Conference Program

Applied Data Science Meeting

July 4-6, 2023

Fudan Institute for Advanced Study in Social Sciences

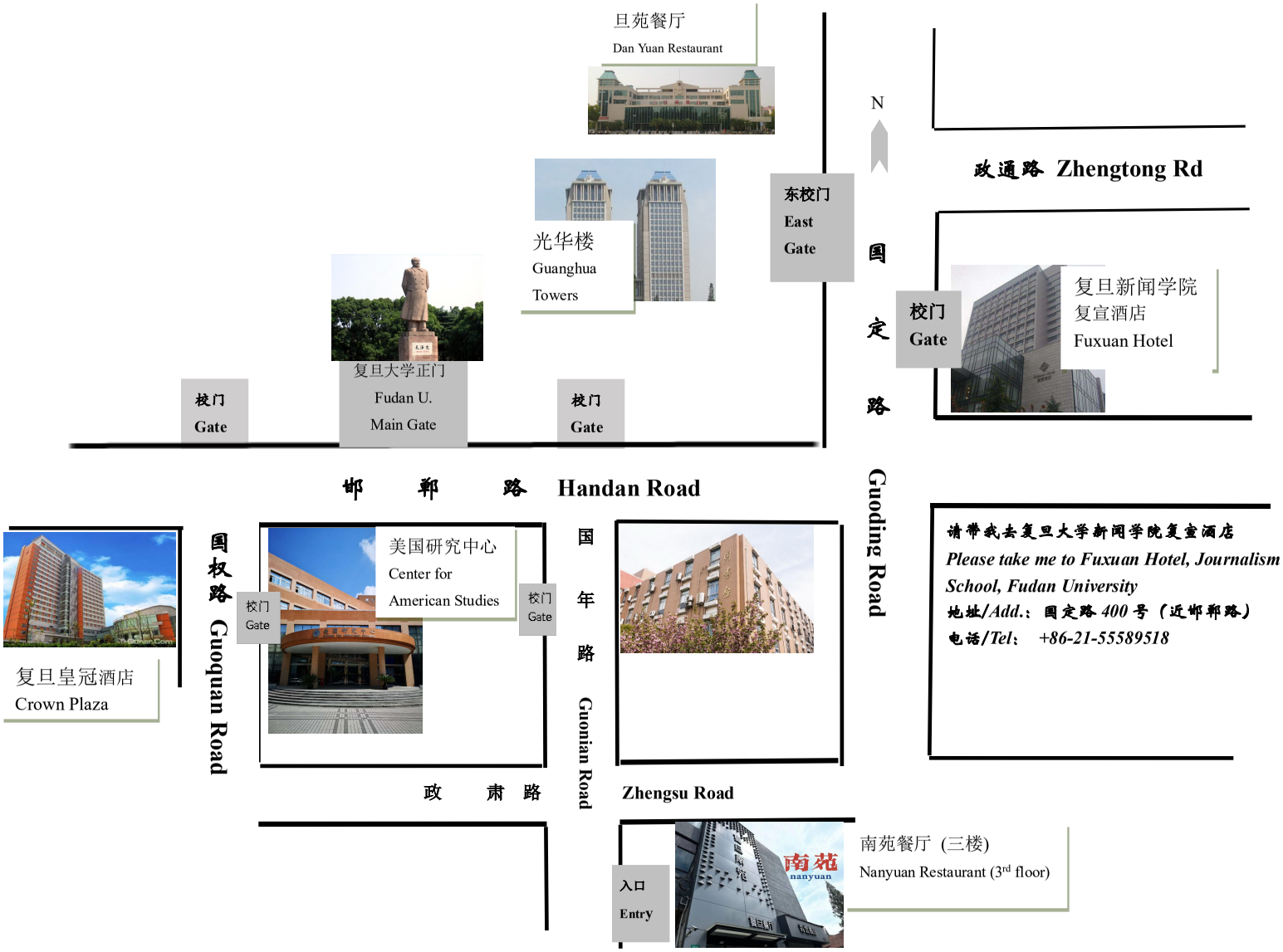
Shanghai, China

Organizing Committee

- Hawjeng Chiou, Distinguished Professor, National Taiwan Normal University
- Sujian Guo, Distinguished Professor, Fudan University
- Karl Ho, Associate Professor, University of Texas at Dallas
- Hongyun Liu, Professor, Beijing Normal University
- Wen Qu, Assistant Professor, Fudan University
- Jiashan Tang, Professor, Nanjing University of Posts and Telecommunications
- Ke-Hai Yuan, Professor, University of Notre Dame
- Zhiyong Zhang, Professor, University of Notre Dame

Sponsored by

Please contact the organizing committee at meeting@isdsa.org for any feedback.


旦苑餐厅
Dan Yuan Restaurant
光华楼
Guanghua
Towers
东校门
East
Gate
N
政通路 Zhengtong Rd
国
定
路
Guoding Road
校门
Gate
复旦新闻学院
复宣酒店
Fuxuan Hotel
复旦大学正门
Fudan U.
Main Gate
校门
Gate
校门
Gate
邯 郸 路 Handan Road
复旦皇冠酒店
Crown Plaza
国权路 Guoquan Road
校门
Gate
美国研究中心
Center for
American Studies
校门
Gate
国
年
路
Guonian Road
政 肃 路
Zhengsu Road
入口
Entry
南苑
nanyuan
南苑餐厅 (三楼)
Nanyuan Restaurant (3rd floor)
请带我去复旦大学新闻学院复宣酒店
Please take me to Fuxuan Hotel, Journalism
School, Fudan University
地址/Add.：国定路 400 号（近邯郸路）
电话/Tel： +86-21-55589518

Schedule

All times are referenced to the time zone used in China.

Conference Venue：Center for American Studies, Meeting Room 104
主会场: 复旦大学美国研究中心 104 会议室

Online (Zoom):
https://notredame.zoom.us/j/99528720917?pwd=ZWtVdzlEV0ZoQTgxYkRyRy9Bd3B5Zz09
Meeting ID: 995 2872 0917
Passcode: 980408

Workshop (Zoom) Schedule at the end
Recording place: Room 2801 East Main Building Guanghua Towers
工作坊录制地点: 光华楼东主楼 2801 室

Lunch: Nanyuan Restaurant 3F (third floor)
午餐：南苑餐厅 3F

July 4, 2023

Time	Title & Presenter
Breakfast	Provided by Fuxuan Hotel
Morning	Session Chair: Dr. Zhiyong Zhang
8:30-8:45	**Welcome and opening remarks** Sujian Guo, Dean and Distinguished Professor of Fudan IAS Zhiyong Zhang, Professor of University of Notre Dame, President of ISDSA
8:45-9:15	**Large-Scale International Education Surveys: Analyses of Slopes Can be More Interesting Than Comparisons of Means** Kit-Tai HAU*; The Chinese University of Hong Kong; HK, China Matthias von Davier; Lynch School of Education and Human Development, Boston College; US
9:15-9:45	**From Narratives to Data: How Data of Social Networks and Criminal Records Change Paradigms of Early China** (on Zoom) Liang CAI*; University of Notre Dame; US

9:45-10:15	**Modeling intraindividual variability as a predictor with longitudinal data: Methods and evaluations** Lijuan Wang*; University of Notre Dame; US Xiao Liu; University of Texas - Austin; US
10:15-10:30	**Break/Group Photo**
10:30-11:00	**Using Baseline Case Characteristics to Predict Successful Treatment Completion** (on Zoom) Xiaolan Liao*; The University of Oklahoma Health Sciences center; US Hairong Song; The University of Oklahoma; US David Bard; The University of Oklahoma Health Sciences center; US
11:00AM-12:00PM	*Speed talks* 1. **A 2-stage Latent-standardization Method in Second-order Latent Growth Modeling** Zhonglin Wen*; South China Normal University; CN Yifan Wang; South China Normal University; CN 2. **PCA, FA, LCA, LPA, LDA, which model is the best choice for your data?** Laura Lu*; University of Georgia; US 3. **Multitask prediction algorithms and their development** (on Zoom) Li Xing*; The University of Saskatchewan; CA 4. **A Comparison of Mathematical and Statistical Modeling with Longitudinal Data: An Application to Ecological Momentary Assessment of Behavior Change in Individuals with Alcohol Use Disorder** Sijing Shao*; University of Notre Dame; US 5. **Psychoperiscope: A data ingestion tool and modeling scale for social and behavioral data science** (on Zoom)

	Joshua Gandi*; Department of Psychology University of Jos, Nigeria; NG
12:00-1:00	Lunch break
Afternoon	Session Chair: Dr. Jiashan Tang
1:00-1:30	**Using response time for compromised item detection** Cheng Liu; University of Notre Dame; US Kyung Han; Graduate Management Admission Council; US Jun Li*; University of Notre Dame; US
1:30-2:00	**scAnnotate: an automated cell type annotation tool for single-cell RNA-sequencing data** Xuekui Zhang*; University of Victoria; CA
2:00-3:00	*Speed talks* 1. **Factors Affecting Reading Growth in Language Minority Learners** Yanran Chen*; University of Notre Dame; US 2. **Automated evaluation and measurement model for the effectiveness of state transitions in problem-solving tasks** Pujue Wang; Beijing Normal University; CN Hongyun Liu*; Beijing Normal University; CN 3. **Fit Multilevel Differential Equation Models to Intensive Longitudinal Data Using Numerical Optimization** Qingshan Liu; Beijing Normal University; CN Yueqin Hu*; Beijing Normal University; CN 4. **A hybrid method: Solve the impact of variable ordering in Bayesian network structure learning** Minglan Li*; Beijing Normal University; CN Fandi Chen; Hangzhou Mingshitang Digital Technology Co., LTD; CN

	Mengwei Wu; Hangzhou Mingshitang Digital Technology Co., LTD; CN Yueqin Hu; Beijing Normal University; CN 5. **Assessing relative importance of multiple mediators** Xun Zhu; East China Normal University; CN Xin Gu*; East China Normal University; CN
3:00-3:15	Coffee break
3:15-3:45	**Causal Effect Estimation and Transportation Using Modified Bootstrap** Tao Liu*; Brown University; US
3:45-5:00	*Speed talks* 1. **Integrating Structural Equation Modeling with Social Networks** Ziqian Xu*; University of Notre Dame; US Zhiyong Zhang; University of Notre Dame; US 2. **Revisiting the Evaluation with Coefficient Alpha on Varying Length Scales and Sample Sizes** Leifeng Xiao*; Department of Educational Psychology, The Chinese University of Hong Kong; HK; China Kit-Tai HAU; Department of Educational Psychology, The Chinese University of Hong Kong; HK; China 3. **How does prior distribution affect model fit indices of BSEM** Yonglin Feng; Department of Psychology, Sun Yat-sen University; CN Junhao Pan*; Department of Psychology, Sun Yat-sen University; CN 4. **How does Prior Variance Affect Local Dependence Detection in CFA** Xinyu Qiao*; Department of Psychology, Sun Yat-Sen University.; CN Junhao Pan; Department of Psychology, Sun Yat-Sen University.; CN

	5. **Synthesizing data from pretest-posttest-control-group designs in mediation meta-analysis** Zhiming Lu; Sun Yat-sen University; CN Zijun Ke*; Sun Yat-sen University; CN Rebecca Cheung; University of Reading; GB Qian Zhang; Florida State University; US 6. **Ordinal Logistic Regression Model in Determining Factors Associated with Household Food Insecurity in Namibia** (on Zoom) Dibaba Bayisa Gemechu*; Namibia University of Science and Technology (NUST), Windhoek; NA Leonard Elifas; Namibia University of Science and Technology (NUST), Windhoek; NA
6:00-	Banquet @ Fuxuan Hotel
July 5, 2023	
Time	Title & Presenter
Breakfast	Fuxuan Hotel
Morning	Session Chair: Dr. Ke-Hai Yuan
8:30-9:00	**Parasocial but Meaningful: Exposure to Foreign Cultures Encourages Pro-Outsider Attitudes** Gong Chen*; Fudan University; CN
9:00-9:30	**Estimation of functional-coefficient autoregressive models with measurement error** Pei Geng*; Illinois State University; US
9:30-10:30	*Speed talks* 1. **Detecting mediation effects with the Bayes factor: Performance evaluation and tools for sample size determination** (on Zoom) Xiao Liu*; The University of Texas at Austin; US 2. **Supervised Machine Learning Applications for Detecting Internet Research Agency Misinformation** (on Zoom) Thomas Wiese*; SUNY Empire State College; US

	3. **Construct Different Uncertainty Indexes to Forecast Macroeconomic Variables** (on Zoom) Yang Liu*; Department of Economics at Rutgers University-New Brunswick; US 4. **An analysis of emotion contagion with longitudinal text data** Xin Tong*; University of Virginia; US Haiyan Liu; UC Merced; US Shelly Tsang; University of Virginia; US Adrienne Wood; University of Virginia; US 5. **Chinese Adolescents' Career Choice in Large-Scale International Studies: Trends and Methodologies** Luyang Guo*; the Chinese University of Hong Kong; CN Kit-Tai HAU; The Chinese University of Hong Kong; CN
10:30-10:45	Coffee break
10:45-12:00	*Speed talks* 1. **A psychometric network analysis of adolescents' mental health status: Evidence for key features for interventions** ANG XIA*; Sun Yat-sen University; CN Ruyi Ding; Sun Yat-sen University; CN 2. **Communication Patterns of Adolescents' Distress in Mother-Adolescent Dyads Predict Adolescents' Suicidal Ideation: Mediated by Adolescents' Internalizing and Externalizing Problems** Zhou Yue; Sun Yat-Sen University; CN Ruyi Ding*; Sun Yat-Sen University; CN 3. **Determinate the number of attributes using neural network** Yali Wang; jiangxi normal university; CN Yujun Li*; Jiangxi Normal University; CN 4. **The InterModel Vigorish for Model Comparison in Confirmatory Factor Analysis with Binary Outcomes**

	Lijin Zhang*; Graduate School of Education, Stanford University; US Benjamin Domingue; Graduate School of Education, Stanford University; US 5. **The Roles of Emotion Regulation Strategies in Mothers' and Adolescents' Depression: Using the Actor-Partner Interdependence Model** (Recording) Zhonghuang Su; Sun Yat-sen University; CN Wenyu Qiu; Sun Yat-sen University; CN Ruyi Ding*; Sun Yat-Sen University; CN Junhao Pan; Department of Psychology, Sun Yat-Sen University; CN
12:00-1:00	Lunch Break
Afternoon	Session Chair: Dr. Hongyun Liu
1:00-1:30	**Two-Parameter State Response Measurement Model for Process Data and Its Mixture Extension** Yue Xiao; East China Normal University; CN Hongyun Liu*; Beijing Normal University; CN
1:30-2:00	**How Does Self-Compassion Associate with Coping Self-Efficacy in Daily Life? A Dynamic Structural Equation Model Analysis** Huihui Wang*; Ningxia University; CN Guifen He; Zhejiang University; CN Enna Wan; Tianjin University; CN Yang Wang; Guangdong University of Finance; CN MingQing Zheng; Ningxia University; CN
2:00-3:00	*Speed talks* 1. **A Tutorial in Longitudinal Network Analysis** Huihui Wang*; Ningxia University; CN Xuzhu Zhou; Ningxia University; CN MingQing Zheng; Ningxia University; CN 2. **Improvements to Regularization Differential Item Functioning Detection Methods: Combining Regularization with Traditional Methods** Kaiyu Gan; Beijing Normal University; CN

	Hongyun Liu*; Beijing Normal University; CN 3. **Evaluating the Threshold of Missingness Percentage in the Modeling of Intensive Longitudinal Data** Zhilin Wan*; Sichuan Normal University; CN Yue Liu; Sichuan Normal University; CN 4. **Estimating Average Treatment Effects in the Context of Outcome-Dependent Sampling Designs with Mismeasured Outcomes** MIN ZENG*; Department of Statistics and Finance, School of Management, University of Science and Technology of China; CN Hong Zhang; University of Science and Technology of China; CN 5. **The Story After Mobility: Debates and Critiques on Methods for Estimating Social Mobility Effects** Peng Wang*; Department of Sociology, The Chinese University of Hong Kong; CN
3:00-3:15	Coffee break
3:15-3:45	**Assessing between- and within-person reliabilities of items and scale for daily procrastination: A multilevel and dynamic approach** Xiaohui Luo; Beijing Normal University; CN Yueqin Hu; Beijing Normal University; CN Hongyun Liu*; Beijing Normal University; CN
3:45-5:00	*Speed talks* 1. **On the Relationship between Factor Loadings and Component Loadings when Latent Traits and Specificities are Treated as Latent Factors** Kentaro Hayashi*; University of Hawaii at Manoa; US Ke-Hai Yuan; University of Notre Dame; US 2. **Multiverse-style analysis: Current bibliometric analysis and tutorial** Shunsen Huang; State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University; CN

	Xinmei Zhao; State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University; CN Xiaoxiong Lai; State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University; CN Xinran Dai; State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University; CN Huanlei Wang; State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University; CN Yun Wang*; Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University; CN **3. L0-regularized high-dimensional mediation analysis** Saijun Zhao*; University of Science and Technology of China; CN Hong Zhang; University of Science and Technology of China; CN **4. The Impact of Machine Learning Approach In Recruitment Process with the use of NLP Algorithm** Sandra Ihemeje*; ISDSA Nigeria; NG **5. Effects of BAS/BIS on person-situation dynamics** Fan Wu; Beijing Normal University; CN Yueqin Hu*; Beijing Normal University; CN **6. Investigating the Factor Structure of Sense of Social and Academic Fit Scale: A Multilevel Bifactor Study** Lixin Wu*; University of Illinois at Urbana-Champaign; US Ge Jiang; University of Illinois at Urbana-Champaign; US
5:00PM -	Dinner on own
July 6, 2023	
Breakfast	Fuxuan Hotel
Time	Title and Presenter
Morning	Session Chair: Dr. Wen Qu
8:30-9:00	**Bayesian Compositional Data Analysis with Informative Priors** Xin Tong*; University of Virginia; US Di Jia; University of Virginia; US Yunli Liu; University of Virginia; US

