes. The goal of the project is to improve understanding of reading comprehension and the role of reader's characteristics, text genre, and comprehension processes in understanding the text through the application of explanatory item response models. These models are advantageous when compared to standard statistical models because they allow for estimation of individual differences, explain probability of correct responses utilizing external variables, and jointly model the probability of correct responses as a function of person and item characteristics. Joint modeling provides a more accurate estimation of individual differences in a latent trait as it estimates probability of correct responses using person and item side of data.

Reader and Text Characteristic's Contribution to Inferential Processing n Adequate and Struggling Comprehenders.

Yusra Ahmed, David Francis, Marcia Barnes, Jack Fletcher, Amy Barth, & Mary York

Making inferences is potentially a factor for successful reading comprehension as research shows a high correlation between the two skills. Inference making ability requires readers to connect different sentences and ideas within the text in order to build a fully integrated and coherent representation of the text. Inferred information thus is based on the text, but not stated explicitly, requiring readers to interpret the text through existing knowledge. However, background knowledge alone is not sufficient to make correct inferences. Other factors that make contributions to inference making are vocabulary, reading strategies and word reading skills. In the present study, student characteristics were modeled in addition to the effect of two text-based features on inferencing ability, textual distance (whether information to be integrated is near versus far) and concept consistency (whether a sentence is a consistent versus inconsistent continuation in a given metal model). Additionally, an explanatory IRT model was contrasted with repeated measures ANOVA methods which use percent correct on inferencing measures.

A Diagnostic Cognitive Model (DCM) of Algebraic Processing.

Justin Neil Young, Tammy D. Tolar, David J. Francis, & Jeffrey J. Morgan

Examination of theories and limited research evidence related to the development of algebra suggests that algebra knowledge may be a multidimensional construct in that it is possible for students to develop some types of algebraic knowledge independently of others. Possible dimensions include procedural versus conceptual knowledge; verbal versus symbolic versus graphical representations; and components of mathematical objects (e.g., intercept versus slope) versus the objects themselves (e.g., linear functions). The purpose of this study was to use diagnostic cognitive modeling (DCM) to determine if there were classes of students who demonstrated mastery of some dimensions independent of the others. The Linear Function (LF) Diagnostic Assessment was developed from an explanatory IRT framework to evaluate the dimensions described above. It was administered to 300 first year algebra students in middle and high schools. Data from the state algebra end-of-course exam was collected for external validation of the experimental assessment. Systematic evaluation of algebra topics is needed to identify cognitive components of algebraic knowledge and skills. This evaluation should include a variety of research and statistical methods in order to 1) advance theory about the development of algebraic thinking, 2) develop diagnostic tools for identifying students' strengths and weaknesses related to algebra, and 3) produce and/or refine hypotheses for evaluation in subsequent studies. This study provides preliminary evidence that DCM is one measurement method that can be used to address each of these three research goals. DCM based measures may be incorporated into different methodological approaches (e.g., experimental and growth studies) as part of a systematic evaluation of the development of algebra knowledge and skills.

Tuesday May 20th - Concurrent Session 3
1:45 pm – 3:15 pm (90 minutes)

Session 3.2: Exploring the lavaan Ecosystem: Packages to Extend the Capability of lavaan (Laurel Hall 202)

***semTools*: Useful Tools for SEM.**

Alexander M Schoemann, Sunthud Pornprasertmanit, & Patrick J. Miller

semTools is an R package that provides useful add-on functions not provided in main SEM packages (e.g. lavaan, OpenMx, sem). semTools includes functions assisting with model evaluation (e.g. providing additional fit indices for a lavaan model), measurement invariance, power analysis, missing data analysis (e.g. a function to multiple impute data, run a SEM using lavaan and combine the results using Rubin's rules), estimate latent interactions with lavaan, and many other functions (e.g. a script translator from LISREL to lavaan). Additionally semTools is designed to be a user supported packages. Users of R and SEM are encouraged to suggest and create new functions relevant to using SEM in R. Currently, the package includes over 15 different contributors and we strongly encourage others to contribute to the package.

Getting the Most out of your Family of Data with the R-package fSRM.

Lara Stas, Felix Schönbrodt, & Tom Loeys

Family research aims to explore family dynamics but is often limited to the examination of unidirectional processes (e.g. parenting, child effects). As the behavior of one person has consequences that go beyond that one individual, the family functioning should be investigated in its full complexity. The Social Relations Model (SRM; Kenny & La Voie, 1984) is a conceptual and analytical model which can disentangle family dynamics at three different levels: the individual level (actor and partner effect), the dyadic level (relationship effects) and the family level (family effect). Nonetheless, its statistical complexity may be a hurdle for family researchers. The user-friendly R-package fSRM almost automatically performs those rather complex SRM analyses. When a round robin design is used (i.e. every family member rates every other member on the same items), the etiology of the obtained dyadic scores can be unraveled using the SRM. In particular, the estimation of the SRM parameters can be based on a confirmatory factor analysis. Therefor fSRM builds on lavaan (Rosseel, 2012).With fSRM, one simple line of R-code suffices to perform the required analysis. The fSRM-output provides easy-to-interpret summary tables of SRM variances, variance decompositions, individual and dyadic reciprocities. SRM means, which may be very informative - but infrequently reported - are straightforwardly obtained and can easily be compared between roles. Moreover, the package is suitable for both single and multigroup studies. Additional options (e.g. intragenerational similarities) are discussed.

***semPlot*: Unified Visualizations of Structural Equation Models.**

Sacha Epskamp

SEM has a long history of representing models graphically as path diagrams, yet no dedicated software for drawing these path diagrams currently exists. I present the freely available semPlot package for R (Epskamp, 2013), which fills the gap between advanced, but timeconsuming, graphical software and the limited graphics produced automatically by SEM software. In addition, semPlot offers more functionality than drawing path diagrams: it can act as a common ground for importing SEM results into R. Any result useable as input to semPlot can be also represented in any of the three popular SEM frameworks, as well as translated to input syntax for the R packages sem and lavaan. Special considerations are made in the package for the automatic placement of variables, using three novel algorithms that extend earlier work of Boker, J. McArdle, and Neale (2002).

***lavaan*.survey.**

Daniel Oberski

Introduces the R package lavaan.survey, a user-friendly interface to design-based complex survey analysis of SEM. By leveraging existing code in the lavaan and survey packages, the lavaan.survey package allows for SEM analyses of stratified, clustered, and weighted data, as well as multiply imputed complex survey data. lavaan.survey provides several features such as SEM with replicate weights, a variety of re- sampling

techniques for complex samples, and finite population corrections, features that should prove useful for SEM practitioners faced with the common situation of a sample that is not i.i.d.

***simsem*: SIMulated Structual Equation Modeling in R.**

Sunthud Pornprasertmanit, Alexander M. Schoemann, & Patrick J. Miller

simsem is an R package that aims to streamline Monte Carlo simulations for SEMs. simsem can generate multivariate normal and non-normal data for any single group SEM. simsem is especially innovative in the simulation and estimation of missing data: MCAR and MAR missingness and attrition can be simulated, as well as several planned missing designs. Model estimation with missing data can be done using both maximum likelihood (FIML) and multiple imputation, including the combination of parameter estimates using Rubin's Rules. The availability of multiple imputation in simulation software is especially valuable, because there is currently no other software that allows Monte Carlo simulations of data using multiple imputation. Maximum likelihood estimation uses the packages lavaan or OpenMx, and multiple imputations are performed with the packages mice or Amelia. simsem includes innovative methods for methodological simulations and power analyses by continuously varying parameters (e.g., sample size), resulting in power analyses for SEM that require fewer resources than traditional power analyses. simsem allows researchers to easily specify model misspecification in the population, resulting in simulations that realistically approximate real world conditions and simsem includes algorithms that control for parsimony error (e.g., users can specify a population RMSEA and simsem will determine the amount of misspecification needed to achieve that RMSEA).

Graphical Structural Equation Modeling with Onyx.

Timo von Oertzen, Andreas Brandmaier, & Siny Tsang

Onyx is a software environment for creating and estimating linear SEMs. A graphical user interface facilitates the intuitive creation of models and a powerful backend performs maximum likelihood estimation of parameters in the model. Models are defined by drawing diagrams, which can be exported to a variety of matrix specifications, model syntaxes and graphical formats. Particularly, Onyx allows users to export diagrams to model definitions for lavaan, OpenMx and Mplus. Interactive code views allow diagrams to be viewed along with the corresponding model definitions; particularly, changes to the diagram are directly reflected as changes in the model syntax. Onyx can serve as a graphical frontend for creating models in lavaan and is particularly helpful to users who are familiar with the concepts of SEM but less so with a specific syntax, e.g., students in a modeling class. Also, matrix specifications in LISREL and RAM are available in the program. For the use in publications, Onyx supports export to a variety of graphical formats, which allow the creation of publication-ready diagrams for print and web, including EPS, PDF, JPEG, PNG and LaTeX. The estimation engine of Onyx is based on a swarm optimization method that in parallel searches the likelihood space with multiple runners, which is particularly useful for models with difficult optimization surfaces. Also, users are alleviated from the need to choose starting values and potential multiple minima of the likelihood space can be discovered. Further features of Onyx comprise multi-group modeling, definition variables, likelihood ratio tests, and data simulation. Onyx aims at complementing existing free software as a graphical modeling tool that provides seamless transition to script-based programs, particularly lavaan and OpenMx.

Session 3.3: New Developments in the Analysis of Incomplete Data (Laurel Hall 205)

A Survey of Missing Data Techniques Employed in Genome-Wide Association Studies.

Gregory J. Matthews & Ofer Harel

Genome-wide association studies (GWAS) have had great success in unlocking some of the genetic secrets of complex diseases. One of the many issues commonly encountered in these types of studies is that of missing data. Many different methods for dealing with the missing data problem in the GWAS context are used. We attempt to survey and catalog the most common methods currently being employed for missing

data in GWAS as well as offer some suggestions for dealing with missing data in this context. Joint work with Ofer Harel.

Clustering Incomplete Data using Normal Mixture Models.

Chantal D. Larose, Ofer Harel, & Dipak Dey

Clustering with Normal mixture models illustrates how data clumps together by using Normal distributions. However, existing methods for such analyses require complete data. One way to address incomplete data is multiple imputation; however, it is difficult to combine multiple imputation and cluster analysis. In this paper, we develop a new methodology for clustering incomplete data using Normal mixture models. We illustrate how our new method outperforms existing methodology with a simulation study, then demonstrate the method by clustering incomplete yeast gene expression data. Joint work with Ofer Harel and Dipak Dey.

Comparison of Transformations Used in the Estimation of R^2 and Adjusted R^2 in Incomplete Data Sets using Multiple Imputation.

Valerie Pare & Ofer Harel

The coefficient of determination (also known as R^2) is a common measure in regression analysis that many use to indicate the strength of a statistical model. Often times this measure is also used as a model selection tool. When data is incomplete, complete case analysis is often used in order to estimate R^2, however such practice may lead to biased results. The utility of using multiple imputation for the estimation of R^2 after a Fisher-z transformation has been previously studied. This presentation will explore how log-log and complementary log-log transformations perform in the estimation of R^2 and adjusted R^2. The method will then be applied to HIV data. Joint work with Ofer Harel.

Session 3.4: Bayesian Modeling (Laurel Hall 206)

Bayesian SEM Perspectives (From the Hills of Asymptotia).

Albert Satorra

In the context of longitudinal models, we compare two alternative approaches for SEM analysis: (classical (frequentist) versus Bayesian. We confront the actual reports arising from both mehtods of analysis on the same model and data set for different sample sizes. The possibly more general SEM arising from the Bayesian perspective is discussed as well as the robustness to distributional assumptions. The role of Bayesian methods as a computational engine for complex SEM modeling is explored, The relation between empirical Bayes, hierarchical models, and classical Bayesian methods is discussed in the context of SEM under different assumptions on the status (fixed versus random) of parameters. Model specification and model modification tools under the classical and the Bayesian approaches are also discussed. We illustrate this comparison using simulated data and also empirical data on firms' profitability. In the comparison, the (frequentist) asymptotic theory SEM is taken as the preliminary benchmark method.

