Ivan Jacob Agaloos Pesigan

January 10, 2026

References

Aalen et al.: Can we believe the DAGs? A comment on the relationship between causal DAGs and mechanisms   **Aalen-Roysland-Gran-etal-2016**

Odd O. Aalen, Kjetil Røysland, Jon Michael Gran, Roger Kouyos, et al. "Can we believe the DAGs? A comment on the relationship between causal DAGs and mechanisms". In: *Statistical Methods in Medical Research* 25.5 (July 2016), pp. 2294–2314. ISSN: 1477-0334. DOI: 10.1177/0962280213520436.

Abstract: Directed acyclic graphs (DAGs) play a large role in the modern approach to causal inference. DAGs describe the relationship between measurements taken at various discrete times including the effect of interventions. The causal mechanisms, on the other hand, would naturally be assumed to be a continuous process operating over time in a cause-effect fashion. How does such immediate causation, that is causation occurring over very short time intervals, relate to DAGs constructed from discrete observations? We introduce a time-continuous model and simulate discrete observations in order to judge the relationship between the DAG and the immediate causal model. We find that there is no clear relationship; indeed the Bayesian network described by the DAG may not relate to the causal model. Typically, discrete observations of a process will obscure the conditional dependencies that are represented in the underlying mechanistic model of the process. It is therefore doubtful whether DAGs are always suited to describe causal relationships unless time is explicitly considered in the model. We relate the issues to mechanistic modeling by using the concept of local (in)dependence. An example using data from the Swiss HIV Cohort Study is presented.

Odd O. Aalen, Kjetil Røysland, Jon Michael Gran, and Bruno Ledergerber. "Causality, mediation and time: A dynamic viewpoint". In: *Journal of the Royal Statistical Society. Series A (Statistics in Society)* 175.4 (2012), pp. 831–861. ISSN: 09641998, 1467985X. DOI: 10.1111/j.1467-985X.2011.01030.x.

Abstract: Time dynamics are often ignored in causal modelling. Clearly, causality must operate in time and we show how this corresponds to a mechanistic, or system, understanding of causality. The established counterfactual definitions of direct and indirect effects depend on an ability to manipulate the mediator which may not hold in practice, and we argue that a mechanistic view may be better. Graphical representations based on local independence graphs and dynamic path analysis are used to facilitate communication as well as providing an overview of the dynamic relations 'at a glance'. The relationship between causality as understood in a mechanistic and in an interventionist sense is discussed. An example using data from the Swiss HIV Cohort Study is presented.

Janne K. Adolf et al. "Optimal sampling rates for reliable continuous-time first-order autoregressive and vector autoregressive modeling". In: *Psychological Methods* 26.6 (Dec. 2021), pp. 701–718. ISSN: 1082-989X. DOI: 10.1037/met0000398.

Abstract: Autoregressive and vector autoregressive models are a driving force in current psychological research. In affect research they are, for instance, frequently used to formalize affective processes and estimate affective dynamics. Discrete-time model variants are most commonly used, but continuous-time formulations are gaining popularity, because they can handle data from longi-

tudinal studies in which the sampling rate varies within the study period, and yield results that can be compared across data sets from studies with different sampling rates. However, whether and how the sampling rate affects the quality with which such continuous-time models can be estimated, has largely been ignored in the literature. In the present article, we show how the sampling rate affects the estimation reliability (i.e., the standard errors of the parameter estimators, with smaller values indicating higher reliability) of continuous-time autoregressive and vector autoregressive models. Moreover, we determine which sampling rates are optimal in the sense that they lead to standard errors of minimal size (subject to the assumption that the models are correct). Our results are based on the theories of optimal design and maximum likelihood estimation. We illustrate them making use of data from the COGITO Study. We formulate recommendations for study planning, and elaborate on strengths and limitations of our approach.

Alwin et al.: The decomposition of effects in path analysis Alwin-Hauser-1975

Abstract: This paper is about the logic of interpreting recursive causal theories in sociology. We review the distinction between associations and effects and discuss the decomposition of effects into direct and indirect components. We then describe a general method for decomposing effects into their components by the systematic application of ordinary least squares regression. The method involves successive computation of reduced-form equations, beginning with an equation containing only exogenous variables, then computing equations which add intervening variables in sequence from cause to effect. This generates all the information required to decompose effects into their various direct and indirect parts. This method is a substitute for the often more cumbersome computation of indirect effects from the structural coefficients (direct effects) of the causal model. Finally, we present a way of summarizing this information in tabular form and illustrate the procedures using an empirical example.