9:00-9:30	**Impact of Temporal Order Selection on VAR-based Clustering of Intensive Longitudinal Data** Hairong Song*; Department of Psychology University of Oklahoma; US Yaqi Li; University of Oklahoma; US
9:30-10:30	*Speed talks* 1. **Small Data Approaches to Link Faster Time Scale Engagement Dynamics with Slower Time Scale Outcomes in Biobehavioral Interventions** Jingchuan Wu*; Pennsylvania State University; US Nilam Ram; Stanford University; US James Marks; Penn State College of Medicine; US Necole Streeper; Pennsylvania State Health Milton S. Hershey Medical Center; US David Conroy; Pennsylvania State University; US 2. **China's COVID Lockdown Policy and Trade with U.S.: A Deep Learning Time Series Approach** (on Zoom) Min Shi*; The University of Texas at Dallas; US Karl Ho; The University of Texas at Dallas; US 3. **Rationality Evaluation and Continuous Improvement of the Curriculum System Using Factor Analysis** Yi Kuang*; Xiangtan University; CN 4. **Can testing/training dataset separation ensure the reproducibility?** Yongtian Cheng*; University College London; GB Xijuan Zhang; York university; CA 5. **Estimation of the learning trajectories with observed and latent heterogeneity: A multi-sample study of growth mixture modelling** Pi-Fang Lin*; School of Teacher Education, National Taiwan Normal University; TW, China Hawjeng Chiou; College of Management, National Taiwan Normal University; TW, China
10:30-10:45	Coffee break

10:45-12:00	*Speed talks* 1. **The Dynamic Impact Between Intraindividual Perceived Stress and Perceptual Speed in Young and Old Individuals: Evidence from an Intensive Longitudinal Study** Xiaoyu Chen*; Beijing normal university; CN Hongyun Liu; Beijing normal university; CN 2. **Using semantic items to detect careless responding and warning in real-time and its effect on questionnaires' quality** Chang Xin*; Beijing Normal University; CN Jian Li; Beijing Normal University; CN 3. **When Predictors Sum to a Constant: Trade-off Effect Analysis Using a Regression Model Based on Isometric Log-ratio Transformation** Jieyuan Dong*; Beijing Normal University; CN Hongyun Liu; Beijing Normal University; CN 4. **A Two-Step Method based on lz* for Identifying Effortful Respondents** Yilan Chen*; Beijing Normal University; CN Hongyun Liu; Beijing Normal University; CN Yue Liu; Sichuan Normal University; CN 5. **Implementation of Gradient Boosting for Survival Analysis with Competing Risk** Md Hasinur Khan*; ISRT, University of Dhaka; BD 6. **Evaluation of the Bass-Ackward Method for Identifying the Number of Factors** Lingbo Tong*; University of Notre Dame; US Zhiyong Zhang; University of Notre Dame; US
12:00-1:00	Lunch Break
Afternoon	Session Chair: Dr. Zhiyong Zhang

1:00-1:30	**University Major Decision and its Effect on Wage: Modeling Interaction between Major Specificity and Education-Job Relevance Using Machine Learning Approach** Tian Hang*; Holly Futures, Co.; CN Karl Ho; University of Texas at Dallas; US Yong Zhou; Jiangsu SOHO Holdings Group CO., Ltd.; CN
1:30-2:00	**Modeling Data with Measurement Errors but without Predefined Metrics: Fact vs Fallacy** Ke-Hai Yuan*; University of Notre Dame, USA; US Zhiyong Zhang; University of Notre Dame, USA; US
2:00-3:00	*Speed talks* 1. **The Re-individualization of the Internet: A Comparison Analysis of Aggressiveness and Politeness in Social Media Posts Before and After the Compulsory GEO-IP Tag Policy** Yujun Li*; Jiangxi Normal University; CN Wenzheng Lin; Jiangxi Normal University; CN Zhihui Lai; Jiangxi Normal University; CN Xu Feng; Jiangxi Normal University; CN 2. **User Priority based Task Allocation and Predictive Scheduling Model for QoS enhancement in Cloud-IoT** Muthusenthil Balasubramanian*; SRM Valliammai Engineering College Department of CSE SRM Nagar Chennai; IN Devi Karuppiah; SRM Valliammai Engineering College Department of CSE SRM Nagar Chennai; IN 3. **Statistical power for linear and quadratic growth curve models with ignorable and non-ignorable missing data** Zhiyong Zhang*; University of Notre Dame; US 4. **Election Winning Prediction Using Demographic and Past Political Data: Data Science Approach** (Recording) Ismail Olaniyi MURAINA*; Lagos State University of Education; NG

	5. **Fuzzy Feature Selection (FFS) Based on Different Experts' Judgments: A case Study Epilepsy's Disease** (Recording) Hojjatollah Farahani*; Tarbiat Modares University; IR Parviz Azadfallah; Tarbiat Modares University; IR Peter Watson; Cognition and Brain Unit, University of Cambridge; GB Marija Blagojević; University of Kragujevac, Faculty of Technical Sciences Čačak; CS Farzin Bagheri Sheykhangafshe; Tarbiat Modares University; IR Mehrnaz SalehiRaj; Tarbiat Modares University; IR
3:00PM	Meeting ends

Workshop Schedule	
July 3, 2023	
8:00AM- 11:00AM	Workshop 1: Introduction to Social Network Analysis Prof. Haiyan Liu; UC-Merced; US
1:00PM- 4:00PM	Workshop 6: Bayesian Longitudinal Data Modeling Prof. Cynthia Tong; University of Virginia, US
8:00PM- 11:00PM	Workshop 2: Longitudinal Data Analysis (In Chinese) Prof. Hongyun Liu; Beijing Normal University; CN
July 5, 2023	
8:00PM- 11:00PM	Workshop 3: Practical Mediation Analysis Prof. Laura Lu, University of Georgia, US Prof. Qian Zhang, Florida State University, US
July 6, 2023	
8:00PM- 11:00PM	Workshop 4: Deep Learning Using R Zhiyong Zhang*; University of Notre Dame; US
July 7, 2023	
8:00AM- 11:00AM	Workshop 5: From Latent Class Model to Latent Transition Model Using Mplus (In Chinese) Hawjeng Chiou; College of Management, National Taiwan Normal University

Invited Workshops

Workshop 1: Introduction to Social Network Analysis
Instructor: Haiyan Liu -- University of California, Merced

Time: 8am-11am, July 3, China Time / 8pm-11pm, July 2, US Eastern Time
Location： Virtual on Zoom

Social network analysis is becoming increasingly popular in social, educational, and psychological sciences. This interactive course intends to provide participants with a detailed introduction, practical examples, and demonstration of analyzing social network data using the free software R. Topics covered include (1) Network Data; (2) Network Visualization; (3) Network Statistics; (4) Basic and Advanced Network Models. Especially, we will cover classical models such as Stochastic Block Model, Exponential Random Graph Model, Latent Space Model, and the newly developed techniques such as Latent Factor Space Modeling and Network Mediation Analysis.

Dr. Haiyan Liu is currently an assistant professor of Quantitative Methods, Measurement, and Statistics at the University of California, Merced. Dr. Liu's research centers broadly on statistical modeling of psychological and behavioral data such as high-dimensional data, longitudinal data, and categorical data. Her recent research includes social network modeling, Bayesian structural equation modeling, and non-parametric modeling of growth curves.

Workshop 2: Longitudinal Data Analysis (In Chinese)
Instructor: Hongyun Liu -- Beijing Normal University

Time: 8pm-11pm, July 3, China Time / 8am-11am, July 3, US Eastern Time
Location： Virtual on Zoom (Register to get the link)

Longitudinal Data Analysis (LDA) is a popular statistical method in various fields, including psychology, education, sociology, management, economics, and medicine. This workshop aims to provide participants with an overview of the applications and capabilities of LDA. Topics include: (a) introduction of longitudinal designs and longitudinal data, (b) latent growth curve models and growth mixed models, (c) cross-lagged panel models, (d) separating within- and between-person effects models, and (e) advances in LDA. The workshop will consist of lectures and software demonstrations using Mplus and R with practical examples.

Dr. Hongyun Liu is a professor in Faculty of Psychology at Beijing Normal University and has been working on the research of quantitative psychology for nearly 20 years. Her main research areas are psychological statistics, the theory and applications of psychological and educational measurement, and quantitative research methods. Her research interests involve the Internet-based test development, the theory and application of educational and psychological assessment, data analysis methods in large-scale assessments, and longitudinal data analysis.

Workshop 3: Practical Mediation Analysis
Instructors: Laura Lu -- University of Georgia; Qian Zhang -- Florida State University

Time: 8pm-11pm, July 5, China Time / 8am-11am, July 5, US Eastern Time
Location： Virtual on Zoom (Register to get the link)

In social sciences, an input variable often affects an outcome variable via a third variable, or mediator. Mediation analysis is used to answer the question of "how"/ "why" the input variable affects the outcome. This workshop will introduce the principles and methods (e.g., multivariate regression, path analysis, multilevel models, etc.) for mediation analyses. Emphasis will be on data analysis and interpretation. After the workshop, attendants should be able to propose research questions related to mediation analyses, use advanced methods such as bootstrap to test mediation effects, decide which type of model is appropriate for mediation, interpret and display empirical findings, and understand some other important issues of mediation research such as longitudinal designs for assessing mediation and missing data problems.

Dr. Laura Lu is an associate professor in the Quantitative Methodologies (QM) program at the University of Georgia (UGA). In general, she has expertise in structural equation modeling (SEM), longitudinal data analysis, hierarchical linear modeling (HLM), and computational statistics such as Bayesian. During her professional career at UGA, her research has focused on developing innovative statistical approaches to address perennial challenges in statistical modeling such as mixture structure, reliabilities, model selection, missing data, outliers, topic modeling, and also promoting statistical models to applied research areas through collaborating, mentoring, and classroom teaching.

Dr. Qian (Jackie) Zhang is an associate professor in the Department of Educational Psychology and Learning Systems at Florida State University. Her research interests focus on causal inferences, longitudinal data analysis, and effect size synthesis using multilevel models and multilevel structural equation models. She is also interested in handling sub-optimal data conditions with missing data, measurement error, and/or confounding variables in statistical analyses. She collaborates with substantive researchers in the areas of early childhood development, sports psychology, and educational research. Her work has been published in prestigious journals such as Psychological Methods, Multivariate Behavioral Research, Structural Equation Modeling, British Journal of Mathematical and Statistical Psychology, and Behavioral Research Methods.

Workshop 4： Deep Learning Using R
Instructor: Zhiyong Johnny Zhang -- University of Notre Dame

Time: 8pm-11pm, July 6, China Time / 8am-11am, July 6, US Eastern Time
Location： Virtual on Zoom (Register to get the link)

Deep learning is a very active area of research in the machine learning and artificial intelligence communities. The foundation of deep learning is the neural network. In this workshop, we will teach how to practically carry out deep learning using the R package keras. Several real-world data sets will be used to illustrate how to build single layer, multiple layer, convolutional, and recurrent neural networks. Basic knowledge of R is expected.

Dr. Johnny Zhang is a professor in Quantitative Psychology at the University of Notre Dame. His research aims to develop better statistical methods and software in the areas of education, health, management and psychology. He has conducted research in the areas of Bayesian methods, Big data analysis, Structural equation modeling, Longitudinal data analysis, Mediation analysis, and Statistical computing and programming. His most recent research involves the development of new methods for social network and text analysis.

Workshop 5: From Latent Class Model to Latent Transition Model Using Mplus (In Chinese)
Instructor: Hawjeng Chiou -- National Taiwan Normal University

Time: 8am-11am, July 7, China Time / 8pm-11pm, July 6, US Eastern Time
Location： Virtual on Zoom (Register to get the link)

Latent class modeling (LCM), a branch of latent variable modeling, has become increasingly popular among social researchers in recent years because of its superior capability for detecting unobserved heterogeneity in data. Unlike the traditional latent variable models that focus on extracting factors, the latent variable in a LCM is discrete and categorical， and its groups, called latent classes, provide a classification structure or statistical taxonomy for recovering the sub-population behind the sample. This workshop will briefly introduce the methodological concepts of LCM and pay more attention to the analytic techniques, using Mplus, for implementing latent class analysis as well as latent profile analysis along with a major extension for longitudinal data analysis, entitled latent transition analysis (LTA), which is used to examine how individuals transition in the latent class membership over time.

Dr. Hawjeng Chiou is a distinguished professor of the College of Management at National Taiwan Normal University. His major interest is in applied psychometrics, particularly for the applications of advanced modeling techniques, such as Structural Equation Modeling, Multilevel Linear Modeling, and Latent Class Modeling, with substantial areas such as Human Resource Management, Organizational Behavior, and Psychological Testing as well as Creativity research. He has published multiple books on a variety of topics.

Workshop 6: Bayesian Longitudinal Data Modeling
Instructor: Xin Cynthia Tong -- University of Virginia

Time: 1pm-4pm, July 3, China Time / 1am-4am, July 3, US Eastern Time
Location： In person at Fudan University and on Zoom (Register to get the link)

Modeling longitudinal data is one of the most active areas of research in social, behavioral, and education sciences because longitudinal data can provide valuable insights into change and causal relationships. The application of Bayesian methods in longitudinal research has gained increasing popularity. This interactive workshop focuses on Bayesian methods in analyzing longitudinal data. Particularly, it will cover topics on growth mixture modeling, missing data analysis, and Bayesian model assessment. Concrete examples will be provided to illustrate how to compute, report, and interpret Bayesian modeling results with empirical psychological data.

[This workshop is supported by Supported by the William K. and Katherine W. Estes Fund that is jointly overseen by the Association for Psychological Science and the Psychonomics Society..]

Dr. Tong is an associate professor at the University of Virginia. Her research focuses on developing and applying statistical methods in the areas of developmental and health studies. Methodologically, she is interested in Bayesian methodology, growth curve modeling, and robust structural equation modeling with nonnormal and missing data. Substantively, she is interested in analyzing the longitudinal development of cognitive ability and achievement skills.

Invited Talks

Track: General Papers [Regular, Online]
User Priority based Task Allocation and Predictive Scheduling Model for QoS enhancement in Cloud-IoT

Muthusenthil Balasubramanian*; SRM Valliammai Engineering College Department of CSE SRM Nagar Chennai; IN
Devi Karuppiah; SRM Valliammai Engineering College Department of CSE SRM Nagar Chennai; IN

Based on IoT data and cloud computing resources, Cloud-IoT provides multiple applications with on-demand services. However, the quality of service suffers as a result of task-virtual machine (VM) mismatches caused by heterogeneous service requests from IoT devices. A User Priority based Task Allocation and Predictive Scheduling Model (UPTA-PS) model is proposed to increase the quality of services offered by cloud-based IoT in organizations. UPTA algorithm is used to prioritize users as well as tasks and Modified-Teaching Learning Based Optimization (M-TLBO) algorithm is to ensure tasks are assigned to VM in order to minimize expected completion times. We have demonstrated a 95% throughput for the UPTA and M-TLBO algorithms with a larger dataset and large number of virtual machines. Furthermore, our approach achieved a make-span of less than 50%. Additionally, the average reliability and deadline violation rate were higher. In the experiments, the proposed approach clearly outperforms other existing approaches.

Track: Data Science in Humanities and Social Sciences [Regular, In-person]
From Narratives to Data: How Data of Social Networks and Criminal Records Change Paradigms of Early China

Liang CAI*; University of Notre Dame; US

This paper aims to achieve two goals. First, I will demonstrate how the structured information on historical figures of early Chinese empires can provide new perspectives in scrutinizing textual narratives. Second, I will contend that systematically extracted data can shift our arguments from analyzing examples to relating a comprehensive history that counts for every recorded case. As a continuation of traditional inquiry with updated tools and technologies, transforming textual sources into data helps revise conventional paradigms.

The Grand Scribe's Records, The History of Western Han, and The History of Eastern Han, are the most consulted historical sources of Qin-Han Empires, the fountainhead of Chinese civilizations. They record the stories of thousands of important historical figures and significant events, which can be transformed into structured biographical and social network data.

In the first case study, I will show that networks based on teacher-disciple relationships served as a social structure channeling political power to Confucians, helping them bypass the hierarchical

ladder of the bureaucracy. This newly constructed data unfold a fundamental historical change that took place when Confucians realized their dream, becoming teachers to the throne. Thereafter, Confucians as "teacher to the throne" became a widely acknowledged political discourse and an entrenched tradition that persisted for almost two millennia until the early twentieth century. The second case study focuses on reconstructed data of criminal records of officials in the Western Han dynasty. I will demonstrate how an astonishing percentage of convicted officials shed light on Confucians' intense criticism of rule by law. Data on criminalizing officials further help us explore the unique understanding of law, and its relationship with the state in Chinese history.

Finally, I will ponder upon the challenge of publishing and re-using the "handcrafted" datasets for broader audiences. I will also briefly discuss the preliminary research result of our project "Digital Empires" regarding how AI (artificial intellectual) could contribute to mine data from narratives and revolutionize the way primary sources are structured, visualized, and extracted.

Track: Data Science Methods in Political Science and Public Policy Research [Speed, In-person]

Parasocial but Meaningful: Exposure to Foreign Cultures Encourages Pro-Outsider Attitudes

Gong Chen*; Fudan University; CN

This article investigates how exposure to foreign cultures affects individual attitudes toward foreign outgroups. According to parasocial contact theory, indirect contact with outgroup members on screen reduces intergroup bias, but this effect is conditioned by the media content and producer. I argue that the consumption of foreign-made cultural products (e.g., TV programs and movies) plays a better role in ameliorating outgroup attitudes. First, group-specific exposure to a foreign culture is associated with decategorization. It deemphasizes group-based differences and strengthens individualized parasocial relationships with the contacted media character. Second, generalized exposure to diverse foreign cultures contributes to recategorization through which a more inclusive supranational identity is constructed beyond national identity, making ingroup members hold favorable opinions of not only a foreign country but immigration and globalization more generally. Analyses of the AsiaBarometer Survey, East Asian Social Survey, and Asian Student Survey lend empirical support to both types of cultural exposure. Using cable TV ownership as an instrument, the basic findings are robust to a set of sensitivity tests. Echoing the psychological microfoundations of constructivism in international relations, a collective "we" identity based on interstate contact lays the mass attitudinal groundwork for international cooperation.