Bayesian Estimation for Structural Equation Models with Sample Weights.

Bengt Muthen

Bayesian Analysis of Multiple Indicator Growth Modeling using Random Measurement Parameters Varying Across Time and Person.

Tihomir Asparouhov

Bayeisan estimation is seeing increasing use in social science applications, but a hindrance to its use has been the lack of methods for handling complex survey data features such as sampling weights. We consider two different approaches to Bayesian estimation in the presence of sampling weights. The first approach is based on frequency weights estimation with a generic standard error adjustment. The second approach is a model based approach where the weights are included in the model. The two approaches are compared in simulation studies for factor analysis models with continuous and categorical measurement data.

Tuesday May 20th - Concurrent Session 3
1:45 pm – 3:15 pm (90 minutes)

Session 3.5: Modeling Treatment and Causal Effects (Laurel Hall 301)

Treatment Effects in Randomized Controlled Trials with Noncompliance and Missing Data.

Shu Xu, Michael F. Lorber, Amy M. S. Slep, Richard E. Heyman, & Danielle M. Mitnick

Randomized Controlled trials (RCT) are the golden standard for assessing causal intervention effects. It has been widely recommended that a complier average causal effect (CACE) should be one of the primary outcomes in a RCT, given the participants under various treatment conditions may not be followed exactly the same way in a study. However, the estimation of a CACE effect may not be straightforward given the presence of both noncompliance and missing data. This study aims to illustrate a method of assessing the intervention causal effect of a RCT given the presence of noncompliance and missing data by taking a structural equation modeling approach (Jo & Mùthen, 2001; Jo, Ginexi & Ialongo, 2010). We start with a brief introduction to an RCT, the potential outcomes framework, and missing data mechanisms. We then demonstrate the CACE estimation with a single continuous outcome. An empirical study is provided as an illustration.

Estimation of Heterogeneity of Treatment Effect Using EM Algorithm.

Evgeniya Reshetnyak, Ying Liu, Barry Rosenfeld, & William S. Breitbart

Heterogeneity of treatment effect (HTE) in randomized clinical trials is referred to the problem of high variability in treatment responses across patients. Average treatment effect within population of patients is often not a reliable measure of efficacy and safety of a tested treatment even in a well-controlled randomized study, since it often masks negative outcomes for subpopulations of patients. Despite the existence of a wide array of statistical methods for HTE assessment, all of them have limitations. We are proposing the statistical model of Heterogeneity of Treatment Effect's estimation using EM algorithm, the approach that so far has been overlooked in the clinical research literature.

Probing Causal Mechanisms and Strengthening Causal Inference by Means of Mixture Mediation Modeling.

Emil Coman, Judith Fifield, Suzanne Suggs, Deborah Dauser-Forrest, & Martin-Peele Melanie

We introduce mixture mediation, a modeling technique aimed at examining whether there are latent classes of cases (people) for which an intervention indirect effect may have differentially augmented or hampered the direct effect on the desired outcome. The models illustrated here use a simple mediation setup intervention –Mediator-Outcome, with covariates, which is estimated for 2 (or more) latent classes. We review options for specification of expected differences in parameters across classes and we present the results of such models from a community-based intervention conducted in Black churches (SisterTalk) aimed to achieve weight loss among women. Results show that mixture mediation can uncover classes of intervention participants for which the indirect effect (along with the direct and total ones) differ, which points to differential mechanisms leading to intended effects in the treatment group. We review some implications like the issue of choosing an appropriate comparison group for specific un-mixed latent classes, and hence that of the causal inference about the 'true impact of an intervention' on participants.

Session 3.6: Modeling and Mediation (Laurel Hall 302)

Parametric Sensitivity Analysis Thresholds for Mediation Effects.

Ben Kelcey, Kenneth Frank, & Michael Seltzer

We develop a new class of sensitivity analysis methods for mediation effects estimated through linear structural equation models. We expand the impact threshold framework developed by Frank (2000) to derive the impact threshold of a confounding variable in mediation analyses. In contrast to other sensitivity analyses, the proposed methods consider each of the simultaneous relationships an unobserved variable must have and quantify specific sensitivity thresholds--the specific minimum simultaneous relationships necessary for an unobserved confounding variable to invalidate an inference concerning a mediation effect.

Tuesday May 20th - Concurrent Session 3
1:45 pm – 3:15 pm (90 minutes)

The methods developed shift the debate from whether or not an inference is sensitive, to the specific magnitude of relationships the unobserved variable needs with the outcome, treatment, and mediator in order to invalidate an inference

The Markov Equivalence Class of the Mediation Model.

Felix Thoemmes, Wang Liao, & Marina Yamasaki

In mediation analysis, it is common practice to present a particular assumed mediation model and provide estimates of indirect effects to support the hypothesis that mediation is in fact present. Less thought is devoted to ruling out alternative models that would fit the data equally well (but would have different causal implications). One reason for this absence of consideration of alternative models might be that researchers are unfamiliar with ways on how to enumerate such alternative models. We use d-separation to enumerate all models in the Markov equivalence class (class of models that fit the data equally well) and use path-tracing rules to derive the expected value of the indirect effect for each of the models. We further show which causal assumptions are necessary to reduce the number of models in the equivalence class.

Mediation Analysis in AB/BA Crossover Studies.

Haeike Josephy, Tom Loeys, & Stijn Vansteelandt

Crossover trials are widely used in psychological and medical research to assess the effect of reversible exposures. In such designs, each subject is randomly allocated to a sequence of conditions, enabling evaluation of treatment differences within each individual. When there are but two possible exposures -each assessed during one time period-, the crossover study is referred to as an AB/BA design. The goal of this presentation is to discuss mediation analysis in such studies. We do so by considering within-subject mediation from a counterfactual-based perspective and by deriving expressions for the (in)direct effect. Employing simulation studies, the performance of several existing methods will be assessed and compared to that of a novel one, we propose. We show that the new method yields unbiased and efficient estimators for the (in)direct effect, under a minimalistic set of 'no unmeasured confounding'-assumptions. Finally, we illustrate the different techniques with data from a neurobehavioral study.

Tuesday May 20th - Keynote Address
3:30 pm – 5:00 pm
Laurel Hall Room 102

Estimation of Causal Effect of Exposures that Change Over Time: Methods and Case Studies

Dr. James Robins, Harvard University

The principal focus of Dr. Robins' research has been the development of analytic methods appropriate for drawing causal inferences from complex observational and randomized studies with time-varying exposures or treatments. The new methods are to a large extent based on the estimation of the parameters of a new class of causal models – the structural nested models – using a new class of estimators – the G estimators. The usual approach to the estimation of the effect of a time-varying treatment or exposure on time to disease is to model the hazard incidence of failure at time t as a function of past treatment history using a time-dependent Cox proportional hazards model. Dr. Robins has shown the usual approach may be biased

whether or not further adjusts for past confounder history in the analysis when:
(A1) there exists a time-dependent risk factor for or predictor of the event of interest that also predicts subsequent treatment, and (A2) past treatment history predicts subsequent risk factor level.

Conditions (A1) and (A2) will be true whenever there are time-dependent covariates that are simultaneously confounders and intermediate variables.

In contrast to previously proposed methods, Dr. Robins' methods can 1) be used to estimate the effect of a treatment (e.g., prophylaxis for PCP) or exposure on a disease outcome in the presence of time-varying covariates (e.g., number of episodes of PCP) that are simultaneously confounders and intermediate variables on the causal pathway from exposure disease; 2) allow an analyst to adjust appropriately for the effects of concurrent non-randomized treatments or non-random non-compliance in a randomized clinical trial. For example, in the AIDS Clinical Trial group (ACTG) trial 002 of the effects of high-dose versus low-dose AZT on the survival of AIDS patients, patients in the low-dose arm had improved survival, but they also took more aerosolized pentamidine (a non-randomized concurrent treatment); 3) allow an analyst to adequately incorporate information on the surrogate markers (e.g., CD4 count) in order to stop at the earliest possible moment, randomized trials to the effect of the treatment (e.g., AZT) on survival.

Dr. Robins teaches at the Harvard School of Public Health.

Tuesday May 20th - Poster Session and Reception 5:00 pm – 7:00 pm Student Union Ballroom

1. **The use of Multilevel Random Coefficient Models for the Analysis of Mediational Change Over Time.**
   Erin H. Arruda & Jennifer L. Krull

Mediation analysis permits researchers to address questions concerning mechanistic processes, positing the relationship between initial (x) and outcome (y) variables is at least partly due to an intervening variable, the mediator (m). Longitudinal mediation models explicitly model temporal precedence among x, m, and y variables, yet do not address a potentially novel research question – Does the mediated effect change over time? This study develops a multilevel modeling method to evaluate mediational change over time (the M-COT model), combining growth modeling and multilevel mediation approaches into a single comprehensive analysis producing an estimate of the rate of change in meditational processes not directly available from a series of single-level analyses. Using simulation techniques, conditions for application were evaluated (e.g. sample size, attrition and form of the trend). Results indicate the multilevel technique performs well across conditions. Estimates of timepoint-specific mediated effects have smaller standard errors, thus more power, than single-level estimates.

2. **A Latent Class Analysis of Transgender Identity Terms.**
   Andrew Bauerband, Katherine Masyn, Sari Reisner, & Kerith Conron

The current research utilized latent class analysis to characterize subgroups within the transgender population. We used the National Transgender Discrimination Survey Data to run LCA on 13 transgender identity terms. The data was separated by sex assigned at birth and randomly split into two subsamples for each sex. LCA was conducted separately for males assigned at birth and females assigned at birth. The second samples served to cross validate the models. LCA yielded an optimal 3 Class Model for both assigned males and females. The classes that emerged were 1) transgender individuals with some non-conforming identification, 2) individuals with primarily a transsexual/transgender identity, and 3) non-transsexual, gender non-conforming individuals. Distinctions between class models and implication for future transgender research will be discussed.

3. **Prospective Links Between Adolescent Perseverance and Adult Health: An Integrative Data Analysis Approach.**
   Lizbeth Benson, Margaret L. Kern, & Laurence Steinberg

For this study, we identified six archival longitudinal datasets that included measures of adolescent perseverance, and adult physical health items at two time points (around age 30 and 50). We applied integrative data analysis techniques to test adolescent perseverance as a predictor of health behaviors in mid-adulthood and self-rated health, cardiovascular function, and energy/ vitality in adulthood. Items in each sample were rated for construct similarity; similar items were selected and included in the analyses. We tested parallel models across the six samples, and the resulting effect sizes were meta-analytically combined to obtain overall effects for each outcome. Next, studies were harmonized by dichotomizing or rescaling each item or scale (e.g., self-rated health was dichotomized to good versus poor health) and then directly combined. Findings supported the protective effect of perseverance, but also highlighted complexities based on samples and measures.

4. **An Exploration of US Students' Math Achievement Growth using Latent Growth Modeling.**
   Menglin Xu

This study explored the factors influencing trajectories of math achievement among US Grade 8-12 students. The data was from the National Education Longitudinal Study (NELS) student questionnaire for the years 1990, 1992, and 1999. SAS 9.3 and Mplus 7.0 were used to analyse data. Conditional latent growth modeling was adopted. Results showed that male students had higher growth rate than females; reading achivement and subjective overall academic performance in Grade 8 significantly and negatively affect the math skills growth rate.