Andrews: Heteroskedasticity and autocorrelation consistent covariance matrix estimation
<div align="right">Andrews-1991</div>

Abstract: This paper is concerned with the estimation of covariance matrices in the presence of heteroskedasticity and autocorrelation of unknown forms. Currently available estimators that are designed for this context depend upon the choice of a lag truncation parameter and a weighting scheme. Results in the literature provide a condition on the growth rate of the lag truncation parameter as $T \rightarrow \infty$ that is sufficient for consistency. No results are available, however, regarding the choice of lag truncation parameter for a fixed sample size, regarding data-dependent automatic lag truncation parameters, or regarding the choice of weighting scheme. In consequence, available estimators are not entirely operational and the relative merits of the estimators are unknown. This paper addresses these problems. The asymptotic truncated mean squared errors of estimators in a given class are determined and compared. Asymptotically optimal kernel/weighting scheme and bandwidth/lag truncation parameters are obtained using an asymptotic truncated mean squared error criterion. Using these results, data-dependent automatic bandwidth/lag truncation parameters are introduced. The finite sample properties of the estimators are analyzed via Monte Carlo simulation.

Andrews: Inconsistency of the bootstrap when a parameter is on the boundary of the parameter space
<div align="right">Andrews-2000</div>

Andrews et al.: An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator **Andrews-Monahan-1992**

Donald W. K. Andrews and J. Christopher Monahan. "An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator". In: *Econometrica* 60.4 (July 1992), p. 953. DOI: 10.2307/2951574.

Antonakis et al.: On ignoring the random effects assumption in multilevel models: Review, critique, and recommendations **Antonakis-Bastardoz-Ronkko-2019**

John Antonakis, Nicolas Bastardoz, and Mikko Rönkkö. "On ignoring the random effects assumption in multilevel models: Review, critique, and recommendations". In: *Organizational Research Methods* 24.2 (Oct. 2019), pp. 443–483. ISSN: 1552-7425. DOI: 10.1177/1094428119877457.

Abstract: Entities such as individuals, teams, or organizations can vary systematically from one another. Researchers typically model such data using multilevel models, assuming that the random effects are uncorrelated with the regressors. Violating this testable assumption, which is often ignored, creates an endogeneity problem thus preventing causal interpretations. Focusing on two-level models, we explain how researchers can avoid this problem by including cluster means of the Level 1 explanatory variables as controls; we explain this point conceptually and with a large-scale simulation. We further show why the common practice of centering the predictor variables is mostly unnecessary. Moreover, to examine the state of the science, we reviewed 204 randomly drawn articles from macro and micro organizational science and applied psychology journals, finding that only 106 articles—with a slightly higher proportion from macro-oriented fields—properly deal with the random effects assumption. Alarmingly, most models also failed on the usual exogeneity requirement of the regressors, leaving only 25 mostly macro-level articles that potentially reported trustworthy multilevel estimates. We offer a set of practical recommendations for researchers to model multilevel data appropriately.

Appelbaum et al.: Journal article reporting standards for quantitative research in psychology: The APA Publications and Communications Board task force report

Appelbaum-Cooper-Kline-etal-2018

Mark Appelbaum et al. "Journal article reporting standards for quantitative research in psychology: The APA Publications and Communications Board task force report". In: *American Psychologist* 73.7 (Oct. 2018), pp. 947–947. ISSN: 0003-066X. DOI: 10.1037/amp0000389.

Abstract: Following a review of extant reporting standards for scientific publication, and reviewing 10 years of experience since publication of the first set of reporting standards by the American Psychological Association (APA; APA Publications and Communications Board Working Group on Journal Article Reporting Standards, 2008), the APA Working Group on Quantitative Research Reporting Standards recommended some modifications to the original standards. Examples of modifications include division of hypotheses, analyses, and conclusions into 3 groupings (primary, secondary, and exploratory) and some changes to the section on meta-analysis. Several new modules are included that report standards for observational studies, clinical trials, longitudinal studies, replication studies, and N-of-1 studies. In addition, standards for analytic methods with unique characteristics and output (structural equation modeling and Bayesian analysis) are included. These proposals were accepted by the Publications and Communications Board of APA and supersede the standards included in the 6th edition of the Publication Manual of the American Psychological Association (APA, 2010).