Track: General Papers [Speed, In-person]

The Dynamic Impact Between Intraindividual Perceived Stress and Perceptual Speed in Young and Old Individuals: Evidence from an Intensive Longitudinal Study

Xiaoyu Chen*; Beijing normal university; CN
Hongyun Liu; Beijing normal university; CN

Many studies have shown that perceived stress influences cognitive ability, the magnitude of effects might vary for different dimensions of perceived stress, and the effects might be moderated by age. However, the dynamic relationships between intraindividual perceived stress and cognitive ability remain insufficiently documented. In this article, we used residual dynamic structural equation modeling (RDSEM) to investigate the reciprocal relationships between intraindividual perceived stress fluctuation and intraindividual perceptual speed fluctuation as well as the difference between young and old individuals. Derived from the COGITO Study, our data involved 204 individuals (101 young people and 103 old people), and each of them was measured repeatedly over approximately 100 days. The results showed that there was a stable carryover effect of intraindividual perceived stress, and this effect was higher in young adults than in old adults. The effects between perceived stress and perceptual speed were unidirectional. For old adults, perceived helplessness (a dimension of perceived stress) had a negative effect on perceptual speed, while perceived self-efficacy (a dimension of perceived stress) showed a positive effect on perceptual speed. However, this was not the case in young groups. During the 100 days, perceptual speed task performance showed an increasing trend. These findings extended the relationship between perceived stress and cognitive ability to the intraindividual level and reflected the sensibility of perceived stress on perceptual speed in older adults.

Track: Data Science in Humanities and Social Sciences [Speed, In-person]
Factors Affecting Reading Growth in Language Minority Learners

Yanran Chen*; University of Notre Dame; US

Language minority (LM) students come from families where a language other than English is spoken. This group encompasses children from a wide spectrum of background and with varying language abilities at the time of kindergarten entry; their participation in language supporting programs also varies. These differences are highly predictive of future growth in language minority students' reading skills but not yet extensively studied in previous research. Taking advantage of a nationally representative dataset in the U.S. (Early Childhood Longitudinal Study - Kindergarten cohort of 2010-2011), we examined LM students' reading growth trajectories from kindergarten through 5th grade. This study employed recursive partitioning in a mixed-effect longitudinal framework to detect relationships between characteristic variables and reading score. It incorporated various levels of interactions among variables to provide more accurate and sensible explanations to growth patterns. Results suggested that a student's score in the screening test administered upon enrollment in kindergarten is the most influential factor in predicting his/her future growth pattern in reading. Socioeconomic status is the second most important variable, while different types of language supporting program do not produce a notable difference in projected trajectories. Students who have a disadvantaged start also suffer from slower improvement over time, emphasizing the importance of early cultivation of literacy and targeted assistance for students from disadvantaged families.

Track: General Papers [Speed, In-person]
A Two-Step Method based on lz* for Identifying Effortful Respondents

Yilan Chen*; Beijing Normal University; CN
Hongyun Liu; Beijing Normal University; CN
Yue Liu; Sichuan Normal University; CN

The person-fit statistic lz* based on likelihood is commonly used to distinguish effortful respondents from those without effort in educational assessment. It impossibly circumvents item parameter estimation, though correcting the original index lz* in terms of the sampling variability to relax the distributional assumption. Biased estimates resulting from data with non-effortful responses undermine the strength of lz*. In this study, we propose a two-step method to improve lz* with the help of data mining through estimating item parameters more accurately. The results suggest that K-means affording less biased item parameter estimates facilitates lz* with regard to the precision of identifying effortful respondents when non-effortful responding is serious.

Track: Data Science in Humanities and Social Sciences [Speed, Online]
Can testing/training dataset separation ensure the reproducibility?

Yongtian Cheng*; University College London;
Xijuan Zhang; York university; CA

In recent years, machine learning techniques have been viewed as a potential solution to solve the replication crisis in psychology. Commonly, a study using machine learning separates the dataset into three parts: training, validation, and testing. A model provided by the machine learning method is usually fitted by the training dataset, modified by the validation dataset, and its performance is evaluated by the testing dataset.

While the training and validation datasets can sometimes exchange with each other, the testing dataset is always isolated from the training and validation datasets. Because the testing dataset is independent of the training and validation datasets, the model selected by machine learning techniques is viewed to have higher performance and reproducibility than that derived using traditional model selection techniques.

However, to our knowledge, no empirical evidence can be used to support that using training/testing separation can ensure the reproducibility of the performance of the model. In order to fill this research gap, this study used Monte Carlo simulation to evaluate a neural network, a popular machine learning method, which can report it found a reasonable pattern in both the validation dataset and testing dataset from a population with a null effect.

In this study, the population with a null effect is simulated by assigning random numbers to the dependent variable, which is the outcome of the neutral network. As the outcome of the neutral network is totally random and has no relationship with the input, the neural network should not find any reasonable pattern. In this study, we set a cut-off as more than 5% of the variance of the validation dataset and testing dataset can be explained. In the case that more than 5% of the

variance can be explained for both validation and testing datasets, we believe this is a model with questionable reproducibility, as this kind of performance cannot be found in the population.

We simulated both normal and non-normal data. We varied the sample size (n=500, 200, 100, 50, or 20) and tolerance of early stopping in the neutral network (5, 10, or 20) with 1000 replication times. We found that it is very unlikely that a neural network can reach a performance of with 5% variance explained from a population with a null effect. In all replication times in all conditions, none of the neural network models provide questionable reproducibility.

Our findings provided empirical evidence that a good performance of neural networks is very unlikely to come from a population with null effect. This study also provides empirical evidence that the training/testing separation procedure can improve the reproducibility of a study.

Track: General Papers [Speed, In-person]
When Predictors Sum to a Constant: Trade-off Effect Analysis Using a Regression Model Based on Isometric Log-ratio Transformation

Jieyuan Dong*; Beijing Normal University; CN
Hongyun Liu; Beijing Normal University; CN

The standard regression model is not feasible when the sum of the independent variables is constant, which is often the case for proportional data or ipsative data. Davison et al. (2022) described a set of reduced-rank regression models in which each regression coefficient can be interpreted as the effect of a predictor trade-off, and applied their method to the frequencies of four response categories in a multiple-choice test. However, the compositional nature of the predictors should not be neglected. Even though one is dropped from the model, the remaining predictors are still not an unconstrained vector but a subcomposition in the simplex where regular statistical operations in real space are not permissible, and the risk of multicollinearity still exists. In this study, from the perspective of compositional data, we transform the predictors into the isometric log-ratio (ILR) coordinates with the help of a planned sequential binary partition process, and build the ILR-transformation-based regression model. Then we come up with an estimation approach to trade-off effects on the basis of the improved model. Instead of directly relying on regression coefficients, the point estimate of a trade-off effect is determined by the difference in the dependent variable before and after the trade-off, from which the 95% confidence interval can be derived as well. Re-analyzing the response categories with the proposed method yields a better regression model and more reasonable estimation results of trade-off effects featuring the flexibility that the effect sizes of pairwise trade-offs do not have to be equal. A new empirical example from the section of learning strategies in PISA2012 is also studied to verify the validity and usefulness of the method in diverse situations like the forced-choice items and the presence of covariates.

Track: Data Science in Humanities and Social Sciences [Regular, Online]
Fuzzy Feature Selection (FFS) Based on Different Experts' Judgments: A case Study Epilepsy's Disease

Hojjatollah Farahani*; Tarbiat Modares University; IR
Parviz Azadfallah; Tarbiat Modares University; IR
Peter Watson; Cognition and Brain Unit, University of Cambridge; GB
Marija Blagojević; University of Kragujevac, Faculty of Technical Sciences Čačak; CS
Farzin Bagheri Sheykhangafshe; Tarbiat Modares University; IR
Mehrnaz SalehiRaj; Tarbiat Modares University; IR

Feature selection is a method for removing redundant and noisy variables in cognitive psychology. There are already some methods for doing this which use approaches such as genetic algorithms and random forest models. The purpose of this paper is to introduce the use of fuzzy feature selection methods using R software. The method in this paper selects the most important variables influencing prediction of a target variable by quantifying the inherent uncertainty expressed by clinicians in linguistic terms.

Method: This method can be used with self-reported variables and experts' opinions in the field or combinations of them. The first step is to measure the features, the second step is compute the mean of the quantitative measures, the third step is to fuzzify the means based on triangular fuzzy numbers (TFNs) which are defined based on experts' opinion. The fuzzy numbers are then analysed using the Fuzzy WASPAS Method. This method identifies the most important features and their weights. The FuzzyMCDM package in R is used. As a case study in this paper, 7 features (age, sleep, diet, family support, emotion regulation, history of trauma, physical exercise) which have an impact on mild cognitive impairment were measured based on 3 criteria including the opinions of 6 psychiatrists', 5 psychologists and 3 neurologists. The data were collected by structured interview and they were measured based on 5-point Likert scales (1=very low,2 =low, middle=3, high=4 and very high=5). The means of the features were defined as triangular fuzzy numbers and using the fuzzy WASPAS method the weights and importance ratio of the features were computed. Each feature was either if the i-th criterion was a benefit or if the i-th criterion was a cost. Lambda was 0.5.

Results: The results indicated that age, diet, emotion regulation, family support, physical exercise, history of trauma and hours of sleep were the most important features, in that order, with variables given higher weights having a larger impact on mild cognitive impairment in epilepsy's disease.

Conclusion: Using this Fuzzy feature selection method we can combine information from different types of qualitative (linguistic terms) and quantitative data when selecting the most important features. This method can be regarded, therefore, as a semi qualitative feature selection method.

Track: General Papers [Speed, In-person]
How does prior distribution affect model fit indices of BSEM

Yonglin Feng; Department of Psychology, Sun Yat-sen University; CN
Junhao Pan*; Department of Psychology, Sun Yat-sen University; CN

Bayesian structural equation model (BSEM) integrates the advantages of Bayesian method into the framework of structural equation modeling and ensures the identification by assigning priors with small variances. Previous studies have revealed that prior specifications in BSEM influence model parameter estimation, but the impact on model fit indices remains unclear and needs to be explored. Therefore, two simulation studies were conducted. Normal distribution priors were specified for factor loadings, while inverse Wishart distribution priors and separation strategy priors were applied for the variance-covariance matrix of latent factors. Conditions included five sample sizes and 24 prior distribution settings. Simulation study 1 compared the model fitting performance of BCFI, BTLI, and BRMSEA proposed by Garnier-Villarreal and Jorgensen (2020), and PP*p* value. Simulation study 2 compared the performance of BCFI, BTLI, BRMSEA, and DIC in model selection between three data generation models and three fitting models. Results showed that prior settings would affect Bayesian model fit indices in evaluating model fitting and selecting models, especially in small sample sizes. Even under large sample size, the highly improper factor loading prior led to poor performance of the Bayesian model fit indices. BCFI and BTLI were less likely to reject the correct model than BRMSEA and PP*p* values under different prior specifications. For model selection, different prior settings would affect DIC on selecting the wrong model, and BRMSEA preferred the parsimonious model. Our findings recommended that the Bayesian approximate fit indices may be better for evaluating model fitting and selecting models under the BSEM framework.

Track: General Papers [Speed, In-person]

Improvements to Regularization Differential Item Functioning Detection Methods: Combining Regularization with Traditional Methods

Kaiyu Gan; Beijing Normal University; CN
Hongyun Liu*; Beijing Normal University; CN

In the measurement research of education and psychology, the test fairness is an important issue. Differential Item Functioning (DIF) is one of the key indicators of test fairness and validity. A variety of methods to detect DIF have been developed. In recent years, with the rise of machine learning technique, the regularization method has been applied to DIF detection and worked well. However, the regularized DIF method has the problem of insufficient interpretability. Researchers cannot easily interpret the meaning of DIF parameters in regularized methods. We propose a method that combines regularization method with traditional method. That is, we use the regularization to select anchor items and use the Mantel–Haenszel (MH) method to identify DIF items. In addition, we tried different anchor selection methods. We conducted simulation and empirical study to compare the performance of the suggested method to the regularization method and MH method (in terms of Type I error and power). The results show that with the increase of sample size and amounts of DIF, the suggest method has better control of Type I error than the MH method. Also, the suggest method has higher power than the regularization method.

Track: Data Science in Humanities and Social Sciences [Regular, In-person]
Psychoperiscope: A data ingestion tool and modeling scale for social and behavioral data science

Joshua Gandi*; Department of Psychology University of Jos, Nigeria; NG

Psychometrics could be explained as the science of psychological measurement which focuses on the latent variables and constructs under consideration, the measurement instruments, the measurement procedures, and the quality of both instruments and procedures. The fact that data science is the process of using advanced analytics to extract valuable information by executing data ingestion, data storage and processing, data analysis, and insights communication and data visualizations consequently makes the insights easily implementable. Integration of the science of psychological measurement and the data science heralded the development of psychoperiscope as a coined nomenclature from psychometrics and periscope. It was constructed, developed and being appropriately standardized. Psychoperiscope has been particularly designed as a 3-version modeling scale which mirrors the mediating-moderating effects of cognitive coping strategies in the relationship between illness and quality of life. Thus, it typically exemplifies the application of data science in the social and behavioral sciences. This 3-version scale proves useful for descriptive analysis, diagnostic analysis, predictive analysis, and prescriptive analysis that are essential functions of data science and psychometric modeling.

Track: Data Science in Humanities and Social Sciences [Speed, In-person]
Ordinal Logistic Regression Model in Determining Factors Associated with Household Food Insecurity in Namibia

Dibaba Bayisa Gemechu*; Namibia University of Science and Technology (NUST), Windhoek; NA
Leonard Elifas; Namibia University of Science and Technology (NUST), Windhoek; NA

Food insecurity is a global issue, and households in a society can experience food insecurity at different levels that could range from being food secure to severely food insecure. In Namibia, the household food insecurity is still a threat to the nation. The severity of food insecurity is an ordinal categorical variable in nature and different types of ordinal logistic regression models could be used to model such variables. The purpose of this study was to identify the socioeconomic and demographic factors associated with household food insecurity in Namibia by fitting an ordinal logistic regression model using the 2015/2016 Namibia Household Income and Expenditure Survey. The proportional odds model (POM) and the partial proportional odds model (PPOM) were fitted and the performance of the two models was also compared. The PPOM was found to be the better model and based on the PPOM result, the study found factors such as the age of the household head, the household size, the source of income of a household, the annual income of the household, the education level attained by a household head and the geographical location of a household to be significant factors associated with severity of household food insecurity in Namibia.

Track: General Papers [Regular, In-person]
Estimation of functional-coefficient autoregressive models with measurement error

Pei Geng*; Illinois State University; US

The functional-coefficient autoregressive (FAR) models are flexible to fit nonlinear features in time series data with covariates. Cai, Fan and Yao (2000) developed an effective local linear estimation procedure under the FAR framework. When the time series data are observed with measurement error, we first derive the asymptotic bias of the naive local linear estimator (LLE) by ignoring the measurement error. Then, we propose a bias-corrected local linear estimation procedure for both the functional coefficients and the autoregressive error variance. Through simulation study, we present that the naive LLE is biased while the proposed estimation method shows superior performance with much reduced bias under various choices of FAR model settings. Furthermore, sensitivity analysis shows the robustness of the proposed estimator under a chosen misspecified measurement error model. The asymptotic properties of the bias-corrected estimators are also established. At last, the proposed approach is applied to a cybersecurity real data example.

Track: Data Science in Humanities and Social Sciences [Speed, In-person]
Chinese Adolescents' Career Choice in Large-Scale International Studies: Trends and Methodologies

Luyang Guo*; the Chinese University of Hong Kong; CN
Kit-Tai HAU; The Chinese University of Hong Kong; CN

Sustainable economic development relies on having an adequate workforce supply in vital professions that drive individual growth and national prosperity. During adolescence, young people begin to grasp the complex nature of career choices, and their career expectations can significantly impact their educational and occupational attainment. Cultural and social contexts also influence career expectations, making it crucial to understand the popularity of elite professions among young people in China and attract prospective candidates for the future.

The Programme for International Student Assessment (PISA) is a large-scale triennial project that measures 15-year-old students' cognitive abilities, career interests, and personal background factors. Using PISA data from 2000 to 2018, linear and multilevel multinomial logistic regressions showed that Chinese girls increasingly aspired to teaching careers, while more Chinese boys preferred business and administration. In contrast, science and medical careers did not receive enough attention compared to many other high-performing economies.

From a country-level perspective, hierarchical linear regressions revealed that financial expenditure, working conditions, and societal evaluations of professions played a crucial role in shaping adolescents' career choices. However, the effects differed for various professions. The relative importance of academic ability, socioeconomic status, psychological attributes, and

environmental support was also examined using multiple methodologies appropriate to the relevant research questions.

The implications of these results were discussed in relation to the demand for skilled workers in the technological digital era. It is essential to prioritize the cultivation of talent in the fields that will drive economic growth, and this can be achieved by understanding the factors that shape young people's career aspirations and working to attract and retain top talent in these vital professions.

Track: General Papers [Regular, In-person]
University Major Decision and its Effect on Wage: Modeling Interaction between Major Specificity and Education-Job Relevance Using Machine Learning Approach

Tian Hang*; Holly Futures, Co.; CN
Karl Ho; University of Texas at Dallas; US
Yong Zhou; Jiangsu SOHO Holdings Group CO., Ltd.; CN

This study investigates university students' choice of majors and such effect on wage in Chinese labor market. Major specificity, measured by the graduates' concentration in occupation selection, indicates the specificity of the knowledge and skills each major bears. This research explores the interaction relationship between major specificity and education-job relevance in the wage function. I hypothesize that choosing education-job relevance and major specificity have a positive correlation with one's starting wage.