5. **A Latent Growth Curve (LGC) Analysis to Dissociate Components of Response-Time (RT) Variance.**
   Natalie Borter, Stefan Troche, Yulia Dodonova, & Thomas Rammsayer

RT on more demanding cognitive tasks predicts psychometric intelligence (g) better than RT on less demanding ones. For any given cognitive task irrespective of task demands the correlation between a person's worst performance (the slowest RTs) and g is larger than between this person's best performance (the fastest RTs) and g. The magnitude of the difference between these correlations was suggested to be greater in more demanding tasks. This finding was previously based on analyses of worst performance without controlling for best performance as well as on task demands without controlling for demand-independent variance. Therefore, with a two level LGC, we dissociated worst and best performance at a first level and variance due to task demands from demand-independent variance at a second level and correlated those components with g. Our results challenged the finding that the effect known as the worst performance rule is greater in more demanding tasks.

6. **Using Latent Transition Analysis to Examine Intervention Effectiveness Among Subgroups.**
   Leslie Brick, Steven F. Babbin, & Wayne F. Velicer

The current study examines two longitudinal transition models using Latent Transition Analysis (LTA) in order to better understand findings from a computer delivered, tailored intervention targeting multiple risk behaviors in adolescents. Sixth graders from 20 schools were randomly assigned to either a smoking and alcohol substance use prevention program (SP) or an energy balance (EB) program. Students in the EB intervention effectively initiated and maintained energy balance behaviors. The SP intervention, however, was not as effective. Separate two-group models were fit to examine status transition differences for smoking and alcohol acquisition between intervention groups and across time. Among the findings, it was discovered that the Most Protected subgroup was consistently the most stable for both intervention conditions and behaviors and had higher probabilities in the SP condition (a positive treatment effect). LTA provides an essential and unique perspective in the understanding of treatment effectiveness by highlighting subgroup transitions across time.

7. **Trajectories of Heavy Episodic Drinking During Emerging Adulthood: An Application of Generalized Linear Mixed Model.**
   Grace Chan, Victor Hesselbrock, & Michie Hesselbrock

Objective: To model trajectory of heavy episodic drinking from adolescence to young adulthood.
Methods: The COGA study has collected extensive data from 1211 (44% male) subjects. Participants were assessed four times in approximately two-year intervals between the ages of 12 and 30. Generalized linear

mixed model (GLMM) was used to model the trajectory of heavy episodic drinking, which was defined as the maximum number of drinks consumed in any 24-hour period during the past six months. Results: The trajectory of recent heavy episodic drinking varies as a quadratic function with age (peak at 22.5 years old) after adjusting for gender and family history, which was defined as having at least one parent with alcohol dependent. Conclusions: The GLMM is an appropriate statistical technique to model trajectory of heavy episodic drinking during emerging adulthood. It is flexible enough to include covariates and take individual variation into consideration.

8. Causal Modeling Approaches to Answer Research Questions About Health Disparities.
Emil Coman, Victor Villagra, Brenda Shipley, & Judith Fifield

We present a series of causal models meant to answer Health Disparities (HD) research questions. HDs are considered avoidable differences in health outcomes between groups like ethnicities. Understanding HD and their causes requires flexible modeling aimed at separating out causal factors and mechanism responsible for observable HDs. To answer HD questions like "What is the size of HD between specific ethnic groups", "What are the key drivers of HD in a specific health outcome", "How much of existing HD is due to ethnic differences, or how much can/cannot not be explained by other mechanisms" we present a sequence of simple models starting with: 1. simple mean comparisons, which can be equivalently specified as 2. regressions of the outcome on the grouping variable; 3. adjusted mean difference models; 4. mediation models of HD; 5. mixture models to gauge the "size of the HD effect" and the power of HD causal models. Extensions and suggestions are provided.

9. Evaluating a Test for Negligible Interaction in Factorial Independent Groups Designs.
Alyssa Counsell, Robert Cribbie, & Chantal Ragoonanan

In many designs, researchers aim to find no/negligible interaction as part of a primary hypothesis of no interaction or to justify dropping the interaction term from their model. Typically a lack of interaction is concluded when the interaction term is not statistically significant. Having the research goal (no interaction) aligned with a null hypothesis is problematic both theoretically and statistically. Instead, equivalence testing methods should be employed. The current simulation study evaluated a test of negligible interaction (Cheng & Shao, 2007) and compared it to erroneously using non-rejection of the null hypothesis as indication of no interaction. Statistical properties were compared under normal conditions as well as conditions with non-normality and unequal variance. Results demonstrate that the test of negligible interaction has good statistical properties and, unlike using a non-significant test statistic, provides a theoretically sound methodological alternative for detecting a lack of interaction.

10. A Note on the Estimation of Factor Scores from Multiple Group Item Response Models: Implications for Integrative Data Analysis and Potential Solutions.
Pega Davoudzadeh, Kevin J. Grimm, Keith F. Widaman, Sarah L. Desmarais, Stephen Tueller, & Richard A. Van Dorn

Data integration refers to obtaining multiple data sets, scaling their measurements, and analyzing them as though they represent a single data set. A first step is to scale measurements to common scales, which is typically done with multiple group item response models. Latent variable scores are then estimated and used in subsequent analyses. This approach was found to produce inconsistencies in latent variable estimates for individuals from different studies with the same response pattern. Monte Carlo Simulations were then conducted to evaluate the accuracy of latent variable estimates from this and other approaches that ignore the nesting of participants in studies. Results suggest that ignoring study differences led to slightly more accurate latent variable estimates that were also consistent. This approach is recommended for data integration and scaling of latent variable scores across studies. Implications for longitudinal data integration are discussed.

11. Invariance of Teacher Ratings of Child Behavior Across Preschool.
Mihaela Ene, Elizabeth A. Leighton, & Christine DiStefano

As screening for emotional and behavioral risk has become more prevalent with young children, information concerning screeners' latent structure is necessary to support use of the scoring system used to identify students who may be at-risk. The purpose of this study is to investigate the latent structure of the Behavioral and Emotional Screening System Teacher Rating Scale-Preschool (BESS TRS-P) and determine if the structure holds for all age groups considered for screening with this instrument. A U.S. representative

sample of ratings of more than 1,200 preschoolers was used for both exploratory and confirmatory factor analyses. Results showed support for a bifactor model, suggesting that although useful, only considering an overall score may be insufficient for an accurate, specific classification of preschoolers into behavioral risk categories. Invariance testing by age and latent mean differences will be used to further examine the consistency of BESS TRS-P structure across age groups.

12. Latent Transition Analysis Approach for Evaluating Parenting Prevention Programs: Example of the Intervention for Parents of Children with Asthma.

Ania Filus, Alina Morawska, & Amy Mitchell

Latent Transition Analysis is a method of modeling change over time in categorical variables. It has been popular in social sciences for many years, yet it is scarcely used for evaluation of parenting prevention programs. To illustrate utility of this approach we applied it to the data from a randomized controlled trial [2 (intervention vs. care-as-usual) x 2 (pre-test, post-test)] designed to test the impact of the intervention (Triple P) in improving parental illness management and child behaviour. LCA revealed two classes of parents: (a) poor and (b) efficient illness management groups. 38.1% of parents classified into poor illness management group at Time 1 transitioned into favorable direction at Time 2. Intervention effect was not significant (OR = .80, 95%CI (0.29 – 2.21)). Parents in the efficient illness management group at Time 2 reported significantly less ($p < .05$) child behaviour problems at Time 2 compared to other group of parents.

13. Investigation of Treatment Effects and Measurement Bias for Mathematics Achievement in Children with Mild Disabilities.

Matthew E. Foster, Lee Branum-Martin, Rose A. Sevcik, Robin D. Morris, & Paul T. Cirino

Very little is known about mathematics achievement in children with mild intellectual disabilities (MIDs). Using data from three time points during a school year for 265 elementary aged children with MIDs and 137 with reading disabilities (RDs), we fit confirmatory factor models for the structure of mathematics achievement, examined across group measurement invariance and differential indicator bias, and evaluated students' response to a mathematics treatment. Results suggest that the structure of mathematics achievement was equivalent in children with MIDs and children with RDs, that mathematics achievement measurement was temporally stable within and between groups, and that students in the mathematics intervention demonstrated more growth over the school year and outperformed children in the reading intervention comparison condition. Coupled with the previous findings, those concerned with differential indicator bias suggested that the KeyMath-R (Connolly, 1988) provided a reliable and valid assessment of mathematics skills.

14. The Impact of Home, Work, and Church Environments on Fat Intake Over Time Among Rural Residents.

Regine Haardoerfer, Michelle C. Kegler, Karen Glanz, & Ann Addison

We investigated stability of individual and environmental determinants of fat intake in the home, at work, and at church in rural southwest Georgia. Multilevel analyses investigated the impact of those determinants on fat intake. Home and work environments varied while church environments remained stable. Age, gender, and self-efficacy in healthy eating were individual predictors of fat intake. In the home, presence of more high fat items was significant. In the work environment, having access to healthy foods as well as healthy eating programs had positive impact as did hearing healthy eating messages and availability of healthy foods at church. Understanding stability and variability of dietary fat intake from a social ecologic perspective will aid in identifying targets of change for intervention. Understanding which components of key behavior settings are dynamic and which are relatively stable will help to disentangle the complexity of multi-level determinants of dietary behavior.

15. An Item Response Theory Model for Periodontal Examination Data with Planned Missingness.

Lauren Harrell

In a study of the oral health status of current methamphetamine users, an item response theory model is proposed to analyze periodontal examination data where data are missing by design. Periodontal disease is typically classified on a binary scale based on one or more observations of high attachment loss or pocket depth, but half-mouth examinations have been shown to underestimate the prevalence of periodontal

disease under this definition. It is hypothesized that periodontal disease can be represented as a latent continuous trait, which is measured by attachment loss and bleeding at four sites per tooth. Unidimensional and multidimensional item response models are calibrated to the data and compared, and periodontal disease status estimates are produced from the expected a posteriori. However, further research is necessary to address the residual local dependence of items due to potential spatial associations and to extend this model to data from the general, methamphetamine-free population.

16. A Multilevel Framework for Investigating Patterns of Breast Cancer Outcomes Among Women in Connecticut.

Jessica Hoag

Increased breast cancer mortality risk in minority and disadvantaged population subgroups in Connecticut (CT) is well-established, but more research is warranted to understand why these patterns exist. Utilizing data from the National Cancer Institute-Surveillance Epidemiology and End Results (NCI-SEER) Connecticut Tumor Registry (CTR) and US Census tract-level data, a three-level hierarchical logistic regression model incorporates individual-level factors (race, age, place of birth, insurance status, tumor characteristics); neighborhood-level factors (income, employment, racial composition); and regional factors (racial and socioeconomic residential segregation indices specific to CT), ultimately allowing for systematic analysis of covariates measured at different levels on treatment outcomes, and corrects for the biases in parameter estimates that result from clustering. This study expands on existing literature which has focused primarily on effects of individual characteristics or socioeconomic determinants separately, by emphasizing the application of multilevel considerations into breast cancer disparities research.

17. Clinical Application of Planned Missingness Designs.

Scott Huff, Shayne Anderson, & Rachel Tambling

We assessed how appropriate it is to use planned missingness designs to collect clinical measures as a way to reduce the burden on clients when completing clinic paperwork. Using data from clients, we simulated scores as though clients had completed a randomized subset of the items in our battery of assessments. Results indicate that under a three-form design, clinicians would lose almost no information about specific clients, including measures of current functioning, clinical cut-offs, and client progress. Under a ten-form design, measures were still generally accurate though less precise. In practice, this means that research involving clinical work can still implement planned missingness designs without significant clinical consequences. Implications will be explored.

18. Characterizing the Developmental Course and Turning Points for Alcohol Consumption, Heavy Drinking, and Marijuana use.

Kristina Jackson & Daniel Bauer

Substance use shows normative age-related trends, typically onsetting in late adolescence, manifesting peak prevalences in emerging adulthood, and decreasing thereafter. The literature is inconsistent, however, with regard to the precise timing of peak use. The present study examined course of alcohol use (frequency of consumption, heavy drinking) and marijuana use in a national sample of adolescents age 12-16 followed into young adulthood. Data were drawn from the National Longitudinal Survey of Youth-1997 (N=8,984; 52% female; 52% non-Black/non-Hispanic). Using latent growth modeling, we modeled course of substance use from age 12 to 30 using a linear spline model with an unknown changepoint. For all substances, we observed a steep linear increase throughout adolescence. Peak use was reached earliest for marijuana use and latest for alcohol consumption. Whereas heavy drinking and marijuana use declined following peak use, alcohol consumption plateaued. Findings have implications for processes underlying the timing of normative maturing out.