Arbuckle: Full information estimation in the presence of incomplete data

Arbuckle-1996

James L. Arbuckle. "Full information estimation in the presence of incomplete data". In: *Advanced structural equation modeling.* Ed. by George A. Marcoulides and Randall E. Schumacker. 1996. DOI: 10.4324/9781315827414.

Arbuckle: Amos 27.0 user's guide   **Arbuckle-2020**

James L. Arbuckle. *Amos 27.0 user's guide.* Chicago: IBM SPSS, 2020.

Arbuckle: Amos 28.0 user's guide   **Arbuckle-2021**

James L. Arbuckle. *Amos 28.0 user's guide.* Chicago: IBM SPSS, 2021.

Arnett: The developmental context of substance use in emerging adulthood

Arnett-2005

Jeffrey Jensen Arnett. "The developmental context of substance use in emerging adulthood". In: *Journal of Drug Issues* 35.2 (Apr. 2005), pp. 235–254. ISSN: 1945-1369. DOI: 10.1177/002204260503500202.

Abstract: The theory of emerging adulthood has been proposed as a way of conceptualizing the developmental characteristics of young people between the ages of 18 and 25. Here, the theory is applied to explaining the high rates of substance use in this age group. Specifically, five developmentally distinctive features of emerging adulthood are proposed: the age of identity explorations, the age of instability, the age of self-focus, the age of feeling in-between, and the age of possibilities. Then, each of these features is applied to an explanation of drug use in emerging adulthood.

Arnold: Stochastic differential equations: Theory and applications   **Arnold-1974**

Ludwig Arnold. *Stochastic differential equations: Theory and applications.* A Wiley-Interscience publication. New York, NY: Wiley, 1974. 228 pp. ISBN: 9780471033592.

Aroian: The probability function of the product of two normally distributed variables

Aroian-1947

Leo A. Aroian. "The probability function of the product of two normally distributed variables".

Abstract: Let x and y follow a normal bivariate probability function with means $\bar{X}, \bar{Y}$, standard deviations σ_1, σ_2, respectively, r the coefficient of correlation, and $\rho_1 = \bar{X}/\sigma_1, \rho_2 = \bar{Y}/\sigma_2$. Professor C. C. Craig [1] has found the probability function of $z = xy/\sigma_1\sigma_2$ in closed form as the difference of two integrals. For purposes of numerical computation he has expanded this result in an infinite series involving powers of z, ρ_1, ρ_2, and Bessel functions of a certain type; in addition, he has determined the moments, semin-variants, and the moment generating function of z. However, for ρ_1 and ρ_2 large, as Craig points out, the series expansion converges very slowly. Even for ρ_1 and ρ_2 as small as 2, the expansion is unwieldy. We shall show that as ρ_1 and $\rho_2 \to \infty$, the probability function of z approaches a normal curve and in case $r = 0$ the Type III function and the Gram-Charlier Type A series are excellent approximations to the z distribution in the proper region. Numerical integration provides a substitute for the infinite series wherever the exact values of the probability function of z are needed. Some extensions of the main theorem are given in section 5 and a practical problem involving the probability function of z is solved.

Ash et al.: Sensitivity, specificity, and tolerability of the BACTrack Skyn compared to other alcohol monitoring approaches among young adults in a field-based setting
Ash-Gueorguieva-Barnett-etal-2022