Employing machine learning techniques, I will take advantage of Lasso model, KNN fit and random forest models for variable selection and measuring feature importance. The interaction relationship is measured by the adjusted Friedman's H^2 statistic. Network plots and heatmap with the leaf sort algorithm are used to visualize interaction relationship (Earle et al., 2015; Inglis et al., 2022).

By this research, three goals are achieved: the first, major specificity receives validation; the second, the interaction between major specificity and education-job relevance is studied and visualized; the third, the result is instructive and informative to both government and individuals.

References

Earle, D., & Hurley, C. B. (2015). Advances in dendrogram seriation for application to visualization. *Journal of Computational and Graphical Statistics*, *24*(1), 1-25.
Inglis, A., Parnell, A., & Hurley, C. B. (2022). Visualizing variable importance and variable interaction effects in machine learning models. *Journal of Computational and Graphical Statistics*, *31*(3), 766-778.

Track: General Papers [Regular, In-person]
Large-Scale International Education Surveys: Analyses of Slopes Can be More Interesting

Than Comparisons of Means

Kit-Tai HAU*; The Chinese University of Hong Kong; HK
Matthias von Davier; Lynch School of Education and Human Development, Boston College; US

Large-scale international education surveys, such as the Programme for International Student Assessment (PISA) and Trends in Mathematics and Science Achievement (TIMSS), have gained popularity and influence due to their ability to compare educational strengths and weaknesses across economies. The general public, policymakers, and educational researchers typically focus on mean comparisons of cognitive tests (e.g., reading, science, mathematics in PISA) or questionnaire scales (e.g., self-esteem, anxiety; thereafter simplified as "scales") when using these surveys.

This study used PISA to demonstrate the complexities of comparing scales across different units (e.g., economies, schools). Specifically, the study examined (a) variance at the student, school, and economy levels, (b) the strength of relationships between various scales and external criteria (e.g., achievement, socioeconomic status [SES], gender) at different levels (student, school, economy), and (c) the strength of relationships among various scales at different levels.

The study found that cognitive tests and SES showed 20%-30% variance between economies and typically 20%-50% variance across schools within each economy (e.g., using the set: Australia, BSJG-China, Canada, Hong Kong-China, Japan, Singapore, UK, USA). Thus, league tables based on cognitive tests and SES may be statistically meaningful. However, for questionnaire scales (e.g., joy in science, self-efficacy, belongingness, anxiety, motivation), the variance across economies was small (2.8% - 14.8%, mean = 7.3%) and across schools within each economy was also small (1.0%-10.0%, mean = 3.8%). These small variances suggest that there are much larger differences within each school than across schools or across economies.

The correlations of scales and achievement within each school are interpretable. Thus, within each school, questionnaire scales (averaged over 14,000-17,000 schools in PISA; joy .18-.24; efficacy .12-.17; belongingness .06-.08; anxiety -.08 to -.10; motivation, .15 - .16) were correlated with achievement (math, reading, science) in the same direction and magnitude as reported in the literature. When aggregated to the school and economy level, reliabilities and correlations were expected to be attenuated as measurement errors were reduced. Importantly and most worrisome was that joy (-.49 to -.54), efficacy (-.12 to -.20), and motivation (-.35 to -.43) were negatively related to achievement in contradiction to what the literature would predict (positive relation with desirable scales).

In conclusion, while league tables based on cognitive tests and SES may be statistically meaningful and interpretable, those based on questionnaire scales have small variances across schools and economies and are susceptible to frame (Marsh & Hau, 2003) and other aptitude-achievement paradox phenomena (Downey et al., 2009). It is important to note that the interpretation of means at the school and economy levels may differ from that at the student level. Thus, examining the relationships between scale scores and achievement, SES, and gender

within each school may provide richer information and meaning consistent with traditional literature.

Downey, D. B., Ainsworth, J. W., Qian Z. (2009). Rethinking the attitude-achievement paradox among blacks. Sociology of Education, 82(Jan), 1-19.

Marsh, H. W., & Hau, K. T. (2003). Big-fish-little-pond effect on academic self-concept: A cross-cultural (26 country) test of the negative effects of academically selective schools. *American Psychologist, 58*, 364-376.

Track: Data Science in Humanities and Social Sciences [Speed, Online]
On the Relationship between Factor Loadings and Component Loadings when Latent Traits and Specificities are Treated as Latent Factors

Kentaro Hayashi*; University of Hawaii at Manoa; US
Ke-Hai Yuan; University of Notre Dame; US

Most existing studies on the relationship between factor analysis (FA) and principal component analysis (PCA) focus on the relationship between common factors and first few components as well as their loadings. Based on a setup in Bentler and de Leeuw (2011), this study examines the relationship between FA loadings and PC loadings when specificities are treated as latent factors. In particular, we will examine the closeness between the two types of loadings when the number of observed variables (p) increases. Parallel to the development in Liang et al. (2015), average squared canonical correlation (ascc) is used as the criterion for measuring the closeness. Results show that the ascc between the two sets of loadings approaches 1.0 as p approaches infinity. The study gives a different perspective on the relationship of PCA and FA, and the results add additional insights on the selection of the two types of methods in the analysis of high dimensional data.

Track: Data Science in Humanities and Social Sciences [Regular, Online]
Universal Kriging using High Performance Tools on a Laptop

Erin Hodgess*; Western Governors University; US

Universal kriging is an excellent tool for spatial data analysis. The underlying assumptions of this model are eminently suitable for real-world applications. There are currently existing functions for universal kriging in R. However, since the model requires inversion of large matrices, these functions can be limited by data size and speed. We have developed a package for universal kriging which uses the Windows Subsystem for Linux, along with NVIDIA graphics cards. We can exploit the CUDA architecture on the NVIDIA cards. We can complete our matrix inversion via CUDA Fortran. We will discuss the speed up and increased data size that CUDA provides. We will see that the high performance tools can be extremely useful for large spatial data sets.

Track: General Papers [Regular, In-person]
Multiverse-style analysis: Current bibliometric analysis and tutorial

Shunsen Huang; State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University; CN
Xinmei Zhao; State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University; CN
Xiaoxiong Lai; State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University; CN
Xinran Dai; State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University; CN
Huanlei Wang; State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University; CN
Yun Wang*; Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, No. 19, XinJieKouWai St., HaiDian District, Beijing 100875, P. R. China; CN

Aim: Multiverse-style analysis (MSA), a method aiming to report combinations of multiple analysis strategies during data analysis and to test the robustness of interested effects in all analytic strategies, has gained an increasing popularity in social science. This study aims to conduct the bibliometric analysis for MSA and provide a three-step analytic tutorial to promote its development and applications in various aspects of Chinese social science. We focused on four main bibliometric issues: (1) what is the annual scientific publication growth in MSA; (2) which authors or articles are the most influential; (3) what the contribution distribution across different countries; and what is its development condition in China; (4) what are the most important topics researchers care about.

Methods: Firstly, literature relevant to MSA from 2015 to 2022 were searched in *Web of Science* database using keywords including "multiverse analysis", "multimodel analysis", "specification curve analysis", "vibriation of effects". Articles whose abstract and title containing one of these keywords were kept, and articles which directly cited the four articles firstly introducing this method (Patel et al., 2015; Steegen et al., 2016; Young & Holsteen, 2017; Simonsohn et al., 2020) were also included. Then, the replicated literature was dropped. Second, all the defined literature (605 articles) was analyzed in R package *Bibliometrix.* Third, the three-step approaches of MSA were introduced as a tutorial.

Results: The bibliometric analysis indicated that the publications related to MSA have rapidly increased, from 6 articles in 2015 to 176 articles in 2022, with an annual growth rate is 62.04%. The top four subjects are psychology-multidisciplinary, neuroscience, public environment & occupational health, sociology. The most cited article across *Web of Science* is the work of Munafo et al. (2017) with 1288 citations, and the most cited article within the identified literature is the work of Steegen (2016) with 450 citations. The most productive countries are USA (224 articles), followed by United Kingdom (76 articles), Germany (71 articles), and Netherland (51 articles). China ranks 12 with 7 articles. The conceptual structure map related to MSA identified four main topics. The first topic is related to mental health, adolescent, causality, fMRI, covid-

19, etc.; the second topic contains robustness, replication, meta-science, etc.; the third topic involves reproducibility, replicability, open-science, transparency, etc.; and the fourth topic includes replication crisis, bias, p-hacking, etc. The MSA tutorial contains: (1) Identify all theoretically and empirically possible analytic combinations, (2) Estimate and describe the effects of all the identified combinations, (3) Statistical inferences.

Discussion: MSA has rapidly developed and been widely applied in social science, with North America and Europe being the main contributors. China's contribution in this field is seriously mismatched with the international status of the country's comprehensive scientific and technological strength. Chinese research in this field is still in its early stages, future development and applications are needed. Improving the robustness and transparency of application research, advocating open science, alleviating replication crisis in social science, and using MSA to guide practice are the most important concerns in this field.

Track: General Papers [Speed, In-person]
The Impact of Machine Learning Approach In Recruitment Process with the use of NLP Algorithm

Sandra Ihemeje*; ISDSA Nigeria; NG

Machine learning technology has numerous applications that give rise to what is known as "intelligent HR", which in reality is data-driven HR for organizations that wish to use these skills in the Human Resource Department with the current environment.

This paper provides a Machine Learning approach to the recruitment process, such as automation tools for Candidate screening, use of NLP algorithm approach for Candidate engagement and Workforce planning. It is not the case that Natural Language Processing systems or Artificial Intelligence systems can replace HR but These systems increase the power of HR Managers within their organization by enabling them to make a far better data-driven decision as they oversee Human Resources. the fact that the flood of resumes can be onerous and time-consuming and the act of sorting through the pile can be really tiresome if candidate resumes are manually screened.

However, this paper will first outline the ideal procedures for gathering data through web scraping to automate the extraction of data from recruiting websites in order to analyze the job market, understand candidate qualifications, review candidates resumes using data scraping and text mining techniques to extract information based on the matching skillset. The Natural Language Processing algorithm describes how well the candidates fit based on their education, work expertise, skills and other alternative requirements.

This research methods consisted of data gathering aiming to label the observations and prepare the datasets for supervised machine learning algorithms' describing how organisations can use different algorithms to get the best accuracy model, algorithm such as Support Vector Machine(SVM), K-Nearest neighbor(KNN), Random Forest, XGBoost Classifier, Artificial Neural Network.

The final aim of the paper provide an end-to-end solution for People's Management with high performance and reduced costs. provide techniques using Natural Language Processing that could help the Human Resources department or People Analytics apply sophisticated data science and machine learning to help manage their people practically and efficiently.

Track: General Papers [Regular, In-person]
Implementation of Gradient Boosting for Survival Analysis with Competing Risk

Md Hasinur Khan*; ISRT, University of Dhaka; BD

For analyzing high-dimensional time-to-event data with competing risk, sophisticated methods are required that consider censoring and the event of interest. The sparsity associated with high-dimensional data introduces even more complexity. For lowdimensional data subdistribution hazard model or cause specific hazard model has been proposed. For high dimensional data, gradient boosting decision tree (GBDT) is a powerful ensemble machine learning method that has potential to identify risk factors associated with micro-array data. with recent algorithmic advances such as Extreme Gradient Boosting (XGB) and Light Gradient Boosting (LGB), GBDT training become more efficient and scalable. In this study we proposed a method to analyse high-dimensional competing risk data with this modern ensemble machine learning techniques. We performed several simulation to judged the performance of our proposed method and compared them with an existing boosting method CoxBoost (Binder et al., 2009). We assessed the models based on their discriminating ability. Gradient boosting methods produced equivalent result almost in every scenarios while being more efficient. But in some scenario, CoxBoost produced more accurate results with a small margin. Finally we applied our method to a publicly avaialable data of bladder cancer and found that there LightGBM outperformed CoxBoost and XGBoost. Hence, we can get almost the same performance as CoxBoost or even better with gradient boosting and we would not need a huge amount of time. That surely will benefit the researchers, scientists, decision maker etc. in all fields where we have to deal with competing risk and censored observations.

Track: General Papers [Speed, In-person]
Rationality Evaluation and Continuous Improvement of the Curriculum System Using Factor Analysis

Yi Kuang*; Xiangtan University; CN

Outcome-based, student-centered, and continuous improvement are the three core concepts of engineering education professional certification. The Washington Accord asserts that the design and continuous improvement of curriculum systems are essential to achieve undergraduate students' graduate attributes, and training goals, which helps in nurturing recognized internationally competent engineers. Traditionally, the design of the "course-graduate attributes "relationship matrix in the curriculum system relied primarily on expert intuition, lacking models and measurement methods. To address this issue, we proposed a continuous improvement method for the curriculum system based on factor analysis to provide a more scientific, objective, and reasonable curriculum system. The graduate attributes entail 12 skills, out of

which five are technical requirements, including engineering knowledge, problem analysis, design/development solutions, research, and the use of modern tools. In this study, we employed factor analysis to refine the matrix of curricula and the five latent technical skills. We collected exam score data from 32 curricula of the Communication Engineering major in the 2018th and 2019th classes. To address the missing data, we generated ten imputed datasets using a Markov Chain Monte Carlo method. Next, we used confirmatory factor analysis to construct the original relationship matrix model, which had a poor fit (RMSEA=0.126). Therefore, we performed exploratory factor analysis using ordinary least squares estimation and CF-varimax rotation on each imputed dataset to extract factors. The RMSEA is 0.008, indicating that the model fits the data well. The factor loadings can be used to refine the "course-graduate attributes" relationship matrix for the five technical graduate attributes and continuously improve the curriculum system.

Track: General Papers [Regular, In-person]

A hybrid method: solve the impact of variable ordering in Bayesian network structure learning

Minglan Li*; Beijing Normal University; CN
Fandi Chen; Hangzhou Mingshitang Digital Technology Co., LTD; CN
Mengwei Wu; Hangzhou Mingshitang Digital Technology Co., LTD; CN
Yueqin Hu; Beijing Normal University; CN

In recent years, the rise of artificial intelligence and the development of machine learning have brought new analytical methods to theoretical research. Bayesian networks are an important tool for uncertainty inference and can be applied to many complex causal systems. It has been applied to automation, biology, medicine and other fields. In psychology, it is often used in risk prediction. A Bayesian network is a probabilistic graph model composed of directed acyclic graph (DAG) where the nodes represent the variables we choose. Typically, DAG are obtained by machine learning after the nodes have been identified by domain experts. However, in machine learning, the order of the input variables can remarkably affect the result (Kitson et al.,2022). Especially when there are few variables, the order can even directly determine the structure of the Bayesian network. However, existing research has paid no attention to this issue. This study proposed a hybrid approach to help identify the optimal order of variables, thus improving the stability and accuracy of the Bayesian network structure. This method randomizes the variables orders and then selects the best fitting order under certain theoretical constraints. In addition, in psychological research, many variables have a bi-directional relationship, while DAG only allows for one-direction relationships. Thus, this study also proposed to combine the above method with the partially directed acyclic graphs (CPDAGs), in which bidirectional relationships are represented by lines. The candidate DAGs found in the first step can be further verified by CPDAGs. Different DAGs with the prediction performance may correspond to the same CPDAG. Therefore, this method can not only obtain stable results, but also helps to clarify whether the relationship is unidirectional or bidirectional. An empirical study predicting depression and aggressive behaviors in adolescents was also used to demonstrate the method proposed in this study to facilitate the appropriate application of Bayesian networks in psychological research.

Track: Data Science in Humanities and Social Sciences [Speed, In-person]
The Re-individualization of the Internet: A Comparison Analysis of Aggressiveness and Politeness in Social Media Posts Before and After the Compulsory GEO-IP Tag Policy

Yujun Li*; Jiangxi Normal University; CN
Wenzheng Lin; Jiangxi Normal University; CN
Zhihui Lai; Jiangxi Normal University; CN
Xu Feng; Jiangxi Normal University; CN

Since August 2022, Since August 2022, the majority of social media sites in mainland China have enacted the policy to tag along the geographical IP information of the users when they post their comments publicly. This policy offers a great opportunity to test the deindividualization process in social psychology. Theories of deindividualization posit that people would act in antinormative manners when their self-identity is deprived. Those theories have been applied to explain the aggressive behaviors online. As the revealing of the personal information by GEO-IP policy exposes one's self-identity to a degree, we hypothesized that it can "re-individualize" people to behave in a more normative manner. To examine this assumption, we planned to compare the degree of aggressiveness and politeness of Internet Speech before and after the enaction of the policy by collecting and analyzing public posts through Web crawler. the majority of SNSs in mainland China have enacted the policy to tag along the geographical IP information of the users when they post their comments publicly. This policy offers a great opportunity to test the deindividualization process in social psychology. Theories of deindividualization posit that people would act in antinormative manners when their self-identity is deprived. Those theories have been applied to explain the aggressive behaviors online. As the revealing of the personal information by GEO-IP policy exposes one's self-identity to a degree, we hypothesized that it can "re-individualize" people to behave in a more normative manner. To examine this assumption, we planned to compare the degree of aggressiveness and politeness of Internet Speech before and after the enaction of the policy by collecting and analyzing public posts through Web crawler.

Track: Data Science in Humanities and Social Sciences [Regular, In-person]
Using Baseline Case Characteristics to Predict Successful Treatment Completion

Xiaolan Liao*; The University of Oklahoma Health Sciences center; US
Hairong Song; The University of Oklahoma; US
David Bard; The University of Oklahoma Health Sciences center; US

Goal: Accurate prediction of treatment completion prior to assignment is crucial in prevention studies because it can reduce healthcare cost as well as identify areas of treatment that may need to be strengthened to increase adherence and successful completion. However, accurate prediction with traditional statistical models is challenging especially when the number of predictors is large. In this study, the researchers evaluated the uses of machine learning techniques in predicting future treatment completion and adopted a two-step approach that incorporated ensemble models.