19. Two-Part Random Effects Model: Application to Direct Behavior Ratings Data.

Janice Kooken, Megan E. Welsh, Faith G. Miller, Gregory A. Fabiano, T. Chris Riley-Tillman, & Sandra M. Chafouleas

Traditional methods of growth modeling which assume normally distributed outcome variables can lead to problems with fit and interpretation when outcome data is semi-continuous. The data utilized in this study are semi-continuous referring to data for which part of the distribution is defined at a single value such as 0 and the remainder is continuous. Growth models measuring initial status, shift in initial status and change over time from a preliminary study used data transformations to adjust for lack of normality. These results

will be compared to two part models based upon Olsen & Schafer (2001) for the zero and continuous parts of the distribution.

20. Using Autoregressive Integrated Moving Average (ARIMA) Modeling to Analyze Daily High School Attendance Rates.

Matthijs Koopmans

Autoregressive Integrated Moving Average (ARIMA) approaches to time series analysis are infrequently used in education. The purpose of this presentation is to illustrate the use this approach using real high school attendance data. The basic concepts and assumptions of ARIMA are discussed, as well as the data diagnostics, residual analysis model selection and interpretation of parameter estimates. The analysis of this particular school reveals correlated errors, a strong weekly cyclical pattern and a heavy skewness in the series due to occasional instances of very low attendance. Most of these trends would not have been captured by traditional models (central tendency, variability, OLS). Implications of the findings and the potential of ARIMA for the education and related disciplines are further discussed.

21. Testing the Latent Mean Difference of Gambling Involvement.

Jaeseok Lee, Chih-Chien Chen, & Choong-Ki Lee

Pathological gambling has been a critical issue that the gambling industry has to confront due to its adverse consequences. The gambling industry wants involved gamblers but not addicted gamblers because problem gambling requires a casino company to spend substantial costs to recover its legitimacy. This study aims to profile involvement pattern across the gambler types (i.e., recreational, moderate, and pathological gamblers) depending on the level of addiction by comparing three psychological involvement facets (i.e., hedonic value, symbolic value, and centrality). The results of an onsite casino survey revealed that (1) the symbolic value and centrality of pathological gamblers are significantly higher than those of moderate gamblers; (2) moderate gamblers have a significantly higher symbolic value and centrality than recreational gamblers. (3) The hedonic value of moderate gamblers is significantly higher than that of recreational gamblers, whereas (4) it is not significantly different between moderate and pathological gamblers. This study contributes to extending understanding of gamblers' involvement patterns with respect to addiction.

22. Assessing Cross-Domain Measurement Invariance for Teacher Efficacy in Science, Literacy, and Math.

Kyung-Sook Lee, Steven J. Pierce, & Laurie Van Egeren

Measurement invariance (MI) confirmatory factor analysis across domains (e.g., science, literacy, and math) was examined in a sample of 66 teachers using Bayesian structural equation modeling (BSEM). The test for exact scalar invariance with Bayesian estimation failed. Therefore, we investigated approximate and partial MI with Muthén and Asparouhov's (2013) two-step Bayesian analysis procedure, where step 1 performs BSEM analysis with approximate invariance priors and step 2 frees non-invariant parameters while holding other parameters exactly invariant. At Step 1, 2 factor loadings were non-invariant between the science and literacy domains, possibly because of a ceiling effect on a literacy item ("I have fun doing literacy and reading activities in the classroom") that reduced its variability. At Step 2, freeing 2 factor loadings across three domains and holding the other 3 factor loadings and all 5 intercepts exactly invariant results in good model fit. This provides evidence of partial invariance.

23. Investigating the Latent Structure of the BESS Preschool Parent Rating Scale.

Elizabeth Leighton, Mihaela Ene, & Christine Distefano

While screening for emotional and behavioral risk has become more prevalent among young children, little is known about the latent structure underlying many of the brief screeners and instruments currently in use. Using exploratory factor analysis (EFA), confirmatory factor analysis (CFA), and exploratory structural equation modeling (ESEM), the current study will investigate the latent structure of the Behavioral and Emotional Screening System (BESS) Preschool Parent Rating Scale. While the use of EFA and CFA is common practice in investigating latent structures, the emerging technique, ESEM, provides advantages beyond both EFA and CFA because it allows for exploring a latent structure while also including information about fit, different rotation methods, and inclusion of covariates (Asparouhov & Muthen, 2009). In addition to identifying the BESS Preschool parent form's latent structure, findings across each modeling technique will be compared and contrasted in the current study.

24. An Extended Multimethod Latent State-Trait Approach to Assess Consistency, Occasion-Specificity, and Method Effects of ADHD Inattention, Hyperactivity, and Academic Impairment.

Kaylee Litson, Christian Geiser, G. Leonard Burns, & Mateu Servera

The present study demonstrates the general usefulness of multimethod LST approaches in studying both states and traits in longitudinal multimethod data in clinical and non-clinical contexts. The purpose of this study was to assess the reliability, consistency, occasion specificity, and rater specificity of Attention Deficit Hyperactivity Disorder (ADHD) symptom ratings by parents and teachers using advanced latent state-trait (LST) modeling techniques. This study represents the first application of a multimethod LST model (Courvoisier et al., 2008) to examining the convergent and discriminant validity of ADHD states and traits. In this application, we also present a slightly extended version of Courvoisier et al.'s model with indicator-specific state residual factors.

25. Analysis of the Factor Structure of the Pediatric Symptoms Checklist (PSC-17) in the Preschool Environment.

Jin Liu & Christine DiStefano

The purpose of this study is twofold: (1) to investigate the underlying structure of the Pediatric Symptoms Checklist (PSC-17) for preschool children rated by teachers in normal education classrooms using modern methods and (2) to compare and contrast results from different methods to uncover latent variables. Exploratory factor analysis and Exploratory Structural Equation Modeling (ESEM) will be used for analysis. ESEM is a relatively new technique that has been used to analyze the internal structure. Participants were 854 preschool children. Results suggested that the PSC-17 has the following three structures: Internalizing problems, Externalizing problems, and Attention problems.

26. Exploring Sources of Bias in Gene-Environment Interaction Models Testing Differential Susceptibility.

Sarah Moore & Felix Thoemmes

Differential Susceptibility Theory is primarily tested through modeling candidate gene by environment interactions (G x E). Often neglected in these studies are sources of gene-environment correlation, as the child's genotype (G) and an environmental factor (E) are often connected to the outcome through shared genetic confounds that influence experiences across development, as well as a child's sensitivity to experiences. In this proposal we clarify the causal processes at work in G x E models through a directed acyclic graph (DAG). Based on this DAG, we simulate data and evaluate the performance of different commonly used analytic models to identify effects of interest in G x E studies. We find that G x E estimates from models routinely used in the literature are only unbiased under unrealistic assumptions of absence of confounders. However true effects can be estimated without bias if models are expanded to include biasing effects of observed covariates.

27. The Effect of Measurement Reliability, Frequency, and Sample Size on the Statistical Power and Estimate Reliability in Longitudinal Growth Curve Models with Missing Data.

Andrew L. Moskowitz, Jennifer L. Krull, & Bruce F. Chorpita

Surveys of past psychological research have illustrated surprisingly low rates of statistical power. The current study addresses one question about longitudinal research: is it better to collect shorter, less-reliable measures more frequently or longer, more-reliable less frequently? Monte Carlo simulation was used to empirically examine how measurement reliability, assessment frequency, and sample size affect power and estimate reliability in the presence of missing data. One thousand repetitions for each 2,240 conditions were simulated and modeled using random coefficient regression. Results indicated that the loss in statistical power and estimate reliability resulting from reduced measurement reliability could be offset by increasing assessment frequency, and vice versa. Power and estimate reliability were reduced by attrition; however, intermittently missing data had no effect on estimate reliability. These results will help inform researchers on the most efficacious ways of designing their longitudinal research.

28. The Relationship Between Out-of-Pocket Spending and Maternal, Infant, and Under 5 Mortality.

Jonathan Noel

The relationship between aggregated, country-level measures of out-of-pocket costs and maternal, infant and under-5 mortality was determined using hierarchical multivariable regression. Two measures of out-of-pocket costs (out-of-pocket costs as a percent of total health expenditure and private health expenditure), and three measures of maternal and child mortality (maternal mortality, infant mortality, and under-5 mortality) were used. Economic and healthcare access covariates were selected using backward-selection multivariable regression. Significant effects of out-of-pocket costs as a percent of total health expenditure on maternal, infant, and under-5 mortality were not observed (p=0.53-0.95). The effects of out-of-pocket costs as a percent of private health expenditure on infant mortality were approaching significance (p=0.051). Out-of-pocket costs as a percent of total health expenditure is likely a proxy variable for prevailing economic conditions. Out-of-pocket costs as a percent of private health expenditure may affect infant mortality but results are confounded by inter-country differences in healthcare financing systems.

29. Modeling of Dropout Prediction Models Across Multiple States.

Lauren Porter & Lauren P. Bailes

High school dropout rates are a concerning trend for educators, administrators, education policy makers, and communities. A promising effort to address dropout rates is an empirical approach known as Early Warning Systems (EWS). The EWS is a prediction model that determines an individual student's risk of dropping out of high school. The prediction model is designed to alert educators and parents to the risk levels at an early stage to allow interventions to be employed early on in order to shift the trajectory of students displaying risks of dropping out.

30. A Latent Mixture Model of Growth Trajectories of Vocabulary Knowledge.

Jamie Quinn & Richard K. Wagner

In this methodological application study, we applied three growth modeling techniques to discover the latent class trajectories of vocabulary knowledge among a cohort of students (n=316) followed longitudinally from first grade to fourth grade. Three methodological techniques were applied: Latent growth curve analysis, to determine the nature of growth (e.g., linear, quadratic, cubic); latent class growth analysis, to determine the number of latent groups; and lastly, growth mixture modeling to determine the nature of growth within these latent groups. Results showed a three-class, non-linear growth pattern, whereby students in the highest performing group grew the most, and the lower-performing students grew the least. Implications for vocabulary knowledge growth and its relation to reading comprehension are discussed.

31. Predictive Relationship Between Residents' Undergraduate and Graduate Humanistic Clinical Skills Measured on a Licensing Examination using a 6-Period Crossover Design.

William Roberts

The purpose of this study is to investigate the relationship between undergraduate and graduate Humanistic clinical skill scores. Forty first-year and second-year osteopathic residents participated in this study. A 6-period crossover design was used to minimize measurement error associated with period, medical case task, and carry-over effects. The random subject effects covariance model showed a statistically and significantly predictive relationship between undergraduate and graduate scores, $F(1, 223) = 29.76$, $p < .001$. The best linear unbiased estimate for residents' undergraduate GPA was 0.70, controlling for period, case task, and cross-over effects. Results indicate a statistically significant case task effect, $F(5, 217) = 4.00$, $p < .05$. Period and cross-over effects were not significant. Generalizability (?^2=0.78;se=0.30) and dependability coefficients (?=0.71;SE=0.37) were similar to previous studies. The crossover design in conjunction with accepted measurement methods can be used to inform further development of measures to assess similar competencies in graduate medical training.

32. Examination of Population Heterogeneity on the Early Childhood Mathematics: A Growth Mixture Modeling Approach.

Ji Hoon Ryoo, Victoria Molfese, Elizabeth Brown, & Dingjing Shi

Development of mathematical skills in young children is important for preparing US student to be mathematically literate. The TEMA-3 (Ginsburg & Baroody, 2003) is useful for information on children's overall performance and cut scores for creating performance groups. Applying the growth mixture modeling (Muthen & Shedden, 1999; Muthen, 2004), growth patterns were explored using a longitudinal date from Head Start children aged 4 to 7 years. Two growth patterns were observed - growth of "average" performing children evidenced a ceiling effect at the end of 1st grade while that of "low" performing children show growth to 1st grade. No gender differences were found.