Abstract: Background: There is a need for novel alcohol biosensors that are accurate, able to detect alcohol concentration close in time to consumption, and feasible and acceptable for many clinical and research applications. We evaluated the field accuracy and tolerability of novel (BACTrack Skyn)

and established (Alcohol Monitoring Systems SCRAM CAM) alcohol biosensors. Methods: The sensor and diary data were collected in a larger study of a biofeedback intervention and compared observationally in the present sub-study. Participants (high-risk drinkers, 40% female; median age 21) wore both Skyn and SCRAM CAM sensors for 1-6 days and were instructed to drink as usual. Data from the first cohort of participants ($N = 27$; 101 person-days) were used to find threshold values of transdermal alcohol that classified each day as meeting or not meeting defined levels of drinking (heavy, above-moderate, any). These values were used to develop scoring metrics that were subsequently tested using the second cohort ($N = 20$; 57 person-days). Data from both biosensors were compared to mobile diary self-report to evaluate sensitivity and specificity in relation to a priori standards established in the literature. Results: Skyn classification rules for Cohort #1 within 3 months of device shipment showed excellent sensitivity for heavy drinking (94%) and exceeded expectations for above-moderate and any drinking (78% and 69%, respectively), while specificity met expectations (91%). However, classification worsened when Cohort #1 devices $\geq$ 3 months from shipment were tested (area under curve for receiver operator characteristic 0.87 vs. 0.79) and the derived classification threshold when applied to Cohort #2 was inadequately specific (70%). Skyn tolerability metrics were excellent and exceeded the SCRAM CAM ($p \leq 0.001$). Conclusions: Skyn tolerability was favorable and accuracy rules were internally derivable but did not yield useful scoring metrics going forward across device lots and months of usage.

Tihomir Asparouhov et al.: Dynamic structural equation models

Asparouhov-Hamaker-Muthen-2017

Abstract: This article presents dynamic structural equation modeling (DSEM), which can be used to study the evolution of observed and latent variables as well as the structural equation models over time. DSEM is suitable for analyzing intensive longitudinal data where observations from multiple

individuals are collected at many points in time. The modeling framework encompasses previously published DSEM models and is a comprehensive attempt to combine time-series modeling with structural equation modeling. DSEM is estimated with Bayesian methods using the Markov chain Monte Carlo Gibbs sampler and the Metropolis-Hastings sampler. We provide a detailed description of the estimation algorithm as implemented in the Mplus software package. DSEM can be used for longitudinal analysis of any duration and with any number of observations across time. Simulation studies are used to illustrate the framework and study the performance of the estimation method. Methods for evaluating model fit are also discussed.

Tihomir Asparouhov et al.: Latent variable centering of predictors and mediators in multilevel and time-series models   **Asparouhov-Muthen-2018**

Tihomir Asparouhov and Bengt Muthén. "Latent variable centering of predictors and mediators in multilevel and time-series models". In: *Structural Equation Modeling: A Multidisciplinary Journal* 26.1 (Sept. 2018), pp. 119–142. ISSN: 1532-8007. DOI: 10.1080/10705511.2018.1511375.

Tihomir Asparouhov et al.: Multiple imputation with Mplus

Asparouhov-Muthen-2022

Tihomir Asparouhov and Bengt O. Muthén. *Multiple imputation with Mplus*. Tech. rep. http://www.statmodel.com, 2022. URL: http://www.statmodel.com/download/Imputations7.pdf.

Babor et al.: The Alcohol Use Disorders Identification Test: Guidelines for use in primary care   **Babor-HigginsBiddle-Saunders-etal-2001**

Thomas F. Babor et al. *The Alcohol Use Disorders Identification Test: Guidelines for use in primary care*. 2nd ed. Department of Mental Health and Substance Dependence, World Health Organization, 2001.

Abstract: This manual introduces the AUDIT, the Alcohol Use Disorders Identification Test, and describes how to use it to identify persons with hazardous and harmful patterns of alcohol consumption. The AUDIT was developed by the World Health Organization (WHO) as a simple method of screening for excessive drinking and to assist in brief assessment. It can help in identifying excessive drinking as the cause of the presenting illness. It also provides a framework for intervention to help hazardous and harmful drinkers reduce or cease alcohol consumption and thereby avoid the harmful consequences of their drinking. The first edition of this manual was published in 1989 (Document No. WHO/MNH/DAT/89.4) and was subsequently updated in 1992 (WHO/PSA/92.4). Since that time it has enjoyed widespread use by both health workers and alcohol researchers. With the growing use of alcohol screening and the international popularity of the AUDIT, there was a need to revise the manual to take into account advances in research and clinical experience. This manual is written primarily for health care practitioners, but other professionals who encounter persons with alcohol-related problems may also find it useful. It is designed to be used in conjunction with a companion document that provides complementary information about early intervention procedures, entitled "Brief Intervention for Hazardous and Harmful Drinking: A Manual for Use in Primary Care". Together these manuals describe a comprehensive approach to screening and brief intervention for alcohol-related problems in primary health care.