Method: The data were collected from a novel intervention treatment program attempting to prevent child removal following an allegation of child abuse or neglect (N = 795). The sample in study included 40 covariates (50% binary, 50% continuous) with 30% missing values, and an imbalanced outcome (success vs. unsuccess: 78% vs 22%). The sample was randomly split into a training (80%) and a test (20%) set for cross-validation. The classification models (i.e., Logistic Regression (LR), ElasticNet, Random Forest (RF), Support Vector Machine (SVM), K-nearest neighbors (KNN), Naïve Bayes (NB), & Artificial Neural Network (ANN) were built using the training data, their performance on accuracy, recall, precision, F1-score, and Area Under Curve (AUC) were evaluated with test data. Variable selection for creating reduced models was based on feature importance. Missing data were imputed using mice package in R. All models were trained and tested using Python sk-learn module.

Results: Results showed all the models had AUC greater than .50 (ranging from .539 to .666). The RF had the largest accuracy (.805) and F1 score (.886). All the models had high recall scores (>. 926) except NB (.884). The NB had the highest precision score (.836). The performance of the ensemble model with majority vote of predictions from all models was evaluated. While reduced models were often (slightly) worse than full models, the reduced ensemble model outperformed its full model on all indices (accuracy: .811 vs .786; recall 1.00 vs .992; precision: .801 vs .784; F1: .890 vs .876; AUC: .605 vs .562).

Conclusion: Results indicated that all the machine learning models performed better than random guessing in classifying successful and unsuccessful completion. While no model unanimously performed better than the others, RF and SVM were the best ones in correctly predicting positive cases. NB was the one with the highest precision score (i.e., ratio of true positive to the sum of true positive and false positive). The reduced ensemble model had the highest scores on all indices except on precision. This study recommends using a two-step approach for variable selection and model building and model selection. Step one: perform RF to obtain feature importance and select top predictors. Step two: build selected models and an ensemble model with selected predictors and select the best model based on performance and practical concerns. Performance of the chosen deployment model with future data may be periodically and dynamically evaluated to ensure that the model in deployment still performs at an acceptable level.

Track: General Papers [Speed, In-person]

Estimation of the learning trajectories with observed and latent heterogeneity: A multi-sample study of growth mixture modelling

Pi-Fang Lin*; School of Teacher Education, National Taiwan Normal University; TW
Hawjeng Chiou; College of Management, National Taiwan Normal University; TW

Identification of superior and inferior group of academic achievement is not only the major concern of academic researchers but also the foundation of policy-making. However, due to the individual differences of learning, the estimation of the growth curve of achievement have to take the observed and latent heterogeneity into account. In this paper, the growth mixture

modeling (GMM) with multiple sample analysis was applied for identifying the growth trajectory on learning achievement. Two sources of secondary data were used: for the high school sample (sample1), a total of 2,806 students (1420 female, 1386 male) selected from the Taiwan Education Panel Survey (TEPS) was included to identify the latent clusters of superiority and inferiority in students' achievement collected in successive four consecutive times. For the undergraduate sample (sample2), a total of 1571 students (1049 female, 522 male) selected from four departments (Chinese, Mathematics, Education, Business) in one public university of Taiwan were included for estimating the learning growth trajectory. These undergraduates were entered this university starting at 2004 to 2011 with four enrollment types. Data was analyzed by Mplus 8.9 with three models for both datasets: (a) a single-class latent growth curve model, (b) an unconditional growth mixture model, and (c) multi-sample (gender groups and enrollment types) growth mixture model. Results shown that the non-linear model is the best fit for both datasets. Analyses of mixture modeling indicated several heterogeneous classes on the trajectories: For high school samples, in which the typical-growth class has gradually improved achievement curve, and the two relatively inferior groups were found by a lower starting level or a decreased trend at the latter waves of achievement measures. Learning achievement of undergraduates exist more variety due to the type of enrollment as well as the departmentization. The fan-shaped pattern of trajectories supported the phenomenon of Matthew effects for high school students, but not for undergraduates. Differences on gender groups as well as enrollment types based on the multi-sample analysis of mixture growth models were reported for both datasets. In addition to the illustrations of analysis of trajectory with heterogeneous, this study also discussed the methodological implications of applying multi-sample analysis on mixture modeling for examining the heterogeneity.

Track: General Papers [Regular, In-person]
Using response time for compromised item detection

Cheng Liu; University of Notre Dame; US
Kyung Han; Graduate Management Admission Council; US
Jun Li*; University of Notre Dame; US

Nowadays, many tests are administered and taken on computers to take advantage in the advances of information technology. Computer-based testing enables continuous testing, which is considerably more convenient to test takers compared to the traditional paper-and-pencil test. However, it also raises concerns about the test security. For instance, examinees who took the test earlier might share the encountered items with other test takers, resulting in item bank leakage and jeopardizing the test's validity and fairness. To address this challenge, this study proposes a novel compromised item detection model that utilizes response time data collected during computer-based testing. The model considers the leakage rate of different scenarios, allowing it to not only flag items that may have been compromised but also estimate the time point at which each item was compromised. The efficacy of the proposed method is investigated through both a simulation study and a real dataset. The comparison with a detection model that only uses item responses but not the response times shows that the newly proposed method maintains high detection power while improving type-I error control.

Track: Data Science in Humanities and Social Sciences [Regular, In-person]
Fit Multilevel Differential Equation Models to Intensive Longitudinal Data Using Numerical Optimization

Qingshan Liu; Beijing Normal University; CN
Yueqin Hu*; Beijing Normal University; CN

Differential equation models can be used to study univariate or multivariate dynamic systems, such as the self-regulation process of emotions, the bidirectional relationship between stress and sleep quality, and parent-child interactions. Based on intensive longitudinal data, this study proposed a one-step method for estimating the parameters of differential equation models, i.e., the model parameters were obtained by fitting differential equations directly to intensive longitudinal data using a numerical optimization method. We have also developed this method to multilevel models, where individual-level parameter estimates can be obtained. An existing dataset, Components of the Index of Consumer Sentiment data collected in the University of Michigan, was used to illustrate this method. We have also developed an R package and a Python package to facilitate the implementation of numerical methods for differential equation models.

Track: General Papers [Regular, In-person]
Causal Effect Estimation and Transportation Using Modified Bootstrap

Tao Liu*; Brown University; US

Observational data, such as electronic health records (EHR) and insurance claim data, represent one of the richest data sources for clinical research. Observational studies (OS) offer an important alternative to RCTs for studying the effect of a treatment on study subjects. Analyses of OS, however, face various challenges such as: 1) The nature of non-randomized interventions leads to confounding and selection bias; and 2) the source data population may not be representative of the research (target) population, a common problem referred to as covariate shift. We proposed a modified bootstrap method that is built on a propensity score model and a semiparametric tilting model to solve the challenges at the same time. The proposed method is computationally straightforward. We verify the validity of the method using simulations.

Track: General Papers [Speed, Online]
Detecting mediation effects with the Bayes factor: Performance evaluation and tools for sample size determination

Xiao Liu*; The University of Texas at Austin; US

Testing the presence of mediation effects is often of interest in social science research. Recently, Bayesian hypothesis testing with Bayes factors (BFs) is becoming increasingly popular. However, the use of BFs for testing mediation effects has been under-studied, despite the

growing literature on Bayesian mediation analysis. In this study, we systematically examine the performance of the BF for testing the presence versus absence of a mediation effect. The performance evaluation measures include the false- and true-positive rates. To calculate the performance evaluation measures, we extend the previous literature of sample size determination with BFs, and develop a new simulation-based approach. Our results showed that the false- and/or true-positive rates of detecting mediation with the BF can be impacted by the prior specification. The prior specification includes the prior odds of the presence of each path (treatment-mediator path, or mediator-outcome path) used in the design stage for data generation and in the analysis stage for calculating the BF of the mediation effect. To facilitate researchers determining the sample sizes for testing mediation effects with the BF, we develop easy-to-use tools (e.g., R functions and a web application). We hope that our study provides insights on the performance of the BF for testing mediation effects and adds to researchers' toolbox of sample size determination for mediation studies.

Track: Data Science in Humanities and Social Sciences [Regular, Online]
Construct Different Uncertainty Indexes to Forecast Macroeconomic Variables

Yang Liu*; Department of Economics at Rutgers University-New Brunswick; US

A number of studies in economic literature have focused on high dimensional data set to forecast macroeconomic variables. We add to the existing literature by creating some new macroeconomic uncertainty indexes and factors for forecasting. We want to examine the predictive ability of these factors and uncertainty indexes from both forecasting performances and trend indication in the graph. And we will also compare these new indexes with VXO to see their differences. In this paper, we use monthly data from FRED-MD dataset to forecast 14 target variables from eight aspects of whole economy. We use dimension reduction and shrinkage methods, such as PCA, Lasso and the combination of them, to construct new factors or predictors. There are four steps in our experiment. First, we forecast target variables using AR(SIC) model augmented with different factors or predictors. Second, we construct individual uncertainty indexes from forecasting error variance of target variables in the first step forecasting, which is inspired by Jurado, Ludvigson, and Ng (2015). Then, we construct different whole economy uncertainty indexes from the aggregation of 14 individual uncertainty indexes. Third, we do model horse racing for different AR(SIC) augmented models comparing with benchmark AR(SIC) model to examine the marginal predictive ability of all these factors and indexes. Last, we compare trend behavior of all the indexes with VXO in graphs. Both of our new factors and indexes can improve forecasting performances in all forecast horizons based on relative MSE results, especially for housing market variables. Real economy variables usually have better forecasting performances than nominal economy variables. Target variables can get the best forecasting performances when adding both factors and indexes together. And our new indexes have trend indication in short term for uncertainty based on the consistent behavior with VXO and housing market variables.

Track: General Papers [Speed, In-person]
PCA, FA, LCA, LPA, LDA, which model is the best choice for your data?

Laura Lu*; University of Georgia; US

Principal Component Analysis (PCA), Factor Analysis (FA), Latent Class Analysis (LCA), and Latent Profile Analysis (LPA) are widely used statistical models in social and behavioral sciences to identify underlying patterns and relationships in large datasets. Recently, Topic Modeling (TM) and the Latent Dirichlet Allocation (LDA) method have gained increasing popularity in these fields for analyzing text data in addition to numerical responses. All of these models aim to uncover hidden structures and dimensions within vast amounts of observed data. PCA is a method used to reduce the number of observed variables while retaining the majority of variance in the dataset through generating principal components. FA can be used to reduce the dimension of observed variables by identifying underlying factors. LCA is used to identify underlying classes in a population based on patterns of response to a set of categorical variables. LPA is used to identify latent profiles underlying patterns of response to continuous variables. TM is a statistical method used to identify the underlying themes or topics present in a large corpus of text, while LDA is a specific type of TM.

Although these methods aim to uncover hidden structures in data, they differ in their statistical theory, algorithms, and applications. This proposal aims to address the differences and similarities between these models and provide practical guidance for researchers in choosing the appropriate model for their unique data. We will introduce each model by presenting its statistical form and related theory, as well as the algorithms or estimation methods involved in its implementation. Examples of data analysis will be provided. We will then summarize and compare the models from various perspectives, highlighting their differences and similarities. Finally, we will provide practical suggestions for researchers in choosing and implementing these models in their own research. This proposal aims to assist researchers in building their own models and effectively analyzing their data.

Track: General Papers [Speed, In-person]
Synthesizing data from pretest-posttest-control-group designs in mediation meta-analysis

Zhiming Lu; Sun Yat-sen University; CN
Zijun Ke*; Sun Yat-sen University; CN
Rebecca Cheung; University of Reading; GB
Qian Zhang; Florida State University; US

This study aims to compare the approaches to computing effect sizes in mediation meta-analysis (MMA) over data from randomized studies with pretest-posttest control group (PPCG) designs. After the quantification of a treatment effect through meta-analyses over data from PPCG designs, researchers are likely to resort to meta-analytic structural equation modeling (MASEM) techniques to investigate how the treatment effect takes place (i.e., mediation meta-analysis). As PPCG designs represent a type of mixed design, the correlation coefficient between the mediator and the outcome variable reported and synthesized is typically based on the pretest-posttest change-scores, which takes into account the pretest-posttest correlation. However, when synthesizing the correlations between the independent variable and the mediator, and between

the independent variable and the dependent variable from "standardized mean change" measures, it is recommended to ignore the pretest-posttest correlation to make the effect sizes independent of study design. This results in an inconsistency when constructing the correlation matrices for MASEM. Using Monte Carlo simulations, this study evaluates two potential approaches, namely posttest-score MMA and change-score MMA (i.e., PSMMA or CSMMA) to address this inconsistency, and compares existing methods for computing effect sizes in each context. This study also assesses the impact of missing data on the performance of MMA. Results showed that CSMMA is preferred to obtain sufficiently powered results. Analyses are illustrated using a meta-analysis on the therapeutic mechanisms of the effect of mindfulness-based interventions on health outcome variables. A discussion and practical guidelines are also included to conclude the study.

Track: General Papers [Regular, In-person]
Assessing between- and within-person reliabilities of items and scale for daily procrastination: A multilevel and dynamic approach

Xiaohui Luo; Beijing Normal University; CN
Yueqin Hu; Beijing Normal University; CN
Hongyun Liu*; Beijing Normal University; CN

Intensive longitudinal data (ILD) has been collected to capture the dynamic fluctuations of procrastination; however, researchers have typically measured daily procrastination by modifying trait measures (e.g., adding a time reference “today”) without adequately testing their reliabilities. Furthermore, existing reliability estimation methods have limitations in reflecting both the multilevel and dynamic nature of ILD. Therefore, the main purpose of this study was to use an advanced approach, dynamic structural equation modeling, to assess the between- and within-person reliabilities of a widely used six-item measure of daily procrastination. In addition, two other noteworthy issues in the analysis and interpretation of the reliabilities of ILD were investigated. A total of 252 participants completed retrospective measures of various types of trait procrastination and daily measures of procrastination over 34 consecutive days. The results showed high between- and within-person reliabilities of the entire scale for daily procrastination, but much lower reliabilities for one item, suggesting that this item may be inappropriate in everyday contexts. Furthermore, we found moderate to strong associations between the latent trait factor of procrastination and trait measures of procrastination, confirming the need to include a latent trait factor in ILD to reflect between-person differences in procrastination. In addition, we identified substantial between-person differences in person-specific reliabilities and explored its relevant factors (i.e., trait procrastination, and within-person standard deviation of procrastination score), contributing to a more comprehensive understanding of the reliabilities. Overall, this study assessed the reliabilities of a daily measure of procrastination, which facilitated future studies to obtain more reliable and consistent results and to better estimate the reliability of ILD.

Track: Data Science Methods in Political Science and Public Policy Research [Speed, Online]
Election Winning Prediction Using Demographic and Past Political Data: Data Science

Approach

Ismail Olaniyi MURAINA*; Lagos State University of Education; NG

It is essential to employ data science in politics and policies; it is all about transforming data or unprocessed facts into valuable and accurate information. To answer some political issues and ultimately forecast the outcome of voting for a candidate in an election, code can be used to process enormous amounts of data from many sources in politics. This study makes use of demographic information on political candidates, such as their age, gender, ethnicity, and religion, as well as information about their previous political activities, including their political accomplishments in previous positions held and adherence to party rules. Each presidential candidate's demographic and political information was submitted to the public for a proper evaluation between 1 and 100 percent. To gather pertinent information from individuals who are well-versed in the candidates running for the position, the instrument was created and administered using Google Forms. The face validity method was used to validate the study instrument, which had a 0.88 reliability index. Findings indicate a high degree of accuracy in predicting who would win the election, even when taking into account factors like purchasing voters' minds to vote for a candidate who goes against their initial beliefs and technical difficulties with the voting equipment. According to the study, enabling voters to grade candidates for prediction is considerably more trustworthy and preferable to using social media since users are always being inundated with fake news, which in most cases weakens the trustworthiness of the election.

Track: General Papers [Speed, In-person]
How does Prior Variance Affect Local Dependence Detection in CFA

Xinyu Qiao*; Department of Psychology, Sun Yat-Sen University.; CN
Junhao Pan; Department of Psychology, Sun Yat-Sen University.; CN

Local independence is a fundamental assumption in confirmatory factor analysis (CFA). However, several factors (e.g., similar expression in the questionnaire) could result in violation of it. The Bayesian structural equation model can be used to deal with local dependence. Researchers can allow the existence of small local dependence by assigning a zero-mean and small-variance prior to the residual covariance matrix. It's still unclear how prior variance affects the detection of local dependence. Therefore, two simulation studies based on CFA model were conducted with continuous and binary data, respectively. The conditions included sample sizes, number of items, the magnitude of factor loadings, and levels of local dependence. Inverse Wishart distribution was applied for the residual covariance matrix. Five levels of prior variances and the default prior setting in M*plus* for residual covariances were considered. Results indicated that models were not identified for continuous data, and the convergence rate was low for binary data under default prior setting. For continuous data, larger prior variance resulted in less Type I error. The power to detect small local dependence (e.g., residual correlation = 0.2) was lower than 0.1 regardless of prior variances. For binary data, the prior variance mainly affects the accuracy of the estimates of the residual covariance. With the increase of prior variance, the mean square error and relative bias decreased. Power for detecting small or moderate

dependence was always less than 0.8 regardless of prior variances. We provided recommendations and two real data analyses to demonstrate the validity of this recommendation.