33. Multiple Imputation on Longitudinal and Multilevel Data: A Case Example From a School-Based Randomized Trial.

Ji Hoon Ryoo, Juliette Berg, & Catherine Bradshaw

The most common approach for addressing missing data is multiple imputation. Although the MI literature emphasizes its applicability to structured data with a large number of variables, regardless of whether they are in the substantive model, there is limited empirical research providing conclusive recommendations to applied researchers. Given the advantages and limitations of MI, in this project we examined (dis-) similarity between data-based imputation and model-based imputation for data that is longitudinal and multilevel. The ultimate goal is to provide applied researchers with selection criteria to decide between two imputation methods. We observed that the amount of bias produced by the two imputations were similar when the same substantive model was fitted, implying that applied researchers may be able to impute the missing information regardless of the substantive model.

34. Impact of Measurement Error when Analyzing a Panel Model in Continuous Time.

Leslie Shaw & Pascal Deboeck

Prior simulations have shown that analyzing a continuous time panel model with the Exact Discrete Model (EDM) is more accurate than Latent Differential Equations (LDE), but those studies have assumed no measurement error in the data. This simulation study will compare the performance of the two models at four different levels of measurement error added to the data. Normally distributed data was generated for two variables across four time points using a variety of sample sizes, auto-effect values, and cross-lag values. The focus of the analysis will be to determine under which conditions one model produces less biased estimates than the other model.

35. Multilevel Modeling of Cognitive Ability in Highly Functioning Adults.

Catherine Trapani

The goal of this research was to study differences in cognitive performance on verbal and quantitative measures among subjects of different ages. Data was gathered on subjects ranging in age from 16 to 80 years of age from birth-cohorts from 1927 to 1990. Multilevel models were built that predict cognitive performance as a function of age of subject at time of testing. Verbal performance rises as the age of the test-taker rises; quantitative performance declines as the age of the test-taker rises.

36. Order-Constrained Hypothesis Testing for Linear Regression Models: Increased Power in Small Samples.

Leonard Vanbrabant, Rens van de Schoot, & Yves Rosseel

Small samples are common in many social research. For example, when drawn from small populations or when the smallest possible sample is desired. When small samples are combined with classical null hypothesis (H0) testing, it appears that many effects may not be found. In other words the test lacks power. Consequently, many research questions remain unanswered. Recently, more sophisticated methods have been developed for testing the hypothesis of interest instead of H0. In this article, we focus on informative hypothesis testing. In such a hypothesis, expert knowledge is included in terms of order constraints. For example, researchers expect the following order between three means: $M1 < M2 < M3$. An advantage of this procedure is that smaller samples are needed to detect effects. However, hardly no literature exists about this gain in sample size reduction. Therefore, the main goal is to study the relationship between the number of order constraints and the possible reduction in sample size.

37. Fixed-Links Modeling as a Means to Separate Multiple Sources of Variance Underlying the Attentional Blink at the Behavioral and Electrophysiological Levels of Performance.

Felicitas L. Wagner, Karl Schweizer, Thomas H. Rammsayer, & Stefan J. Troche

This study introduces fixed-links modeling (FLM), a variant of structural equation modeling, to address the still unsolved impurity problem arising from the fact that hit rates obtained from experimental tasks usually are influenced by multiple sources of variance such as perception and attention. This leads to diminished correlations with other variables of interest. FLM allows for the separation of multiple processes underlying performance assessed by repeated-measures designs. FLM was applied to behavioral and electrophysiological data obtained from an attentional blink (AB) task. AB describes impaired detection of the second of two visual targets presented at the same position in close temporal proximity. A similar decrement of P300-amplitudes was previously reported, but correlations with second-target hit rates were low and inconsistent. However, substantial correlations were observed after separating multiple independent processes. FLM led to a clearer description of the processes of interest and higher convergent validity than a simple correlational approach.

38. Using Multilevel Modeling to Investigate College Women's Dual Method Contraception Use.

Jennifer Walsh, Robyn L. Fielder, Kate B. Carey, & Michael P. Carey

Dual method contraception use is effective but understudied. No prior studies have taken an event-level approach to examining characteristics associated with dual method use among college women. Twelve monthly surveys resulted in 1843 sexual intercourse events from 296 women at a Northeastern university. Women reported on their use of condoms and hormonal contraception during all events. Associated characteristics were assessed at the event-, month-, and person-level. Multi-level models assessed variables associated with dual method use. Women used hormonal contraception during 53% of events and condoms during 63% of events. Of events involving hormonal contraception use, condoms were used in 53%. Dual method use was more likely with partners who were friends and for women who were more religious and who reported a previous STI. Dual use was less likely when women were using less reliable contraceptive methods, had more experience with hormonal contraception, were older, or reported more sexual partners.

39. Measurement Invariance of Second-Order Factor Model of the Multifactor Leadership Questionnaire (MLQ) across K-12 Principal Gender.

Lihua Xu, Trae Stewart, & Zane Wubbena

The aim of this study was to investigate the factor structure and the measurement invariance of the Multifactor Leadership Questionnaire (MLQ) across gender of K-12 school principals (N=6,317) in the United States. Nine first-order factor models and four second-order factor models were tested using confirmatory factor analysis (CFA). The results suggested that the nine-factor model provided the best fit for the data. Further examination revealed that most constructs lacked convergent validity and discriminant validity. The second-order factor model with two higher-order factors (i.e., transformational and transactional leadership) was deemed the best fit and was then tested for measurement invariance between females and males. The measurement model was found to be invariant across gender. Findings suggested that female school principals demonstrated significantly greater transformational leadership behavior, while male school principals demonstrated significantly greater transactional leadership behavior.

40. Measurement Invariance of the Servant Leadership Questionnaire Across Gender.

Lihua Xu, Trae Stewart, & Paige Haber-Curran

The Servant Leadership Questionnaire (Barbuto & Wheeler, 2006) was applied to 956 k-12 principals in the United States. The 5-factor structure identified by the authors of the instrument was replicated in this study and the factor structure fit the data in both groups satisfactorily. The step-wise measurement invariance test showed the questionnaire has configural invariance, metric invariance and scalar invariance. The subsequent latent factor mean comparisons suggested that female principals displayed higher emotional healing, persuasive mapping, wisdom and organizational stewardship. There was no gender difference in latent mean of altruistic calling.

41. Interrater Agreement Statistics with Skewed Data: Evaluation of Alternatives to Cohen's kappa.

Shu Xu & Michael F. Lorber

Behavior observation is an important research method in behavioral sciences. Various types of inter-rater agreement statistics (IRAS) can be computed for assessing the agreement of raters' observations. These statistics all function equivalently when a behavior's base rate is 0.5. However, when the behavior base rate is low, the performance of these statistics varies. Given that there is no gold standard to determine "true agreement", we argue that an optimal IRAS should (1) provide some level of protection from chance agreement, (2) be minimally sensitive to distributional skew, and (3) maintain these qualities with biased and unbiased raters and at varying sample sizes. In this study, we evaluate the performance of IRAS at varying base rates and when two raters may have different rating biases.

42. Predictive Data Mining Modeling in Very Large Data Sets: A Demonstration and Comparison under Model Ensemble.

Hongwei Yang

The paper demonstrates the effectiveness of predictive data mining modeling under model ensemble. Three types of predictive models (decision tree, neural network, and regression) are each fitted using benchmark applications. Both the case of a categorical outcome and that of a numerical outcome are discussed. Two model ensemble techniques (bagging and boosting) are each added to the modeling process to pool predictions from individual component models in an effort to improve prediction accuracy and stability.

43. A 2-2-1 Multilevel Mediation Model to Study Peer Effects on Language Growth.

Gloria Yeomans-Maldonado, Jessica Logan, & Laura Justice

There is evidence suggesting that the average language skills of a child's classmates influences the child's language growth over an academic year. However, the processes underlying this growth remain unclear. The present study aims at understanding the way these peer effects occur, specifically looking at a direct effect (i.e. children's language skills are affected by their interactions with other children in the classroom) and an indirect effect (i.e. children's language skills at the beginning of the school year affect the quality of teacher instruction which in turn affect children's skills). Using data representing 586 preschool-aged children enrolled in 83 Early Childhood Special Education (ECSE) classrooms from multiple school districts in a single Midwestern state, authors used a 2-2-1 multilevel mediation model to understand the transmission process of language skills within a PreK-classroom.

Wednesday May 21st - Keynote Address
8:00 am – 9:30 am
Laurel Hall Room 102

Simple Methods for Handling Non-Randomly Missing Data

Dr. Sophia Rabe-Hesketh, University of California, Berkeley

In multiple linear or logistic regression, multiple imputation has become increasingly popular for handling missing covariate values. The much simpler approach of listwise deletion or complete-case analysis is often dismissed as making overly strong assumptions. However, I will point out that complete-case analysis is consistent and performs better than multiple imputation for many types of non-random missingness mechanisms. In longitudinal data analysis, dropout or intermittently missing responses are typically dealt with by specifying a joint model for the responses, such as a growth-curve/hierarchical/multilevel model, and estimating the parameters by maximum likelihood. This approach is consistent if missingness of a response depends on observed responses for the same individual but not if if it depends on the response itself or on the random effects in the model. One way of handling such non-random missingness is to model missingness jointly with the response variable of interest, but these joint models are complex, require

specialized software, and make unverifiable assumptions. I will suggest simple fixed-effects approaches that are consistent if missingness depends on the random effects and, in the case of binary responses, if missingness depends on the response itself or previous (observed or unobserved) responses.

Sophia Rabe-Hesketh is a Professor of Education and Biostatistics at the University of California, Berkeley. Her research interests include hierarchical/multilevel models, item-response theory, structural equation models, and generalized latent variable models. She has developed a general model framework "Generalized Linear Latent and Mixed Models," that unifies and extends these models and corresponding software, gllamm, that has been used in over 550 different journals. She is the current president-elect of the Psychometric Society. For more information about Sophia Rabe-Hesketh, check out her website: **www.gllamm.org/sophia.html**

Wednesday May 21st - Concurrent Session 4 9:45 am – 10:45 am (60 minutes)

Session 4.1: New Developments in Variable Selection for Regression Models (Laurel Hall 106)

An Introduction to Variable Selection for Regression Models.
Ofer Harel

An Empirical Bayes Approach to Variable Selection and QTL Analysis.
Haim Bar

We develop a model-based empirical Bayes approach to variable selection problems where the number of predictors is very large, possibly much larger than the number of responses (the so-called "large p, small n" problem). Motivated by QTL (quantitative trait loci) studies, we consider the multiple linear regression setting, where the response is assumed to be a continuous variable, and it is a linear function of the predictors. The explanatory variables in the linear model can have a positive effect on the response, a negative effect, or no effect. Thus, we model the effects of the linear predictors as a three-component mixture, where each component follows a normal distribution with mean μ, -μ, or 0. A key assumption in our approach is that only a small fraction of the candidate predictors have a non-zero effect on the response variable. By treating the putative variables as random effects we get shrinkage estimation, which results in increased power. This approach is computationally efficient because the number of parameters that have to be estimated is small, and remains constant regardless of the number of explanatory variables in the linear regression model. The model parameters are estimated using the EM algorithm which leads to significantly faster convergence, compared with simulation-based methods. Furthermore, we employ computational tricks which allow us to increase the speed of our algorithm, to handle a very large number of putative variables, and to avoid multicollinearity in the regression model.