Baker et al.: Addiction motivation reformulated: An affective processing model of negative reinforcement   Baker-Piper-McCarthy-etal-2004

Abstract: This article offers a reformulation of the negative reinforcement model of drug addiction and proposes that the escape and avoidance of negative affect is the prepotent motive for addictive drug use. The authors posit that negative affect is the motivational core of the withdrawal syndrome and argue that, through repeated cycles of drug use and withdrawal, addicted organisms learn to

detect interoceptive cues of negative affect preconsciously. Thus, the motivational basis of much drug use is opaque and tends not to reflect cognitive control. When either stressors or abstinence causes negative affect to grow and enter consciousness, increasing negative affect biases information processing in ways that promote renewed drug administration. After explicating their model, the authors address previous critiques of negative reinforcement models in light of their reformulation and review predictions generated by their model.

Bakk et al.: Relating latent class membership to external variables: An overview
Bakk-Kuha-2020

Zsuzsa Bakk and Jouni Kuha. "Relating latent class membership to external variables: An overview". In: *British Journal of Mathematical and Statistical Psychology* 74.2 (Nov. 2020), pp. 340–362. ISSN: 2044-8317. DOI: 10.1111/bmsp.12227.

Abstract: In this article we provide an overview of existing approaches for relating latent class membership to external variables of interest. We extend on the work of Nylund-Gibson et al. (Structural Equation Modeling: A Multidisciplinary Journal, 2019, 26, 967), who summarize models with distal outcomes by providing an overview of most recommended modeling options for models with covariates and larger models with multiple latent variables as well. We exemplify the modeling approaches using data from the General Social Survey for a model with a distal outcome where underlying model assumptions are violated, and a model with multiple latent variables. We discuss software availability and provide example syntax for the real data examples in Latent GOLD.

Baltes et al.: History and rationale of longitudinal research   Baltes-Nesselroade-1979

Paul B. Baltes and John R. Nesselroade. "History and rationale of longitudinal research". In: *Longitudinal research in the study of behavior and development.* Ed. by John R. Nesselroade and Paul B. Baltes. New York, NY: Academic Press, 1979. ISBN: 9780125156608.

Abstract: Within the context of developmental psychology, longitudinal research is defined and reviewed from a historical perspective. Longitudinal research is shown always to include repeated-measurement methodology as the defining attribute, with individuals being the entity under study in developmental psychology. Additional characterizations vary, depending on historical and theoretical contexts. The need for longitudinal research was recognized at least as early as the nineteenth century. Terminology and specification of rationale, however, did not appear until the second or third decade of the twentieth century. The term longitudinal was initially identified in the context of age-based definitions of development. Recent decades, however, have seen an expansion of developmental theory beyond monolithic views to include age-irrelevant and multidirectional conceptions of the nature of development, particularly if a life-span perspective is taken. Such a pluralistic conception of behavioral development implies a more generic definition of longitudinal methodology than is associated with the traditional age-developmental view. Finally, it is important to recognize that the objective of longitudinal methodology is not only the descriptive identification of change. The objective includes explanatory goals also. Only recently has the unique strength of longitudinal research for explanatory efforts been recognized. In the second section of this chapter, a series of rationales for longitudinal research are outlined. These rationales are developed within the context of developmental psychology. They deal with (1) the direct identification of intraindividual change; (2) the identification of interindividual differences in intraindividual change; (3) the analysis of interrelationships in behavioral change; (4) the analysis of causes (determinants) of intraindividual change; and (5) the analysis of causes (determinants) of interindividual differences in intraindividual change. In a third section, selected issues in longitudinal designs and analysis are briefly reviewed. The need for complex longitudinal designs and control groups is emphasized to help counteract the rather widespread assumption that simple longitudinal studies are invariably sufficient for answering developmental questions. Furthermore, general limitations on aspects of developmental research associated with the study of assigned variables such as age, sex, or cohort are outlined. These limitations place constraints on design purity and mandate the use of and familiarity with alternative quasi-experimental designs. As an example, some of the problems associated with causal analysis

involving distal (delayed, mediated) influences and the use of lagged paradigms and causal modeling are discussed.