Track: Data Science in Humanities and Social Sciences [Speed, In-person]
A Comparison of Mathematical and Statistical Modeling with Longitudinal Data: An Application to Ecological Momentary Assessment of Behavior Change in Individuals with Alcohol Use Disorder

Sijing Shao*; University of Notre Dame; US

The utilization of Ecological Momentary Assessment (EMA) has become prevalent in behavioral research for collecting real-time longitudinal data. Various analytical methods have been employed to investigate the changes in behavior, motivation, and emotions on a daily or within-day basis. However, one of the challenges faced in the behavioral field is determining the appropriateness of these methods for specific research questions. In this paper, we compare two commonly used methods, Generalized Estimating Equations and Generalized Linear Mixed Models, with three less frequently used methods, Markov Models, Generalized Linear Mixed-Effects Markov Models, and Differential Equations, which are widely used in other fields. Our objective is to demonstrate the application of these five distinct analytical methods to an intensive longitudinal dataset on drinking behavior, emphasizing the usefulness of each approach.

Track: Data Science Methods in Political Science and Public Policy Research [Speed, In-person]
China's COVID Lockdown Policy and Trade with U.S.: A Deep Learning Time Series Approach

Min Shi*; The University of Texas at Dallas; US
Karl Ho; The University of Texas at Dallas; US

COVID-19 has had a catastrophic effect on human health and the global economy. Since its outbreak, more than 759 million positive cases and 6.8 million deaths due to COVID-19 have been reported to the World Health Organization (WHO) (WHO 2023). Due to its airborne transmission, high infection rate, and severity of damage to people's health, many governments adopted prolonged shutdowns, travel restrictions, international border and port closures, and other virus-containment measures to slow down the spread of the virus and thus alleviate pressure on strained and vulnerable health systems, which also resulted in the loss of livelihoods and a rippling effect on the economy (Blake & Wadhwa 2020, Cascella et al. 2022, Khorana et al. 2022).

International trade, one sphere of economic activity, is heavily impacted by the COVID-19 pandemic. Research on the effect of the pandemic on global trade reveals sizeable negative international spillovers to trade from supply disruptions due to shutdowns, either from importer or exporter sides (Aiyar et al. 2022, Barbero et al. 2022, Bas et al. 2022, Berthou & Stumpner

2022). As we all know, China insisted very strict Covid Zero policy during the pandemic. Some researchers found that COVID-19-related deaths have a causal relationship with the imports and exports of China, especially to countries at high risk (Zhang et al. 2022, Che et al. 2020). Furthermore, some scholars focused on the negative effect of COVID-19 lockdowns at firm-level based on the data of Spanish, Portuguese, and Colombian firms and pointed out that lockdown stringency in destination markets reduced firms' exports (Lucio et al. 2021, Pimenta et al. 2021, Dueñas et al. 2021). As the United States' largest supplier of goods imports, what's the effect of China's lockdown of borders and ports on U.S. trade at the firm level?

This paper adds to the studies of the impacts of China's lockdown due to COVID-19 on U.S. trade and its further influence on U.S. firms using a deep learning time series approach. Counterfactual predictions using deep learning could help evaluate how different sets of actions can impact target trajectories from a historical or a forecasting perspective (Lim & Zohren 2021). We build deep learning models under two scenarios: a COVID-19 circumstance, and a non-COVID-19 counterfactual setting based on the monthly data of COVID-19, U.S.-China trade, the revenues of U.S. firms, S&P 500 stock index, the U.S. Producer Price Index (PPI) and the U.S. Consumer Price Index (CPI). The differences in the predictions are compared to evaluate the impact of China's lockdown policies due to COVID-19 on U.S. trade at the firm level.

Track: Data Science in Humanities and Social Sciences [Regular, In-person]
Impact of Temporal Order Selection on VAR-based Clustering of Intensive Longitudinal Data

Hairong Song*; Department of Psychology University of Oklahoma; US
Yaqi Li; University of Oklahoma; US

Model-based methods for clustering intensive longitudinal data is a powerful tool to study similarity and dissimilarity in dynamic characteristics among multiple individuals. When implementing such clustering techniques, researchers need to set the temporal orders of the within-individual processes to be identical for all individuals. Two methods have been proposed here. One applies the most complex structure or highest order (HO) for all individual processes, while the other chooses the most parsimonious structure or the lowest order (LO) for all individuals. How each method, HO vs. LO, would perform in clustering multivariate dynamic processes under various data conditions? To our best knowledge, this question has not been well answered yet, leaving the practices of HO and LO as well as the model-based clustering techniques unwarranted. In this study, we introduced a two-step VAR-based clustering procedure and investigated the performance of HO and LO using this procedure with Gaussian mixture model (GMM) and K-means clustering algorithms. Our simulation study showed that LO generally performed better than HO on recovering true number of clusters and cluster membership. In addition, the GMM algorithm exhibited an improved accuracy over the K-Means on recovering numbers of clusters and cluster membership; however, the GMM algorithm demonstrated a higher sensitivity to the selection of temporal orders and showed worse performance on clustering dynamic processes with high-order VAR models.

Track: Data Science in Humanities and Social Sciences [Speed, Online]
The Roles of Emotion Regulation Strategies in Mothers' and Adolescents' Depression: Using the Actor-Partner Interdependence Model

Zhonghuang Su; Sun Yat-sen University; CN
Wenyu Qiu; Sun Yat-sen University; CN
Ruyi Ding*; Sun Yat-Sen University; CN
Junhao Pan; Department of Psychology, Sun Yat-Sen University; CN

Parent-adolescent emotion dynamics have attracted increasing attention in recent years because adolescence is a challenging period for both adolescents and their parents. However, how emotions are coconstructed between parent and adolescent is less clear. To provide insights into this issue, this research examined whether mothers' and adolescents' emotion regulation strategy use is linked with their own (actor effects) and each other's (partner effects) depression using Actor-Partner Interdependence Model (APIM).

One hundred and seventy-three adolescent (M_{age} = 13.00 years old, SD = 0.90) and mother (M_{age} = 43.05 years old, SD = 3.78) dyads, who completed self-report assessments of emotion regulation strategy and depression. APIM was conducted using Mplus. Actor effects indicated that the more mothers used cognitive reappraisal, the lower their depression levels were, and the more mothers and adolescents used expressive suppression, the higher their levels of depression were. Additionally, the results showed that maternal expressive suppression was associated with adolescent's depression. Moreover, there was a significant mothers' expressive suppression × adolescents' expressive suppression interaction on adolescents' depression. To be specific, for mothers with a high frequency of expressive suppression use (+1 SD) in the family, their adolescents' expressive suppression use was significantly positively related to adolescents' depression (b = 1.03, $S.E.$ = 0.21, p < .001), while for those mothers with low frequency of expressive suppression use (-1 SD), there was no significant correlation between adolescents' expressive suppression use and adolescents' depression. The findings point out that it is important to conceptualize parent-adolescent dyads as ties, that is, considering the contributions of two developing individuals to a broader dyadic system.

Track: General Papers [Speed, In-person]
Evaluation of the Bass-Ackward Method for Identifying the Number of Factors

Lingbo Tong*; University of Notre Dame; US
Zhiyong Zhang; University of Notre Dame; US

Factor analysis is commonly used for identifying latent factors behind the observed variables. Goldberg (2006) proposed a method named Bass-Ackward (BA) that led to a hierarchical structure of latent factors based on correlations among factors. While it has been used in research exploring the hierarchical structures among factors, no previous work has evaluated its performance in retaining the correct number of latent factors. In this study, we conduct a set of Monte Carlo simulations to evaluate the performance of BA in identifying the proper number of factors. Specifically, we compare two termination criteria designed for BA (BA-maxLoading and BA-cutoff) to two traditional factor retention methods (the K1 rule and parallel analysis). Our

results show that while parallel analysis achieved the overall best performance, BA-maxLoading outperformed parallel analysis in identifying the number of factors when there are cross-loadings and large factor correlations. Recommendations on when to use different methods and an online tool for visualizing the factor structures through the BA method are provided.

Track: General Papers [Regular, In-person]
Bayesian Compositional Data Analysis with Informative Priors

Xin Tong*; University of Virginia; US
Di Jia; University of Virginia; US
Yunli Liu; University of Virginia; US

Compositional data, or sometimes called ipsative data, are nonnegative data with a constant-sum constraint on the sample values. Examples include percentages that sum to 100, hours in a day that sum to 24, etc. Although compositional data allow us to study the relative importance of the compositions, they are not commonly used in social and behavioral sciences because they introduce the exact collinearity problem where regression coefficients cannot be identified. Traditional methods typically transform the compositions out of the simplex, e.g., using log-ratios or probability-ratios. However, model interpretation after the transformation is not straightforward. In this study, we propose a Bayesian approach with informative priors to conduct the compositional data regression analysis directly using proportions as predictors and to select important compositional predictors. The performance of the proposed method is evaluated by Monte Carlo simulations and is compared with existing penalized approaches.

Track: Data Science in Humanities and Social Sciences [Speed, In-person]
An analysis of emotion contagion with longitudinal text data

Xin Tong*; University of Virginia; US
Haiyan Liu; UC Merced; US
Shelly Tsang; University of Virginia; US
Adrienne Wood; University of Virginia; US

Emotion contagion occurs when people match the emotion states of their interaction partners. Although cross-sectional evidence showed that emotion contagion is positively related to social bond strength, few studies have investigated how this relationship persists over time. In this study, 59 pairs of previously unacquainted participants had conversations with each other once a week for six weeks. The recorded conversations were converted into text data. We conduct emotion detection text analysis to investigate the overall emotion of the dyads and the variance of the emotion scores at each time point. Latent growth models are then fitted to the emotion data to investigate the emotion movement (how much a person's emotion changes over the course of the friendship formation) and the change of emotion distance (how much partners' emotion differs) over time. The results are compared with the analysis of the self-reported emotion measures on the same participants to verify the accuracy of the longitudinal text data analysis.

Track: General Papers [Speed, In-person]
Evaluating the Threshold of Missingness Percentage in the Modeling of Intensive Longitudinal Data

Zhilin Wan*; Sichuan Normal University; CN
Yue Liu; Sichuan Normal University; CN

The Experience Sampling Method (ESM) is a common approach in psychological research for collecting intensive longitudinal data. Autoregressive modeling is widely used to analyze ESM data, which assumes that the time intervals between consecutive measurements are all equal. However, ESM data is often unequally spaced, which result from certain popular sampling schemes that are based on semi-random measurement occasions. Violation of the assumption can cause problematic interpretation of lagged relations and biased parameter estimates. Missing data insertion between realized observations is often used to obtain almost equal measurements. In addition, missing data is also frequently observed due to the participant noncompliance in many intensive longitudinal studies. The quality of estimation based on ESM models deteriorates as the amount of missing data increases. Consequently, many approaches have been proposed to handle missing data in ESM data. For instance, a two-step partial multiple imputation (2PMI) method has been shown to produce consistent estimates under certain conditions. The primary goal of this current study is to see how the percentage of missing data affects the performance of 2PMI method based on a multilevel autoregression model, and investigate the threshold of the amount of missing data to achieve acceptable model results under varies sample sizes.

A simulation study is conducted based on multilevel autoregression model containing time-varying and time-invariant covariates. Three factors were manipulated and fully crossed: the percentage of missing data (50%, 60%, 70%, 80%), number of subjects (50, 100, 200), and number of time points per subject (200, 300, 500). Missingness was imposed on each of the complete data sets, conforming to the missing at random (MAR) mechanism. 2PMI method was used to deal with missing data. Specifically, missing data in the covariates were multiply imputed at first. Afterwards, a Kalman filter approach was used to deal with missing data in the dependent variables. Parameters in the multilevel autoregression model were estimated applying a Bayesian approach. A total of 500 Monte Carlo replications were run for each condition. The results were evaluated in terms of convergence rate, power precision of the point estimates for the fixed effect, the standard error estimates, and the random effect estimates. The results showed that 2PMI method produced better overall estimation results as the percentage of missing data decreased. Practical implications about the threshold of missing data percentage were also provided across different conditions of sample size. The results may shed light on the decision about whether data collection should be continued based on the current percentage of missing data.

Track: Data Science in Humanities and Social Sciences [Regular, In-person]
How Does Self-Compassion Associate with Coping Self-Efficacy in Daily Life? A Dynamic Structural Equation Model Analysis

Huihui Wang*; Ningxia University; CN
Guifen He; Zhejiang University; CN
Enna Wan; Tianjin University; CN
Yang Wang; Guangdong University of Finance; CN
MingQing Zheng; Ningxia University; CN

Objectives The positive association between high self-compassion and coping self-efficacy has been well documented. However, at the level of short-term dynamics, the extent to which state self-compassion or coping self-efficacy persists in daily life is under-explored. This experience sampling study aimed to explore the temporal relationship between self-compassion and coping self-efficacy.

Methods Participants ($N = 240$, $M_{age} = 18.98 \pm 0.99$ years, 44.8% female) completed 14 consecutive daily diaries (twice a day, for 28 measurement intervals) on self-compassion and coping self-efficacy (for a total of 6,438 observations). We combined the two interrelated time series into one model using dynamic structural equation modeling. Both effects (self-compassion on subsequent coping self-efficacy and coping self-efficacy on subsequent self-compassion) were modeled simultaneously while acknowledging the autoregressive characteristics of both states (inert properties).

Results We found evidence for the reciprocal effects of self-compassion and coping self-efficacy, and both processes showed autoregressive relationships. Self-compassion (b = 0.094, 95% CI [0.047, 0.136]) and coping self-efficacy (b = 0.212 , 95% CI [0.164, 0.258]) demonstrated stability through autocorrelations. Self-compassion was a significant predictor of subsequent coping self-efficacy (b = 0.078, 95% CI [0.036, 0.131])), and coping self-efficacy was a significant predictor of subsequent self-compassion (b = 0.048, (95% CI [0.005, 0.088]).

Conclusions Together, at the level of short-term dynamics, we found that coping self-efficacy can be viewed as both an antecedent and an outcome of self-compassion. This means there has been a virtuous cycle between state coping self-efficacy and self-compassion.

Track: Data Science in Humanities and Social Sciences [Regular, In-person]
A Tutorial in Longitudinal Network Analysis

Huihui Wang*; Ningxia University; CN
Xuzhu Zhou; Ningxia University; CN
MingQing Zheng; Ningxia University; CN

Longitudinal network analysis which can estimate network structures at the item level and dynamic changes in individual attributes is now receiving more attention. However, one barrier to broader application of longitudinal network analysis was the lack of the methodological tutorial explaining what types the method is divided into, what the statistical

principles are, and how to implement it with software. As such, in this paper, we present a longitudinal network analysis model to reveal item-level longitudinal effects that occur within and across constructs over two or more time points. Firstly, we introduced the statistical analysis principles of the longitudinal network analysis involving two time points, i.e., cross-lagged panel network model, and the longitudinal network analysis involving multiple time points. Then we use data on depression symptoms and self-efficacy in low-income mother to conduct cross-sectional data network analysis, cross-lagged panel network analysis involving two time points and longitudinal network analysis involving four time points. We discussed the difference in interpreting the results of these models and proposed estimation methods to obtain these network models, which we implement in the R packages graphical*VAR* and ml*VAR*. I think this article could provide methodological guidance for researchers who want to carry out longitudinal social network analysis.

Track: General Papers [Regular, In-person]
Modeling intraindividual variability as a predictor with longitudinal data: Methods and evaluations

Lijuan Wang*; University of Notre Dame;
Xiao Liu; University of Texas - Austin;

In many areas of psychology, researchers are interested in studying whether intraindividual variability (IIV) is predictive of behavioral and health outcomes after controlling for the intraindividual mean. To the end, IIV indicators such as the observed intraindividual standard deviation (ISD) or observed intraindividual variance (IVAR) are often modeled as a predictor in regular regression analysis. However, observed ISD or IVAR scores have been found to have the low reliability problem especially when the number of occasions is small. The low reliability problem can lead to incorrect statistical inferences about the coefficient of ISD or IVAR from regular regression. In this study, we examined the statistical features (mean and variance) of the observed ISD and IVAR. The results revealed that regular regression can yield (1) more accurate results about the coefficient of IVAR but (2) worse results about the coefficient of ISD when the number of occasions increases. Furthermore, we compared the performance of three alternative modeling methods including the time parceling, single indicator latent variable, and Bayesian variability modeling methods to that of regular regression for modeling ISD or IVAR as a predictor. Simulation results were consistent with the analytical results and further suggested that our proposed (a) time parceling with bootstrapping and (b) Bayesian variability modeling approaches performed well and better than regression for modeling IVAR as a predictor. Only the proposed Bayesian variability modeling method performed well for modeling ISD as a predictor. Implications of the results and recommendations were discussed.

Track: General Papers [Speed, In-person]
The Story After Mobility: Debates and Critiques on Methods for Estimating Social Mobility Effects

Peng Wang*; Department of Sociology, The Chinese University of Hong Kong

Social mobility is one of the core research topics in sociology, and its implications have garnered increasing attention. However, accurately measuring the effects of social mobility has been challenging due to issues related to model identification. Over the years, significant advancements have been made in methods for measuring social mobility effects, progressing from Duncan's Square Additive (SA) model to Hope's Diamond Additive (DA) model, and subsequently to Sobel's Diagonal Reference Model (DRM). The DRM, in particular, has been widely utilized and considered the "gold standard" for estimating social mobility effects. Nevertheless, scholars have recently criticized the DRM and proposed alternative methods, such as the Mobility Contrast Model (MCM), for measuring mobility effects. This study critically examines the debates surrounding these methods, systematically comparing them in terms of theoretical assumptions and methodological settings, with a specific focus on the widely employed DRM and the newly introduced MCM. Through simulations and empirical analysis, this study highlights that the primary distinctions between these methods lie in their underlying theoretical assumptions and their capacity to address diverse theoretical concerns, rather than the statistical superiority or inferiority of the techniques themselves. Consequently, the suitability of theoretical assumptions in practical contexts should be taken into consideration when selecting a method. Lastly, this study identifies key considerations in applying these methods and suggests approaches for conducting robustness tests.