Model Selection through Sparse Estimation in Finite Mixture Regression Models.
Elizabeth D. Schifano, Robert L. Strawderman, & Martin T. Wells

In regression modeling, the goal is to relate a response variable y to a set of covariates x=(x1,…,xp). The usual approach requires estimating a single set of regression coefficients, shared by all of the observed samples (y1, x1),…,(yN,xN). It is often the case, especially with a large number of covariates, that the N observed samples are not adequately modeled using the same set of regression coefficients; that is, a set or subset of coefficients may be different for different subgroups of observations. Additionally, it may be possible for some coefficients in some (or all) subgroups to be zero. In this talk, finite mixture regression model fitting is explored when the number of mixture components is potentially unknown and the regression coefficients within each component are allowed to be differentially sparse. A penalized likelihood framework and fitting algorithm is proposed which can estimate both the mixture parameters and the penalized regression coefficients. To demonstrate applicability of the method, simulation results and an

analysis of the well-known ozone meteorological dataset (Breiman and Friedman, 1985) are also presented. Joint work with Robert L. Strawderman and Martin T. Wells.

Session 4.2 (Laurel Hall 107)

Structural Equation Modeling and Factor Indeterminacy.

Edward Rigdon

Under general conditions, a given factor model is consistent with an infinite number of different and potentially diverse sets of factor scores--a phenomenon known as factor indeterminacy. As a discipline, structural equation modeling has tended to ignore or disregard factor indeterminacy. This presentation reviews the phenomenon, assesses its likely magnitude in practice, examines reasons why the discipline has tended to disregard the phenomenon, and argues that factor indeterminacy is an important phenomenon. The presentation concludes by describing changes needed to account for factor indeterminacy in structural equation modeling practice.

Session 4.3: Longitudinal Models (Laurel Hall 108)

Fit Criteria Performance and Parameter Estimate Bias in Growth Curve Models with Small Samples in the SEM Framework.

Dan McNeish

When modeling repeated measures data, researchers often rely on three statistical methods: model-based methods such as multilevel models (MLMs), design-based methods such as generalized estimating equations (GEEs), and latent curve models (LCMs). Recently, studies have addressed the small sample properties of MLMs (Maas & Hox, 2005) as well as GEEs (Morel, Bokossa, & Neerchal, 2003). However, no such studies have yet to address small sample properties of LCMs or the small sample performance of global model fit criteria, one advantage of modeling repeated measures, with LCMs.

Issues in Latent Growth Modeling with Longitudinal Public-Release Data.

Ming Li, Jeffrey Harring, & Laura Stapleton

Despite the explosion of empirical longitudinal research in recent years that has employed basic

LGMs, their flexibility to extend to situations of complex data analytic conditions often found in practice has been under-recognized. Some of these alternative designs include cohort-sequential designs and other planned missingness (Duncan, Duncan, & Stryker, 2006; Little, 2013), and unbalanced designs with measurements not obtained at equally-spaced occasions and possibly idiosyncratic to the individual (Cudeck, 1996; Harring, 2009). For researchers who design and collect their own data, some of these design and measurement challenges are addressed during the data collection phase. However, for researchers using secondary data, understanding what longitudinal design was implemented, how missing data was handled, what sampling plan was used and the nuances of the measurements themselves is crucial. With this as a backdrop, in this paper we discuss four issues to address when conducting an LGM analysis with public-release data: measurement, model choice, missing data, and sampling design. We use the ECLS-K data set to illustrate these issues with an end goal of providing a roadmap and checklist. Specifically, examples will include growth of child body weight over time, growth of cognitive ability, and change in parental involvement.

Wednesday May 21st - Concurrent Session 5
11:00 am – 12:00 pm (60 minutes)

Session 4.4: Effect Size and Confidence Intervals (Laurel Hall 109)

An Investigation of Accuracy and Precision of the Generalized Eta-Squared Effect Size Based on Various Research Designs.

Patrice Rasmussen, Patricia Rodriguez de Gil, Anh Kellerman, Thanh Pham, Jeanine Romano, Yi-Hsin Chen, & Jeffrey Kromrey

Generalized eta-squared (Olejnik & Algina, 2003) has been proposed as a variance-explained effect size that is comparable across a variety of ANOVA designs, including any number of between- and within-subjects factors. However, no empirical investigation has evaluated the accuracy and precision of this effect size index. The current study used simulation methods to investigate a variety of sample sizes, distribution shapes, population effect sizes, and ANOVA research design characteristics (number and levels of between- and within-subject factors, measured and manipulated variables). Findings conclude that substantial bias exists in effect size estimates with small samples, and more bias is evident in interaction effects versus main effects.

Robust Confidence Intervals for Effects Sizes in Multiple Linear Regression.

Paul Dudgeon

The literature on confidence intervals for effect sizes has arguably paid more attention to ANOVA designs than to multiple linear regression over the last decade. Appropriate intervals have been recently proposed for standardized regression coefficients (Yuan & Chan, 2011) and for the semipartial correlation (Aloe & Becker, 2012), but only under conditions of normality and homoscedasticity. This talk presents methods for calculating robust confidence intervals for typically-reported effect sizes under violation of linear regression assumptions within a covariance structure analysis framework. Two approaches to calculate robust standard errors under model misspecification (Yuan & Hayashi, 2006) are investigated by using either a HC1 or HC3 covariance matrix of model parameters (Long & Ervin, 2000). Results of a Monte Carlo simulation indicate that confidence intervals from the HC3-based estimator remain robust for different effect sizes in the vast majority of conditions that were investigated.

Session 4.5: Latent Transition Analysis (Laurel Hall 110)

Implementing the 3 Step Latent Transition Analysis in MPLUS using a Sub-Population from a Complex Sample Design.

Rafael R Ramirez, Jose Noel Caraballo, & Carmen Rivera Medina

MPLUS software will be used in a 3- step Latent Transition Analysis using data from a national complex sample of n=1,179 incarcerated adults in Puerto Rico. The 3- step method has been recently advocated by Vermunt (2010), and further elaborated by Asparouhov & Muthen (2013)to address the problematic issue of latent class's composition and meaning changing when auxiliary variables are included in the mixture model either to explain class composition, as distal outcomes or to explain subject's transition across time between the classes. Our illustration will guide the audience step by step in conducting a 3- step two-waves LTA without measurement invariance. The complete and annotated Mplus syntax will be provided and discussed to enable attendees to be able to implement the method with their own data.

Session 4.6: Advanced Modeling of Healthcare Data (Laurel Hall 111)

Hierarchical Bayesian Exploratory Factor Analysis for Health Care Quality Utilization and Quality Data.

Alan M. Zaslavsky & A. James O'Malley

We describe a Bayesian Exploratory Factor Analysis methodology for hierarchically structured (clustered) multivariate data. Our methodology consists of the following steps: (1) estimation of cluster mean vectors and their sampling variance-covariance matrices, (2) sampling of the level-2 covariance matrix under a

generic prior, (3) estimation of factor loadings and rotations for each draw using standard methods, (4) postprocessing these estimates with relabeling and further rotations to align factor loadings across draws, and (5) summarization of uncertainty of factor solutions. We illustrate its use in developing composite metrics for two distinct sets of health care quality data: Medicare utilization metrics for geographical areas, and a new survey of pediatric experiences with hospital care.

Modeling Multilevel Data with Cross-Classified Outcome Variables using Kronecker-Structured Covariance Matrices, with Applications to Multivariate-Outcome Random-Coefficient Models for Healthcare Quality Data.

Alan M. Zaslavsky & Laura A. Hatfield

Abstract: Consumer-reported survey measures of health care quality are associated with respondent characteristics, specifically age, health status, and education. Furthermore, these coefficients vary across health plans and are correlated. We seek to parsimoniously summarize the covariance structure of the concatenated vector of coefficients of the corresponding regressions for several outcomes. Conventional factor analysis methods essentially assume a one-way layout of variables, but these coefficients constitute a cross-classified structure replicated across plans; such structures are potentially common. We model these structures using Kronecker products, where one factor represents the association among predictors and the other the association among outcomes. These factors are estimated by conditional maximizations embedded in an EM algorithm for the underlying hierarchical model. Potential extensions of this modeling strategy extend the specification to sums of Kronecker products, relaxing the proportionality assumption of the basic model.

Session 5.1: Causal Modeling in Communication (Laurel Hall 106)

Iterative Meta-Causal Analysis: Modeling the Impact of Job Loss on Communication and Personality.

Mark Hamilton

Causal analysis of meta-analytic data presents unique challenges when compared to the modeling of primary source data. An iterative process that involves the concurrent consideration of primary and secondary data as well as static and longitudinal studies is considered. The process begins with theory-driven models of primary source data and extends to the synthesis of secondary data. Strategies for addressing variance due to third variables and artifacts across studies are suggested. Heterogeneity due to moderator variables requires techniques for differentiating additive and multiplicative effects. The iterative meta-causal analysis was applied to data on the effects of job loss and financial strain on social support and social undermining. The two communication variables partially mediate the effect of these employment-related stressors on self-esteem, depressed affect, and anxiety. The concept of mental health is evaluated as a second-order factor. Trauma is examined as a moderator variable. Several artifacts are identified.

Testing the Viability of Alternative Structures with a Distributed Computing System.

James Watt & Mark Hamilton

Structural Equation Modeling (SEM) is a common technique used in behavioral and social sciences. SEM fit tests will classify any particular structural model as either plausible (goodness-of-fit indices that indicate that the covariances predicted by the model are close to those actually observed in the data) or implausible. Finding a model implausible is reason to reject it, but a plausible finding does not mean that it is the single correct model to represent the observed data. Unfortunately, this unwarranted conclusion is common, as identifying all alternative and plausible models would require specifying and analyzing a daunting number of models (e.g. 4096 models with 4 variables; 1,048,576 models with 5 variables). In this presentation we describe a new multiprocessor cluster computer system that specifies all possible SEM models and evaluates and ranks each by goodness-of-fit, thus identifying all plausible SEM models that fit the observed data and may be alternative explanations.

Wednesday May 21st - Concurrent Session 5
11:00 am – 12:00 pm (60 minutes)

The Influence of Synchrony and Sensonry Modality on the Person Perception Process in Computer-Mediated Groups.

Kristine Nowak & James Watt

This study examined the effects of synchrony and the number of cues on the person perception process in computer-mediated communication. One hundred and forty-two participants in groups of three or four engaged in collaboration over five weeks to develop oral reports, using alternate versions of communication systems or meeting face to- face. Consistent with the hyperpersonal model, those using low cue media felt their partners were more credible, and reported more social attraction, less uncertainty, and more involvement in the interaction than those using high cue media. People interacting with synchronous media felt increased social attraction, self-reported involvement, and certainty. They also felt that their conversations were more effective, although this effect appeared mainly in low cue groups. Results of an exploratory path analysis suggest that future research should focus on causal chains rather than direct effects, and that intervening variables (such as involvement) may be central to our understanding of the effects of communication technology systems.

Analyzing and Modeling Behavioral Interaction Data.

Arther Vanlear & Teharan Davis

This paper deals with using sequential and time-series analyses and multi-level modeling, including the modeling of nonlinear dynamical systems

Session 5.2 (Laurel Hall 107)

An Introduction to Integrative Data Analysis in the Behavioral and Social Sciences.

Jennifer Walsh

Integrative data analysis (IDA) involves the analysis of multiple datasets that have been merged. IDA differs from meta-analysis in that it involves the combination of original data (not summary statistics), and it has a number of advantages in diverse fields. These advantages include increased sample size and statistical power; better representation of subgroups and low base-rate (rare or infrequent) behaviors; a broader, more valid assessment of constructs of interest; the ability to test hypotheses not considered in the original studies; and the potential to identify sources of between-study heterogeneity, which may inform theory and intervention development. This methodological illustration will provide an introduction to IDA using an example from HIV prevention interventions. We will discuss the benefits and challenges of IDA and potential applications of IDA in behavioral and social sciences research. Additionally, methods for integrating discrepant measures across studies and for testing cross-study replication will be demonstrated in Mplus.

Session 5.3 (Laurel Hall 108)

A Demonstration of a New Linear Modeling Procedure In SPSS Statistics: Automatic Linear Modeling (*linear*).

Hongwei Yang

This paper demonstrates the new Automatic Linear Modeling (LINEAR) procedure in SPSS 22.0 and introduces it as a new analytical tool for researchers who regularly use linear regression. To that end, the paper uses benchmark applications to examine two of its main features: 1) Automatic data preparation, and 2) automatic subset selection. Additionally, the paper also discusses several limitations of the LINEAR procedure, ways to improve it, and the model ensemble capability of the procedure as a possible topic for future research.

Wednesday May 21st - Concurrent Session 5
11:00 am – 12:00 pm (60 minutes)

Session 5.4 (Laurel Hall 109)

DataToText: Consumer-Oriented Dyadic Data Analysis using R.

David Kenny

DyadR is a program, written in R (a free open-source software environment), that offers the user a wide range options for dyadic data analysis. The program allows the user to restructure the data and conduct a wide range of analyses. Moreover, the program not only conducts the analysis but also provides a text summary and a graphical description of the results. The text contains warnings and cautions, as well as sensitivity analyses. As an illustration, an analysis of dyadic dataset examines whether gender can be ignored in the analysis, a power analysis, and an Actor-Partner Interdependence Analysis using both multilevel and structural equation modeling.

Session 5.5: Measurement Modeling (Laurel Hall 110)

An Investigation of the Alignment Method for Detecting Measurement Non-Invariance Across Many Groups with Dichotomous Indicators.

Jessica Kay Flake, Erin Strauts, Betsy McCoach, Jane Rogers, & Megan Welsh

In this study we evaluated the alignment method (Asparouhov & Muthen, 2013) for the accuracy of estimating factor models across many groups when the indicators are dichotomous. Further, we present research in which we investigate the correct decision rate for the ad-hoc procedure of testing invariance. Consistent across conditions the absolute bias was greater than .10 for the estimates of the loadings, though the bias decreased as sample size increased. Conversely, the thresholds exhibited less bias,

particularly when there were fewer groups. We also found consistent bias in factor variance estimates. However, the factor mean was better estimated in all large sample conditions. These results suggest that different parameters in the model may be better estimated under different conditions. Our results also show that the detection of non-invariance is far better for thresholds than for loadings. Implications for the appropriate use of the alignment with dichotomous data are discussed.

Dimensionality at Multiple Levels: Examining NAEP Mathematics with an Exploratory, Multilevel Item Factor Analysis Model.

Nathan Dadey & Gregory Camilli

Using exploratory multilevel multidimensional item response theory (i.e., exploratory full information item factor analysis), we examine student- and state-level dimensions of academic achievement on the 2009 National Assessment of Educational Progress fourth grade mathematics assessment. The goal is to define and estimate state-level dimensions that can be examined in light of the influences of educational policy and practice.

Session 5.6 (Laurel Hall 111)

Evaluating Measurement Equivalence and Translation Effectiveness of a Customer Engagement Instrument Across National Cultures and Types of Customers.

Dan Yu & Yongwei Yang

International companies that want to measure, compare, and manage customers' attitudes cross cultures face several challenges: (1) national culture affects consumers' attitudes and how they respond to customer survey; (2) Translation of survey questions also may introduce construct irrelevant variance; (3) B2B and B2C customers may evaluate and respond to survey questions differently. In this study, we applied several statistical methods to assess the across-culture measurement equivalence of a customer engagement instrument, including factor analysis, reliability analysis, weighted multidimensional scaling (MDS), and differential item functioning (DIF) indices using both Analysis of covariance (ANCOVA) and Hierarchical

Wednesday May 21st - Concurrent Session 5
11:00 am – 12:00 pm (60 minutes)

Multiple Regression (HMR). We studied 14 language-country combinations and 2 types of customers (B2B vs. B2C). The results indicated that the instrument functions similarly across different national cultures, and it holds true for both B2B and B2C customers. Utility of these methods for evaluating measurement equivalence and translation effectiveness of an instrument is also discussed.

A Mixture Model for Nuptiality Data with Long-Term Survivors.

Paraskevi Peristera & Gebrenegus Ghilagaber

The tacit assumption in the analysis of duration data with censored observations is that censoring time is independent of event time. This implies that the individuals who have not experienced the event of interest by the end of the study do not differ in any systematic manner from those who have experienced the event. This assumption is violated in many situations. For instance, in the analysis of data on family formation, individuals with a tendency to remain single over long periods may be overrepresented among the censored observations. In this work, we use a mixture model where parameters of a binary logistic regression model are jointly estimated with those of a continuous intensity model for family formation. The model allows for incorporation of a frailty-term for unobserved heterogeneity. Preliminary results show that failure to account for long term survivors may yield misleading results and plague the purpose of the analysis.

Lunch, 12:00 pm –1:00 pm
In the Student Union Ballroom

Session 6.1 (Laurel Hall 106)

A Workshop on Bayesian Nonparametric Regression Analysis.

George Karabatsos

Regression analysis is ubiquitous in educational research. However, the results of a regression analysis can become inaccurate, when data violate one or more assumptions of the specified regression model. Meanwhile, over the past decade, important progress has been made towards the development of Bayesian nonparametric (BNP) regression models. They are flexible, infinite-mixture models that make few and realistic assumptions about the data. However, most educational scientists are not aware of the inferential power of BNP regression models. This is because workshops on BNP analysis, and user-friendly BNP data analysis software, have not been made available to this audience. This methodological illustration workshop will provide pedagogical illustrations of how to perform flexible Bayesian nonparametric regression analysis of data, using software that I have developed. The software is menu-driven and user-friendly (like SPSS), and is freely-available on the internet. Through real-data illustrations, people who attend this workshop will learn how to use BNP models, and the software, for the regression analysis of continuous, binary, and ordinal dependent variables; and for multi-level analysis, quantile regression analysis, causal analysis, censored data analysis, meta-analysis, and item response analysis. Acknowledgement: This work is supported by NSF grant SES-1156372.

***Attendees are welcome to bring laptops to follow along with the computations**

Session 6.2: Propensity Score Analysis: Empirical Investigations of Common Problems and Their Impacts on Treatment Effect Estimates (Laurel Hall 107)

Impact of Measurement Error in Propensity Score Analysis.

Eun Sook Kim

Monte Carlo methods were used to investigate the impact of covariate measurement error on the efficacy of propensity score (PS) methods. Seven factors were crossed in the simulation design: number of covariates

(3, 9, 15, 30), population treatment effect (0, .2, .5, .8), covariate relationship to treatment (.1, .2, .4), covariate relationship to outcome (.1, .2, .4), correlation among covariates (0, .2, .5), sample size (50, 100, 250, 500, 1000), and covariate reliability (.4, .6, .8, 1.0). Each sample (5000 replications) was analyzed using seven PS methods (matching with and without a caliper, ignoring covariates, ANCOVA, PS as a covariate, stratification, and PS weighting). Outcome measures included treatment effect bias, RMSE, 95% CI coverage and width. Results indicate that even low levels of measurement error lead to substantial statistical bias in treatment effect estimates and reduction in CI coverage. Such effects were evident across conditioning methods and effects increased with greater amounts of measurement error, larger numbers of covariates, and greater strength of relationship between the covariates and both the treatment assignment and the outcome variable.

Treatment of Missing Data in Propensity Score Analysis.

Patricia Rodriguez de Gil

Missing data is a ubiquitous problem in applied research that, unless treated, can lead to bias and spurious inferences. Many approaches are available to handle missing data in PSA, but multiple imputation (MI) presents many advantages to other strategies. Using simulation methods, this study examined four approaches to MI within the context of PSA related to what is imputed and the treatment of the imputed data sets. For the target of imputation, we examined imputation of covariates only vs. imputation of the propensity score itself. This factor was crossed with two methods of treating the imputed data sets: averaging PS across imputations prior to conditioning and conditioning each imputed data set and averaging the results across imputations. The amount of missing data was investigated in both the percentage of observations presenting missing data (10% - 50%) and the percentage of variables missing for each case (10% - 50%). Outcome variables included statistical bias and RMSE of point estimates, and both confidence interval coverage and width. Results suggest that imputation of the propensity score is superior to imputation of covariates only and that separate conditioning of each imputed data set is superior to averaging propensity scores prior to conditioning.

Single-Level vs. Multi-Level Propensity Scores with Nested Data.

Patricia Rodriguez de Gil & Jeffrey Kromrey

Propensity score (PS) methods provide viable strategies for reducing selection bias in non-experimental (observational) studies. Most research on PS methods model the treatment assignment so that the estimated probability of receiving treatment allows for the identification of comparable individuals based on their individual characteristics. However, in nested data structures selection bias might result not only from differences in the characteristics of the individuals but also from differences in group membership. This study investigated differences in PS results from single-level and multi-level models. Data from an NSF funded project included school transcripts, demographics, enrollment, and achievement data. The impact of special educational programs on advanced mathematics course enrollment was investigated. Data were analyzed by comparing PS distributions, estimating the correlations between the two sets of propensity scores, and comparing the estimates of treatment effects. Results suggest a strong correlation between the PS obtained from single-level and multi-level models and only modest differences in resulting score distributions and estimates of treatment effects.

Covariate Balance in Propensity Score Models: Much Ado About Nothing?

Jeffrey Kromrey

Conventional wisdom indicates that sample covariate balance is critical to obtain unbiased estimates of treatment effects in propensity scores models. Consequently, researchers have been encouraged to report balance statistics before and after conditioning on the propensity scores. Further, sample evidence of imbalance after conditioning suggests that researchers should revise the propensity score model (e.g., alter the functional form of the model) in a search for a model that provides better balance in the covariables. However, the technical literature provides little empirical evidence of the relationship between covariate balance in a sample and the quality of the treatment effect estimate provided by that sample. The present study used simulation methods to investigate this relationship. The factors investigated included the correlation among the covariates ($r12$ = 0, .2, .5), the strength of relationships between the covariates and both treatment assignment and outcome (βj = .1, .2, .4), the number of covariates (k = 3, 9, 15, 30), the magnitude of the population treatment effect (Δ = 0, .2, .5, .8), sample size (n = 250, 500, 1000), and accuracy of model specification (correct specification, omitted covariates, incorrect functional form). Each

sample simulated was analyzed for both the degree of balance in the covariates and the deviation in the treatment effect estimate (i.e., the difference between the sample estimate of Δ and the parameter itself. Results indicate that for naïve models of the treatment effect (i.e., ignoring the covariates), samples with better covariate balance provide estimates of the treatment effect that are closer to the parameter than samples with poor covariate balance. Conversely, for propensity models that attempt to adjust for covariate differences (i.e., through matching, stratification, or ANCOVA on the propensity score), no relationship was evident between the sample covariate balance and the deviation of the treatment effect estimate from the parameter. Results are interpreted in terms of (a) the distinction between sample estimates and population parameters, and (b) the potential for our sample balance estimates to provide useful information about the quality of our propensity score model.

Session 6.3: Missing Data (Laurel Hall 108)

A Further Look into Planned Missing Data Designs in Analysis of Change.

Fan Jia, Wei Wu, Mijke Rhemtulla, & Todd D. Little

This study borrowed the insights from optimal design literature to evaluate planned missing designs in growth curve modeling. Monte Carlo approach was used to estimate the efficiency of growth curve parameters for a large selection of planned missing data designs in comparison to complete data designs with varying number and location of repeated measures. Different cost ratios of recruiting new participants and repeated measure, numbers of measurement occasions, shapes of change trajectory, additional attrition rates were taken into account in the simulation study. The preliminary result on a subset of the conditions suggested that complete data designs generally outperformed planned missing data designs when estimating parameters in linear models. However for quadratic models, while complete data designs were superior to planned missing data designs in estimating random effects, certain planned missing data designs turned out to be more efficient to estimate fixed effects. Cost ratio and attrition rate had more influence on selection of efficient designs for random effects than fixed effects.

A Latent Variable Chained Equations Approach for Multilevel Multiple Imputation.