Track: General Papers [Speed, In-person]

Automated evaluation and measurement model for the effectiveness of state transitions in problem-solving tasks

Pujue Wang; Beijing Normal University; CN
Hongyun Liu*; Beijing Normal University; CN

For the computer simulation interactive tasks designed with finite state automata as the prototype, there are complex state transitions between initial and target states, which makes it difficult to determine the effect of each state transition on task completion. Among the measurement models for analyzing process data that emerged in recent years, the effectiveness of the state transition is also an important parameter in the models. So far it is evaluated manually in a dichotomous way, which limits the application scenarios of the measurement models. To solve these two problems, this study proposes a new index that conforms to the definition of problem-solving, called the change in the shortest distance to target states (CSDT). CSDT can automatically calculate the effectiveness of any state transition and can produce the polytomous score. Furthermore, a new measurement model combining CSDT is proposed, called the response evaluation measurement model (REMM). REMM is suitable for more types of problem-solving tasks, including those with a single path to complete the task and those with multiple paths to success.

The main idea of CSDT is described as follows: For each state, the shortest distance to any target state of the task is obtained as the effectiveness of the state, which is the minimal number of transitions required to complete the task. Then the change in this distance caused by a transition between one state to another can be found as the effectiveness of this transition, which is the number of steps this transition moves forward or backward to achieve the task. Taking CSDTs of all transitions as the effectiveness parameters, REMM can be established for all subjects under the framework of the nominal response model. REMM allows a more detailed evaluation of the tendency for each transition and the latent ability of problem-solving for each subject.

In this study, two computer-based interactive tasks were selected, namely, the Tickets task from PISA 2012 with a single path to the target state, and the Balance task from ATC21S with multiple paths to multiple target states. The algorithms to automatically calculate CSDT of all transitions for two tasks were designed. The evaluation for all transitions was quickly finished for all subjects. Then REMM was fitted to estimate the tendency parameters for all transitions, as well as the abilities of all subjects in the Tickets task and of all groups in the Balance task. The tendencies were consistent with the values of CSDT and theoretical difficulties. The response sequences with high and low ability estimates conformed to the task settings. The abilities were highly correlated with the proportion of effective and ineffective transitions subjects performed evaluated by CSDT.

Track: General Papers [Regular, In-person]
Determinate the number of attributes using neural network

Yali Wang; jiangxi normal university; CN
Yujun Li*; Jiangxi Normal University; CN

During a typical learning process, to what extent can timely feedback benefit learner hinges upon accurate diagnosis of learner's skill profile. To that end, cognitive diagnosis analysis (CDA) serves as a useful tool that assesses students' mastery of skills, known as attributes in CDA, from their performance on a cognitive diagnostic test. The determination of an appropriate number of attributes underpins the accuracy of the test in use, and it is also widely considered as one of the most critical issues in CDA (Xiong et al.,2022; Ma et al.,2022). The majority of existing studies in CDA are aiming for a scenario with a relatively small number of independent attributes (Philipp et al.,2016). However, attributes in real educational assessment can be rather large in number and have inherent hierarchical structures (Han et al.,2020; Tang et al.,2018), thus, it is imperative to develop a new extraction method to fill this gap. In this article, we present a new approach to estimating the number of attributes from a data-driven perspective. A neural network model is trained using data simulated under a broad range of realistic conditions. The new method is then compared with 3 commonly used extraction methods of its accuracy in determining the correct number of attributes for both simulated data and real data. In the simulation experiment, a wide-range design is employed for factors including the number of attributes, the ratio of number of items to number of attributes, and the ratio of number of items to sample size. The hierarchical structures among attributes are randomly generated. To mimic small-scale in-class tests, attributes can be randomly generated as either

independent or having hierarchical structures when the number of attributes is less than 8. The results show that, a) for simulated data, our method outperforms existing methods with accuracy as high as 94%, while the highest accuracy of existing methods is capped at 16%; b) for real data, our method yields identical results with that of field experts, while the other methods all underestimate the number of attributes to a certain degree.

Track: Data Science Methods in Political Science and Public Policy Research [Speed, In-person]

A 2-stage Latent-standardization Method in Second-order Latent Growth Modeling

Zhonglin Wen*; South China Normal University; CN
Yifan Wang; South China Normal University; CN

Latent growth models (LGMs) are powerful tools for analyzing longitudinal data and have attracted the attention of searchers in psychology and other social science disciplines. For a latent variable measured by multiple indicators, the second-order latent growth models (2nd-order LGMs) are superior to the univariate LGMs (also called the first-order LGM) based on composite scores in many aspects.

Under the condition of strong measurement invariance across time, the estimation of growth parameters in the 2nd-order LGMs depends on the scaling method of factors/latent variables. There are three scaling methods: the scaled-indicator method (also called the marker-variable identification method), the effect-coding method (also called the effect-coding identification method), and the latent-standardization method. The estimation given by the latent-standardization method when the first time point as the reference can be regarded as the standardized solution.

The existing latent-standardization method needs to know the reliability of the scaled-indicator or the composite scores at the first time point. In this study, we propose an operable 2-stage latent-standardization method. In the first stage, a confirmatory factor analysis (CFA) with strong measurement invariance is conducted by fixing the mean and variance of the latent variable at the first time point to 0 and 1 respectively. In the second stage, estimated loadings are used and the 2nd-order LGM is established. If the standardization is based on the scaled-indicator method, the loading of the scaled-indicator is fixed to that obtained in the first stage, and the intercept of the scaled-indicator is fixed to the sample mean of the scaled-indicator at the first time point; If the standardization is based on the effect-coding method, the sum of loadings is constrained to the sum of loadings obtained in the first stage, and the sum of intercepts is constrained to the sum of the sample mean of all indicators at the first time point.

By a simulation study, the 2-stage latent-standardization method is compared with the existing latent-standardization method where the reliability of the scaled-indicator or the composite scores is given. The result shows that the 2-stage latent-standardization method performs almost as good as the existing latent-standardization method, but the former does not require knowing the reliability, and thus has more operability.

Track: Data Science Methods in Political Science and Public Policy Research [Speed, Online]
Supervised Machine Learning Applications for Detecting Internet Research Agency Misinformation

Thomas Wiese*; SUNY Empire State College; US

Misinformation has shifted political narratives across the globe. Because information shared over social media platforms lack traditional publishers and editors, the public is more susceptible to consuming information that is untrue. During the 2016 U.S. presidential election, the Russian government sponsored information operatives to spread misleading and/or false claims through social media. This study defines a method for automated detection of misinformation on social media using machine learning.

Track: Data Science in Humanities and Social Sciences [Regular, In-person]
Effects of BAS/BIS on person-situation dynamics

Fan Wu; Beijing Normal University; CN
Yueqin Hu*; Beijing Normal University; CN

Personality dynamics is a burgeoning research field which emphasizes personality process in daily life and its underlying mechanisms. Under this perspective, we proposed two approaches to test hypotheses in personality dynamics. The first approach focuses on the relation between personality states and situations, and the second approach focuses on the underlying stable mechanisms by which personality states fluctuate across situations. We then illustrated these approaches with two studies using experience sampling method.

In study 1, participants (n = 97) received questionnaires measuring the Big Five personality states and the subjective Situational Eight characteristics three times per day for 14 days. Multilevel dynamic structural equation modeling was used to analyze autoregressive and cross-lagged effects of situation characteristics and personality states. Findings indicated that subjective situation characteristics and personality states both have moderate stability. For cross-lagged associations, results showed a pattern of directionality as situation characteristics$_{(t-1)}$ can significantly predict personality states$_{(t)}$ in most person-situation links but not the other way around. Neuroticism stood as an exception as it has significant reciprocal relations with situation characteristics.

In study 2, we aimed to examine how dynamic situational appraisals and stable personality traits interact to predict personality manifestations in daily life. Participants (n = 154) received the same daily measures as in study 1. In addition, we measured BAS (behavioral approach system) and BIS (behavioral inhibition system) sensitivity as the indicator of stable endogenous personality variables. The results of autoregressive and cross-lagged paths replicated that of study 1. Furthermore, we found significant cross-level moderation effects. Specifically, BAS served as a protective force as it alienated the effect of negative situation characteristics$_{(t-1)}$ to neuroticism state$_{(t)}$ and strengthened the effect of positive situation characteristics$_{(t-1)}$ to

extraversion state$_{(t)}$. Based on both studies, we concluded that trait-like personality inclinations interact meaningfully with state-like situational appraisal in influencing how individual acts in daily life.

Track: General Papers [Speed, In-person]
Small Data Approaches to Link Faster Time Scale Engagement Dynamics with Slower Time Scale Outcomes in Biobehavioral Interventions

Jingchuan Wu*; Pennsylvania State University; US
Nilam Ram; Stanford University; US
James Marks; Penn State College of Medicine; US
Necole Streeper; Pennsylvania State Health Milton S. Hershey Medical Center; US
David Conroy; Pennsylvania State University; US

Introduction. Two fundamental challenges of behavioral intervention development involve engaging participants with intervention content and initiating behavior change based on that content. Digital technologies are increasingly used to sense contextual variation and repeatedly engage participants with intervention content at key moments in the natural context of their daily lives. The longitudinal patterning of participants' engagement with those interventions on fast time scales (e.g., daily) may be an advance indicator of clinical outcomes on slower time scales (e.g., monthly). With large, labeled datasets, deep learning approaches can detect subtle but critical patterns in engagement that impact clinical outcomes. However, many important clinical and biobehavioral problems are characterized by small data, intensive longitudinal data from limited numbers of participants, that are not amenable to deep learning. In those cases, alternative approaches, such as time series clustering or feature engineering, might be used to identify patterns of engagement that are associated with improved health outcomes. In this study, we conducted secondary analysis of data from a one-month, single-group trial of a digital intervention to promote fluid intake among patients with kidney stones. For these patients, increasing urine output is a core clinical guideline for preventing a costly and painful recurrence of kidney stones. Patients are recommended to consume enough fluids to produce 2.5L of urine daily but less than half of patients adhere to that guideline. Based on patients reports that a lack of thirst and forgetting to drink were key barriers to adherence, the sipIT intervention was developed to (1) engage patients in semi-automated tracking of fluid intake and (2) deliver just-in-time text message reminders to trigger fluid intake when participants lapsed in attaining hourly fluid intake goals. The semi-automated tracking system incorporates a connected water bottle for automated tracking and a mobile app for manual tracking of fluid intake. Analyses were conducted to evaluate whether urine volume (a surrogate biomarker of kidney stone risk in this population) differed among participants with different patterns of engagement with the connected water bottle and mobile app over a one-month period. Both time series clustering and feature engineering approaches were applied.

Methods. Patients with a history of kidney stones who were identified as low urine producers (verified by recent 24-hour urine volume in the electronic medical record; n=26) were recruited from a urology specialty clinic. Timestamped data from every automated bottle recording and manual app entry were downloaded at the end of the study and aggregated into two daily scores

representing the frequency of use for each method. The primary clinical outcome was 24-hour urine volume collected after 1 month of using the sipIT intervention. The Tscluster package in R was applied to identify clustered time series based on daily tracking behavior with the connected water bottle and app. The Dynamic Time Warping distance metric measured the similarity between the time series data, and hierarchical clustering was performed on the dissimilarity matrix. A two-way ANOVA was conducted to examine the effects of time and cluster membership on urine volume. In parallel, theoretically-informed features were extracted from daily engagement data to characterize each participants' time series: mean, standard deviation, maximum, minimum, latency to maximum, latency to minimum, and frequency of daily disuse (all intraindividual). Linear multiple regressions were estimated to test whether urine volume was associated with the extracted features of each time series.

Results. The time series clustering analysis identified two clusters of time series for both the connected water bottle and mobile app. For both clusters, size was imbalanced (bottle: 92% vs 8%; app: 19% vs 81%). An ANOVA was conducted to investigate the differences in urine volume between the two categorical variables, cluster of bottle use and cluster of app manual input. The results showed no significant main effect of the cluster of bottle use on urine volume ($F(1, 23) = 0.319$, $p = 0.577$). However, a significant main effect of the cluster of app manual input on urine volume was observed ($F(1, 23) = 7.98$, $p = 0.01$). A post hoc analysis using Tukey's HSD test revealed a significant difference in urine volume. Group 1 (High Engagement Group) produced 1220 mL (95% CI: 317.8, 2122.3) more urine than group 2 (Low Engagement Group; $p = 0.01$).

In the linear regression analysis, urine volume was associated with certain features describing engagement with the connected water bottle and mobile app. For the connected water bottle, the only marginally significant relationship observed was between urine volume and the frequency of daily bottle disuse ($b = 1871.2$, $p = 0.06$). Participants who accumulated more days in which they did not use the bottle produced greater urine volume. No other features of bottle engagement were associated with urine volume. For the mobile app, a marginally-significant relationship was found between urine volume and the mean frequency of app input ($b = 458.0$, $p = 0.05$). Participants who engaged more with the app more frequently produced greater urine volume. No other features of mobile app engagement were associated with urine volume in this model.

Discussion. This study compared two methods for linking intensive engagement data from a digital intervention on a daily time scale with a health outcome on a slower time scale. Substantively, results provided a more nuanced understanding of engagement dynamics. Contrary to widespread assumptions that engagement increases intervention effects, engagement in tracking fluid intake may be a double-edged sword for the sipIT intervention because intervention doses are aliased to a lack of engagement. Participants who fail to engage with the bottle are more likely to receive hourly messages reminding them to drink, and participants who engage regularly with the app are likely to meet their hourly fluid intake goals and consequently less likely to receive messages reminding them to drink. Future research should aim to replicate these findings in larger, more diverse samples and investigate the specific mechanisms through which engagement affects clinical outcomes. Methodologically, this study demonstrated two established approaches that have not been widely applied to link engagement with intervention

outcomes. These methods provide a bridge between established big and small data approaches. Time series clustering was somewhat more powerful in these analyses but both approaches provided unique insights and warrant consideration by applied data scientists investigating digital health interventions.

Track: General Papers [Regular, In-person]

Investigating the Factor Structure of Sense of Social and Academic Fit Scale: A Multilevel Bifactor Study

Lixin Wu*; University of Illinois at Urbana-Champaign; US
Ge Jiang; University of Illinois at Urbana-Champaign; US

Although the Sense of Social and Academic Fit (SSAF) scale is popularly used to measure college students' sense of university belonging, its factor structure has yet to be examined extensively. Indeed, there have been very few published validation studies regarding the psychometric properties of the SSAF scale, which has concerned multiple scholars. For example, Pyne et al. (2018) noted, "Walton and Cohen (2007)...references a validation manuscript-in-preparation for the scale ... [that] remains unpublished". Knekta et al. (2020) also noted that "no validation studies or theoretical framework for the questionnaire has been published so far". To the best of our knowledge, only one study has investigated the validity and reliability of the SSAF scale (Maghsoodi et al., 2022).

We conducted the present two studies to investigate the factor structure of the SSAF scale. Specifically, in Study 1, we collected longitudinal data (repeated responses nested within individuals) on college students' university belonging. In Study 1m, once a multilevel study is warranted, we employed the multilevel bifactor analysis (ML-BA) technique in the present studies, including multilevel exploratory and confirmatory bifactor analysis. In Study 2, we recruited another cohort of college students and collected longitudinal data. In cases where multilevel modeling is guaranteed, we adopted the confirmatory bifactor analysis to cross-validate the findings of Study 1. In Study 1, a total of 377 students, including undergraduate and graduate students, from a public Midwestern four-year university were initially recruited in Study 1. In Study 2, the participants were 58 undergraduate students from the same university. During the data collection, students were expected to complete the 17-item Sense of Social and Academic Fit (SSAF) scale, which was designed to evaluate how much college students feel they belong in various social and academic situations, such as when navigating the rules in campus and interacting with professors.

In Study 1, in cases where multilevel modeling is warranted, we would conduct a series of multilevel bifactor analyses (ML-BAs) in three stages to investigate the factor structure of the SSAF scale at the BI and WI levels. In the first stage, we examined the factor structure Maghsoodi et al. (2023) proposed with multilevel confirmatory bifactor analysis (ML-CBA). In the second stage, we conducted a set of multilevel exploratory bifactor analyses (ML-EBAs) to

identify the factor structure should Maghsoodi et al.'s model(s) not perform satisfactorily. In the third stage, we conducted ML-CBA to validate the factor structure(s) suggested by ML-EBAs. We performed these analyses based on the polychoric correlation matrices (Olsson, 1979) because the 17 SSAF items were measured on a 5-point Likert-type scale that produced categorical responses. For the same reason, we used the weighted least square with mean and variance adjustment (WLSMV) estimator across all ML-BA models. In Study 2, if multilevel modeling is warranted, we would conduct a set of ML-CBAs to cross-validate the findings of Study 1.

Overall, the intraclass correlation coefficients warranted the multilevel modeling in the present studies. Overall results showed that a bifactor model with two specific factors at both levels fit the data well. Besides the negative wording effect added at the BI level, the factor structures differed across the two levels within a multilevel framework. Specifically, the SSAF scale was rather unidimensional than multidimensional at the BI level. In contrast, the SSAF scale included a general factor and two specific dimensions at the WI level. Our findings supported using the SSAF scale to measure college students' sense of university belonging in longitudinal studies.

Track: Data Science in Humanities and Social Sciences [Speed, In-person]
A psychometric network analysis of adolescents' mental health status: Evidence for key features for interventions

ANG XIA*; Sun Yat-sen University; CN
Ruyi Ding; Sun Yat-sen University; CN

Psychosocial problems are prevalent during adolescence. In the past decades, Chinese urban and rural regions have experienced a huge difference in economic development, educational resources, and cultures and norms. Adolescents' psychosocial problems have been documented to differ in the two regions, nevertheless, no consistent conclusion has been reached concerning whether the urban or the rural adolescents experience a higher level of psychosocial problems. This research aims to upfold the characteristics of Chinese urban and rural adolescents' psychosocial problems using psychometric network analysis. Strength and Difficulties Questionnaire (SDQ), which is a commonly used scale to assess adolescents' psychosocial problems, was used.