Craig K. Enders & Brian T. Keller

During the last decade, the methodological literature has advocated for two missing data handling approaches, maximum likelihood estimation and multiple imputation. Although these methods are widely available for single-level analyses, fewer sophisticated missing data handling options are available for multilevel data structures. Multiple imputation provides an ideal solution for multilevel models. Despite the advantages, comprehensive imputation routines for multilevel data are either not yet available in popular general-use computer packages or possess important limitations. The overarching goal of this research is to describe a flexible chained equations multiple imputation algorithm for two-level data structures that can account for nominal, ordinal, and normally distributed variables and random slopes. Our simulation studies suggest that we generally obtain relative bias values (i.e., bias as a proportion of the true complete-data standard error) under the 0.40 threshold that Collins, Schafer, and Kam (2001) deem acceptable for a 15% missing at random condition. Furthermore, our normal-theory confidence interval coverage values were generally accurate as well.

Session 6.4: Item Response Theory (Laurel Hall 109)

Unipolar Item Response Models.

Joseph F. Lucke

I propose a new, unipolar item response model (UIRM) for psychological traits, such as addiction, depression, or chronic pain, for which the levels of the traits must be assumed non-negative. The UIRM is essentially a psychometric implementation of Stevens's psychophysical power law. I present a general UIRM together with three specific models --- the log-logistic, the lognormal, and the Weibull, along with their item information functions and general properties. I also present an analysis that exemplifies these model as viable alternatives to the standard (bipolar) IRT models.

Robustness of Mixture Item Response Models to Two Correlated Sources of Differential Item Functioning.

Erin Strauts & Jessica Kay Flake

This study investigates the practicality of using Mixture Modeling to detect uniform DIF in the case of polytomous indicators and one or two sources of DIF. This method has been implemented in a variety of contexts, but recovery of correct class membership in MixIRT models has thus far been found to be poor in simulation studies. In the current study the Mixture Item Response Theory (MixIRT) model did not choose the correct number of classes. The misclassification appears to be due to improper reliance on trait level to create classes. Using a two stage procedure where impact of the known group is assessed in a MIMIC framework before the mixture modeling is conducted improved classification, but classification is still poor. This study implicates that more work is needed before using this methodology in practice.

Realistic IRT Item Parameter Generation for Monte Carlo Simulation Studies.

Ling Ning, Cindy Walker, & Bo Zhang

In IRT Monte Carlo simulation studies, item parameters may be generated using an array of probability distributions. However, item parameters sampled from these distributions show much variability and deviate greatly from each other. In addition, there are big differences found between item parameters obtained randomly from probability distributions and those obtained from actual test data. This study looks into how the selection of different underlying probability distributions impacts the simulation study results. The other goal is to explore the optimal way of item parameter generation.

Session 6.5: Modeling Dyadic Data (Laurel Hall 110)

Modeling Growth in Dyads at the Group Level.

Thomas Ledermann & Siegfried Macho

For the study of growth in dyads, methods have been developed to analyze growth at the level of the dyad members. In this article, we present a novel approach that we call the Common Fate Growth Model (CFGM). This model permits an analysis of growth at the level of the dyads when members are either distinguishable (e.g., heterosexual couples) or indistinguishable (e.g., lesbian couples). To estimate the model, we describe the use of structural equation modeling (SEM) for both distinguishable and indistinguishable members. For indistinguishable members and small groups, such as families, we provide details for the use of multilevel SEM (MSEM). For both SEM and MSEM, we address the issue of measurement invariance (MI) and the estimation of group-level means. The models are illustrated with data from couples collected at seven measurement occasions. To aid the estimation of the models, Mplus code and Amos setups are provided.

A Structural Equation Model of Dyadic Discrepancy Over Time.

Holly Laws, Aline Sayer, Paula Pietromonaco, & Sally Powers

The working alliance, or collaborative psychotherapy relationship, is an established predictor of psychotherapy outcome. Yet few studies have measured the alliance taking both therapist and patient ratings into account. The present study examined similarity in patient and therapist perceptions of the working alliance, hypothesizing that greater similarity in alliance ratings would predict outcome in a study of psychotherapy for chronic depression. We propose to present a structural equation model that simultaneously measures dyadic discrepancy at multiple timepoints and relates these measures to outcome. This model is a version of a second-order latent growth model (Sayer & Cumsille, 2001) that is applied to change in discrepancy over time. The model simultaneously measures dyadic discrepancy in alliance ratings, models change in discrepancy over time, and tests whether alliance discrepancy change predicts depressive symptomatology. Preliminary analyses supported our hypotheses, indicating that increased alliance similarity predicted lower depression at the end of psychotherapy.

The Actor-Partner Interdependence Model for Categorical Dyadic Data: An Introduction to GEE.

Tom Loeys, William Cook, Olivia De Smet, Anne Wietzker, & Ann Buysse

The actor-partner interdependence model (APIM) has been widely used for the analysis of pairs of individuals who interact with each other. The goal of this presentation is to detail in a non-technical way how the APIM for binary or count outcomes can be implemented and how actor and partner effects can be estimated using generalized estimating equations (GEE) methodology. Both SPSS- and SAS-syntax needed

to estimate the model and the interpretation of the output are illustrated using data from a study exploring the effect of satisfaction with the relationship before the break-up on unwanted pursuit behavior after the break-up in formerly married partners. The use of this GEE method will allow researchers to test a wide array of research hypotheses.

Session 6.6: Modeling Secondary Data (Laurel Hall 111)

Using the Pair-Wise Likelihood Method to Analyze Large Datasets with Discrete Responses.

Maria T. Barendse, Frans J. Oort, Marieke E. Timmerman, & Y. Rosseel

Factor analysis of discrete data often relies on the assumption that the responses are manifestations of underlying normally distributed continuous scores. For small numbers of variables (e.g., six or less) it is already difficult to maximize the likelihood of multivariate response patterns, as it requires numerical evaluation of high-dimensional integration over all underlying normally distributed continuous scores. Alternatively, the sum of the likelihoods of the bivariate response patterns (i.e., pair-wise) of two-way contingency tables can be maximized. This relatively new pairwise maximum likelihood (PML) method performs satisfactorily in small data sets with small numbers of variables. Little is yet known about how well the PML estimation method works with larger numbers of variables that are often encountered in empirical data. In this presentation we use a simulation study to examine the performance of the PML estimation method in large datasets. We will investigate parameter estimates and fit indices. The performance of the PML estimation method will be compared to the performance of a stepwise robust weighted least squares estimation method.

Maximum Likelihood Adjustment of Anticipatory Covariates in Analyzing Retrospective Survey Data.

Gebrenegus Ghilagaber & Rolf Larsson

We propose a Maximum Likelihood (ML) procedure to estimate parameters of a multiplicative hazard model in the presence of anticipatory covariates. The issues are illustrated by estimating effects of educational level on risks of divorce in a piecewise-constant hazard model. For individuals with anticipatory educational level, we compute conditional probabilities of having attained the reported level before marriage. These probabilities are then used as weights in their contribution to the likelihood from which the adjusted parameters are derived. The results show that anticipatory analysis is harmless because the adjusted estimates of relative risks do not differ significantly from those of anticipatory analysis. This differs from previous Bayesian analysis of the same data set but re-estimation of relative risks using Bayes-estimated covariate-model parameters restored and confirmed our present result.

Wednesday May 21st – Closing Keynote Address
2:45 pm – 4:15 pm
Laurel Hall Room 102

Accounting for Individual Heterogeneity in Treatment Effect Analysis

Dr. Edward Vytlacil, New York University

Edward Vytlacil received his PhD in Economics from the University of Chicago in 2000. He is currently a Professor of Economics at New York University, having previously been a faculty member at Stanford University, Columbia University, and Yale University. He is a Co-Editor of the Journal of Applied Econometrics, and an Associate Editor for Econometrica and the Journal of Econometrics. Vytlacil's work has focused on the micro-econometric methodology for treatment effect and policy evaluation using

disaggregate data. A theme in his work has been in allowing for the effects of a treatment to vary across people, and allowing individuals to have some knowledge of their own idiosyncratic treatment effect and to act upon that knowledge. In addition to his work in econometric methodology, he has published empirical work in labor economics and health economics evaluating the returns to schooling, the returns to job training programs, and the effectiveness of medical interventions.

Modern Modeling Methods Conference
May 19-20 2015
University of Connecticut

Call for Papers - Proposals due January 30th, 2015

The Modern Modeling Methods (M^3) conference is an interdisciplinary conference designed to showcase the latest statistical modeling methods and to present research related to these methodologies. The fifth annual M^3 conference will be held May 19-20, 2015 at the University of Connecticut.

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, longitudinal modeling, and item response theory are especially encouraged. Given the interdisciplinary focus of the conference, it is completely acceptable to present papers that have been published or presented elsewhere. Presenters may select the length of the session that they prefer: 30 minutes, 60 minutes, or 90 minutes. We also welcome proposals for multi-paper symposia on thematically grouped topics. Generally, symposia sessions will be 90 minutes in length. Conference proposals for the Modern Modeling Methods conference may fall into one (or more) of four categories: Methodological Innovation, Methodological Application, Methodological Illustration, or Methodological Evaluation. Methodological Innovation proposals introduce a new technique. Methodological Evaluation proposals present the

results of empirical research evaluating a methodology. Most often, these will involve simulation studies. Methodological Application proposals present the methods and results of a real research study in which the technique was used. Methodological Illustration proposals provide a pedagogical illustration of when and how to use the technique; these papers are designed to help the audience be able to implement the technique themselves. 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 1,000 words and should include a description of the methodology to be illustrated as well as an outline of the paper/talk.

There are three different types of presentations: Paper sessions (in which authors submit a paper), Symposia (in which a group of authors submit a set of related talks/papers), and posters. Proposals for symposia should be include titles, authors, and brief descriptions/abstracts for all of the paper presentations within the symposium. Symposium proposals may be longer than 1000 words if needed, but they should be less than 2000 words. In addition, we will hold a poster session and reception on May 21st. Graduate students are also encouraged to submit proposals, especially poster sessions. All proposals should be submitted electronically.

Proposals for the 2015 conference are due January 30th, 2015. Notifications of presentation status will be emailed by February 18th, 2015. If you have any questions about the conference, please email D. Betsy McCoach at betsy.mccoach@uconn.edu.

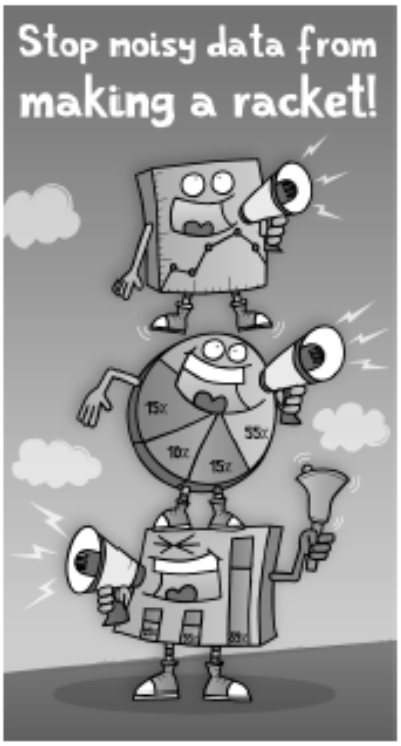
Statistical Innovations

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.

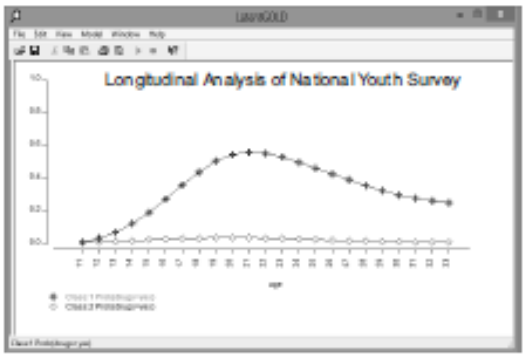
Segmentation with Longitudinal Data has Never Been Easier!

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