Data were analyzed based on reports from adolescents in urban region (n = 371; Males = 182, Females = 189, average age = 14.02 years old) and rural region (n = 371; Males = 183, Females = 188, average age = 13.98 years old).

In general, the results showed that the urban network and rural network had a similar pattern in SDQ as a whole and its subscales. However, differences emerged in the centrality. In particular, the results showed that the urban network centered on items " Helpful if someone is hurt [reversed] ", " Often volunteers to help others [reversed]" and "Often fights with other children", while the top three items most central in the rural network were "Helpful if someone is

hurt [reversed]", " Rather solitary, tend to play alone" and "Often lies or cheats". Besides, in peer problems subscale and prosocial subscale, urban network and rural network had different key items. The most important items of urban network and rural network in peer problems subscale were "Picked on or bullied by other children" and "Has at least one good friend" respectively. And in prosocial subscale, urban network centered on "Often fights with other children" while rural network centered on "Steals from home, school or elsewhere". In general, the urban network and rural network had a similar pattern in SDQ as a whole and its subscales, while there were still some different details. As central items highlight aspects of adolescents' mental health status that influence other aspects, these findings could be used to help intervention to safeguard adolescent mental health.

Track: General Papers [Regular, In-person]
Revisiting the Evaluation with Coefficient Alpha on Varying Length Scales and Sample Sizes

Leifeng Xiao*; Department of Educational Psychology, The Chinese University of Hong Kong; HK
Kit-Tai HAU; Department of Educational Psychology, The Chinese University of Hong Kong;

There is a growing interest on short scales in social science research, partly because survey time is limited, and researchers are interested to work on many dimensions simultaneously. However, researchers might misunderstand short scales should have similar internal reliability as long ones. In this research, we demonstrated and discussed the appropriate use of internal coefficient alpha (alpha) on varying scale lengths. We started with an overview of the mathematical definitions relating alpha and scale length, then we used examples and mathematical derivations to illustrate some common challenges. Notably, first, the impact of additional items (e.g., a low-quality item) was greater for short and weak scales; measurement (precision) was more accurate for longer scales and larger sample sizes (N>=500). Second, we proposed a rough but simple guideline in using the "alpha-if-item-deleted" procedure to decide when to retain and drop items when the scale is long (or short). Specifically, it is advisable to remove an item if alpha decreases by less than .02 upon its removal particularly for short scales. Whereas we will keep an item if alpha decreases more than .05 upon its removal. Third, considering the trade-off between brevity and scale quality, we recommended dynamic benchmarks for varying scale length, and suggested using a higher reliability benchmark for longer scales and larger samples. The .80 benchmark is relatively safe for most scenarios. We concluded that using one golden standard for varying scale lengths is inappropriate. Qualitative analyses (e.g., inspection of item content, face validity, content coverage) would be critical in ensuring scale quality.

Track: General Papers [Regular, In-person]
Two-Parameter State Response Measurement Model for Process Data and Its Mixture Extension

Yue Xiao; East China Normal University; CN
Hongyun Liu*; Beijing Normal University; CN

Process data recorded in computer-based interactive assessments reveals in detail how respondents reach the final outcomes and provides a more comprehensive picture of respondents' problem-solving competence and cognitive processes. In addition, it contains rich information about the task's dynamic characteristics as the task scenario unfolds or changes with ongoing problem-solving process.

Recently, new methods have been proposed to infer individual problem-solving competence based on process data (e.g., Chen, 2020; Lamar, 2018; Liu et al., 2018; Shu et al., 2017). However, most of them do not consider or care about task characteristics at the process level that may facilitate the understanding of population behavioral features and of the tasks. In addition, all of those methods assume homogeneity, whereas distinct subgroups with different problem-solving behavioral patterns (e.g., Liu et al., 2018; Kerr et al., 2011; He et al., 2019) have been identified through cluster analysis or discrete mixture versions of traditional psychometric models.

In this study, we first developed a two-parameter state response (2P-SR) model with MCMC estimation, which is an extension of the state response (SR) model proposed by Xiao and Liu (2023) that defines dynamic task characteristics but considers only the state easiness parameters. The 2P-SR model assumes that the action choice in each task state during problem solving depends on both the person's ability and the easiness and discrimination of that state. The easiness and discrimination quantify the difficulty of making the correct choice and the difference in performance between high- and low-ability respondents in that state, respectively, helping to identify critical states and better understand population behavioral patterns. Then we extended the 2P-SR model to a mixture model with class-specific state parameters to detect latent subgroups with different response propensities towards task states, which is called the mixture two-parameter state response (Mix2P-SR) model.

Parameter estimation for both models was examined using two small simulation studies. In general, the 2P-SR model provided accurate estimation, which could be improved with larger sample size and longer sequences. Besides, the estimation accuracy of 2P-SR model was similar to that of SR model when discrimination parameters were equal across states within the same task, and was much higher with unequal state discrimination levels. However, the application of Mix2P-SR model places demands on the data. When the action sequences were too short or the between-class distinction was small, the model performance was unsatisfactory in terms of convergence, classification, and parameter estimation.

An empirical example using process data from two problem-solving tasks in PISA 2012 demonstrates the application of the two models. The 2P-SR model provided interpretable parameter estimates for all task states, and there was a great variation between the discrimination of task states, suggesting the need to consider this parameter for task states. When using the Mix2P-SR model, two latent classes of respondents were detected, who differed largely in the easiness and discrimination parameters of some task states, reflecting the different behavior characteristics of the two groups in the problem-solving process.

Track: General Papers [Speed, In-person]

Using semantic items to detect careless responding and warning in real-time and its effect on questionnaires' quality

Chang Xin*; Beijing Normal University; CN
Jian Li; Beijing Normal University; CN

Careless responding (CR) is the common source of measurement errors in self-administered data, especially in surveys and questionnaires. Recent studies have used real-time warnings to reduce and prevent CR (Gibson, 2019; Sun et al., 2022; Zhang & Conrad, 2018), but there are several limitations of the existing methods. First, the detection indexes are limited, which most of them only used reaction time (RT) and non-difference index (e.g., longstring, LS). Second, the quality criteria are also inadequate, including RT, non-difference index, and items having correct answers. In current study, we used NaoDao platform (Chen et al., 2023) and jsPsych (de Leeuw, 2015) to implement the real-time warning function. We added semantic index to screen the respondents in real-time, including bogus items (BI), instructed response items (IR), semantic synonym & antonym (SS & SA), and even-odd consistency. This could find the omitted CR when using only RT and non-difference index. Besides the CR index mentioned above, this study also included other CR index as the quality criteria, including Mahalanobis distance (MD), self-report items and psychometric synonym & antonym. Furthermore, the criterion validity and construct validity were calculated, evaluating the impact of real-time warning on questionnaires' quality.

Track: General Papers [Regular, In-person]
Multitask prediction algorithms and their development

Li Xing*; The University of Saskatchewan; CA

For simultaneous prediction of multiple labels, we propose variations of a stacking algorithm which borrow information among multiple prediction tasks to improve multivariate prediction performance. We apply the methods to real-world data to show their better prediction performance and other attritive features, such as flexibility in handling various outcome types and easy ensembling of diverse single learners. Those methods have been built into an R package, MTPS, available for public users.

Track: General Papers [Speed, In-person]
Integrating Structural Equation Modeling with Social Networks

Ziqian Xu*; University of Notre Dame; US
Zhiyong Zhang; University of Notre Dame; US

Structural equation modeling is a widely used statistical technique in social sciences and has been traditionally applied to continuous, ordinal, and categorical data. However, with the rapid development of digital technologies, a vast variety of new data types such as text data, biometric data, and network data are now available in social sciences. As a result, it has become

increasingly important to extend the applicability of structural equation modeling to these data types. In this project, we focus on adapting structural equation models to include social networks as variables. Social networks can provide valuable information such as interpersonal relationships and popularities that can enhance the interpretability of structural equation models in explaining how behaviors, cognitive patterns, and psychological features are linked to social relationships. We discuss two approaches for using social networks as variables in structural equation models: using network nodes as variables in individual-centered structural equation models and using network dyads as variables in relationship-centered structural equation models. These approaches can be applied directly to network features, but they can also be extended to different network models such as the latent space model. Finally, we provide empirical examples of how to conduct structural equation modeling with network variables in practice.

Track: General Papers [Regular, In-person]
Modeling Data with Measurement Errors but without Predefined Metrics: Fact vs Fallacy

Ke-Hai Yuan*; University of Notre Dame, USA; US
Zhiyong Zhang; University of Notre Dame, USA; US

Data in social and behavioral sciences typically contain measurement errors and also do not have predefined metrics. Structural equation modeling (SEM) is commonly used to analyze such data. This article discusses issues in latent-variable modeling as compared to regression analysis with composite-scores. Via the analysis of a real dataset, several notions related to bias, parameter efficiency, standardization, and result interpretation are shown to be fallacy instead of fact. The clarifications facilitate better understanding of the strength and limitations of SEM and regression analysis with weighted composites, and are expected to advance social and behavioral data science.

Track: General Papers [Regular, In-person]
Latent Class Dynamic Mediation Model

Ying Yuan*; University of Texas MD Anderson Cancer Center; US
Jing Huang; The University of Pennsylvania; US

Traditional mediation analysis assumes that a study population is homogeneous and the mediation effect is constant over time, which may not hold in some applications. Motivated by smoking cessation data, we propose a latent class dynamic mediation model that explicitly accounts for the fact that the study population may consist of different subgroups and the mediation effect may vary over time. We use a proportional odds model to accommodate the subject heterogeneities and identify latent subgroups. Conditional on the subgroups, we employ a Bayesian hierarchical nonparametric time-varying coefficient model to capture the time-varying mediation process, while allowing each subgroup to have its individual dynamic mediation process. A simulation study shows that the proposed method has good performance in estimating the mediation effect. We illustrate the proposed methodology by applying it to analyze smoking cessation data.

Track: Data Science in Humanities and Social Sciences [Speed, Online]
Communication Patterns of Adolescents' Distress in Mother-Adolescent Dyads Predict Adolescents' Suicidal Ideation: Mediated by Adolescents' Internalizing and Externalizing Problems

Zhou Yue; Sun Yat-Sen University; CN
Ruyi Ding*; Sun Yat-Sen University; CN

Adolescents' suicide has become a prevalent problem worldwide. There has been an increasing call for research of examining the protective and risky factors of adolescents' suicide. Considering that adolescents' distress is one of the major reasons for their suicide-related behaviors, this research examined whether communication patterns of adolescents' emotional distress (CPAED) could predict adolescents' suicide ideation using a longitudinal study. CPAED refers to when adolescents experience emotional distress, the open communication (i.e., successful initiation of communication) or defective communication (i.e., unsuccessful initiation of communication) conducted in parent-child dyads. A total of 836 adolescents (Mean age=15.85, 43.9% female) participated in this research in two waves (one year apart). The results showed that active and reactive emotion sharing (i.e., adolescents' active sharing of emotion or reactive sharing in the face of maternal inquiry) at wave 1, negatively predicted adolescent suicidal ideation at wave 2, while lack of response by adolescents (i.e., describing adolescents' avoidance of communication even when asked by mothers) at wave 1 positively predicted adolescent suicidal ideation at wave 2. Moreover, this research found that adolescents' internalizing and externalizing problems at time 2 both significantly mediated the relationship between the two communication patterns and adolescent suicidal ideation. In conclusion, this

research points out the importance of CPAED in adolescents' suicide ideation and provides information about the underlying mechanism that CPAED predict adolescents' suicide ideation.

Track: General Papers [Speed, In-person]
Estimating Average Treatment Effects in the Context of Outcome-Dependent Sampling Designs with Mismeasured Outcomes

MIN ZENG*; Department of Statistics and Finance, School of Management, University of Science and Technology of China; CN
Hong Zhang; University of Science and Technology of China; CN

Outcome-dependent sampling designs are extensively utilized in various scientific disciplines, including epidemiology, ecology, and economics, with retrospective case-control studies being a specific example of such designs. Additionally, if the outcome used for sample selection is also mismeasured, then accurate estimation of the average treatment effect (ATE) can be even more challenging. To our knowledge, no existing method can be used to address these two issues simultaneously. In fil this gap, we propose a novel method for estimating ATE in the context of generalized linear models. A computationally simple estimator is developed based on the method of estimating equations, which is shown to be consistent under some regularity conditions. To relax the model assumption, we also consider generalized additive models and propose to estimate ATE with penalized B-splines and establish asymptotic properties for the proposed estimator. Our methods are evaluated through extensively simulation studies and the application to a semi-synthetic dataset based on the UK Biobank, with alcohol intake as the treatment and gout as the outcome.

Track: General Papers [Regular, In-person]
The InterModel Vigorish for Model Comparison in Confirmatory Factor Analysis with Binary Outcomes

Lijin Zhang*; Graduate School of Education, Stanford University; US
Benjamin Domingue; Graduate School of Education, Stanford University; US

Confirmatory factor analysis (CFA) has been widely used to assess the fit of a theoretical measurement model to observed data. Here, we introduce a novel index, InterModel Vigorish (IMV), which measures the value (which is expressed in units that can be thought of as dollar amounts) of using one model over another based on prediction accuracy. We extend the IMV into CFA models with binary outcomes and demonstrate that IMV provides a unique perspective for model comparison. Three simulation studies were conducted to evaluate its effectiveness in model selection and compare IMV with traditional fitting indices. Results showed that IMV provides both model-level and item-level information for detecting model misspecification. It is also not sensitive to changes in sample size. The IMV and traditional fitting indices differ in what they evaluate: traditional indices (e.g., CFI, TLI) focus on the fit to the current dataset, while IMV focuses on model predictions and can penalize overfitted models. An empirical

analysis illustrates utilization of the IMV in practice. The IMV has practical implications for researchers and practitioners using CFA models and can be extended to structural models in future research.

Track: General Papers [Regular, In-person]
scAnnotate: an automated cell type annotation tool for single-cell RNA-sequencing data

Xuekui Zhang*; University of Victoria; CA

Single-cell RNA-sequencing (scRNA-seq) technology enables researchers to investigate a genome at the cellular level with unprecedented resolution. An organism consists of a heterogeneous collection of cell types, each of which plays a distinct role in various biological processes. Hence, the first step of scRNA-seq data analysis is often to distinguish cell types so they can be investigated separately. Researchers have recently developed several automated cell type annotation tools, requiring neither biological knowledge nor subjective human decisions. Dropout is a crucial characteristic of scRNA-seq data widely used in differential expression analysis. However, no current cell annotation method explicitly utilizes dropout information. Fully utilizing dropout information motivated this work.

We present scAnnotate, a cell annotation tool that fully utilizes dropout information. We model every gene's marginal distribution using a mixture model, which describes both the dropout proportion and the distribution of the non-dropout expression levels. Then, using an ensemble machine learning approach, we combine the mixture models of all genes into a single model for cell type annotation. This combining approach can avoid estimating numerous parameters in the high-dimensional joint distribution of all genes. Using 14 real scRNA-seq datasets, we demonstrate that scAnnotate is competitive against nine existing annotation methods. Furthermore, because of its distinct modelling strategy, scAnnotate's misclassified cells differ greatly from competitor methods. This suggests using scAnnotate together with other methods could further improve annotation accuracy.

Track: Data Science in Humanities and Social Sciences [Regular, In-person]
Statistical power for linear and quadratic growth curve models with ignorable and non-ignorable missing data

Zhiyong Zhang*; University of Notre Dame Notre Dame, IN 46556 USA; US

The present study investigates how missing completely at random (MCAR), missing at random (MAR), and missing not at random (MNAR) missing data affect parameter bias, the likelihood ratio test, type I error, and power in detecting linear change and treatment effect on quadratic growth. Based on four simulation studies, it is demonstrated that MNAR may cause parameter bias and incorrect likelihood ratio test statistics although MAR and MCAR have no effect. In addition, the likelihood ratio test yields correct Type I errors under MAR and MCAR. While any missing data reduce power, the reduction gets larger as the ratio of MAR to MCAR increases,

irrespective of the overall rate of missingness. Instructions on how to conduct statistical power analysis are provided.

Track: General Papers [Speed, In-person]
L0-regularized high-dimensional mediation analysis

Saijun Zhao*; University of Science and Technology of China; CN
Hong Zhang; University of Science and Technology of China; CN

Mediation analysis is a widely-used tool in various scientific fields, such as psychology, economics, social sciences, and health sciences. As data collection techniques advance, there has been increasing interest in developing high-dimensional mediation analysis methods that employ various regularization techniques. Despite this, the method based on L0 regularization remains unexplored. To address this gap, we present a novel high-dimensional mediation analysis approach, referred to as the L0-regularization High-Dimensional Mediation Analysis (L0HMA) method, under the linear mediation model. Furthermore, we extend the proposed L0HMA method to survival models, including the Cox proportional hazards model and accelerated failure time model. The simulation results demonstrate that our L0HMA method owns good performance in mediator selection and mediation effect estimation. Finally, we apply the L0HMA method to a real dataset to illustrate its practical use.

Track: Data Science in Humanities and Social Sciences [Regular, In-person]
Assessing relative importance of multiple mediators

Xun Zhu; East China Normal University; CN
Xin Gu*; East China Normal University; CN

Mediation analysis is widely used in social sciences to identify the relationship between independent and dependent variables through mediators. Assessing the relative importance of mediators in parallel mediation models can help researchers better understand mediation effects and guide interventions. The traditional coefficient-based measures of effect size may reach problematic results for merely focusing on partial effects. This study develops a method of measuring the importance of multiple mediators and presents new measures of mediation effect in terms of relative importance. Three $R2$ measures of indirect effect proposed by MacKinnon (2008) in simple mediation models are extended to multiple mediation models. Dominance analysis, a popular method for relative importance, is applied to decompose the $R2$ indirect effect and attribute it to each mediator. Both frequentist and Bayesian methods are used to make statistical inference for the importance of multiple mediators. Simulation studies show the performance of the new measures and their inference. A real data example illustrates how to assess relative importance in multiple mediation models.