Barker et al.: Habitual alcohol seeking: Modeling the transition from casual drinking to addiction   Barker-Taylor-2014

Abstract: The transition from goal-directed actions to habitual ethanol seeking models the development of addictive behavior that characterizes alcohol use disorders. The progression to habitual ethanol-seeking behavior occurs more rapidly than for natural rewards, suggesting that ethanol may act on habit circuit to drive the loss of behavioral flexibility. This review will highlight recent research that has focused on the formation and expression of habitual ethanol seeking, and the commonalities and distinctions between ethanol and natural reward-seeking habits, with the goal of highlighting important, understudied research areas that we believe will lead toward the development of novel treatment and prevention strategies for uncontrolled drinking.

Barlow et al.: Toward a unified treatment for emotional disorders - republished article   Barlow-Allen-Choate-2016

Abstract: Over 40 years of development of cognitive behavioral approaches to treating anxiety and related emotional disorders have left us with highly efficacious treatments that are increasingly widely accepted. Nevertheless, these manualized protocols have become numerous and somewhat complex, restricting effective training and dissemination. Deepening understanding of the nature

of emotional disorders reveals that commonalities in etiology and latent structure among these disorders supercedes differences. This suggests the possibility of distilling a set of psychological procedures that would comprise a unified intervention for emotional disorders. Based on theory and data emerging from the fields of learning, emotional development and regulation, and cognitive science, we identify three fundamental therapeutic components relevant to the treatment of emotional disorders generally. These three components include (a) altering antecedent cognitive reappraisals; (b) preventing emotional avoidance; and (c) facilitating action tendencies not associated with the emotion that is dysregulated. This treatment takes place in the context of provoking emotional expression (emotional exposure) through situational, internal, and somatic (interoceptive cues), as well as through standard mood-induction exercises, and differs from patient to patient only in the situational cues and exercises utilized. Theory and rationale supporting this new approach are described along with some preliminary experience with the protocol. This unified treatment may represent a more efficient and possibly a more effective strategy in treating emotional disorders, pending further evaluation.

Barnard et al.: Nonparametric standard errors and confidence intervals: Discussion   Barnard-Collins-Farewell-etal-1981

George A. Barnard et al. "Nonparametric standard errors and confidence intervals: Discussion". In: *The Canadian Journal of Statistics / La Revue Canadienne de Statistique* 9.2 (1981), pp. 158–170. DOI: 10.2307/3314609.

Barnett: Alcohol sensors and their potential for improving clinical care   Barnett-2014

Nancy P. Barnett. "Alcohol sensors and their potential for improving clinical care". In: *Addiction* 110.1 (Dec. 2014), pp. 1–3. ISSN: 1360-0443. DOI: 10.1111/add.12764.

Abstract: Alcohol sensors are used successfully to monitor alcohol offenders in criminal justice, but their potential clinical applications with other populations are untapped. Sensors may im-

prove provider understanding about patients' patterns of alcohol use, augment current treatment approaches and identify when alcohol may be interfering with the treatment of other conditions.

Baron et al.: The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations   Baron-Kenny-1986

Reuben M. Baron and David A. Kenny. "The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations". In: *Journal of Personality and Social Psychology* 51.6 (1986), pp. 1173–1182. DOI: 10.1037/0022-3514.51.6.1173.

Abstract: In this article, we attempt to distinguish between the properties of moderator and mediator variables at a number of levels. First, we seek to make theorists and researchers aware of the importance of not using the terms moderator and mediator interchangeably by carefully elaborating, both conceptually and strategically, the many ways in which moderators and mediators differ. We then go beyond this largely pedagogical function and delineate the conceptual and strategic implications of making use of such distinctions with regard to a wide range of phenomena, including control and stress, attitudes, and personality traits. We also provide a specific compendium of analytic procedures appropriate for making the most effective use of the moderator and mediator distinction, both separately and in terms of a broader causal system that includes both moderators and mediators.

D. J. Bauer et al.: Probing interactions in fixed and multilevel regression: Inferential and graphical techniques   Bauer-Curran-2005

Daniel J. Bauer and Patrick J. Curran. "Probing interactions in fixed and multilevel regression: Inferential and graphical techniques". In: *Multivariate Behavioral Research* 40.3 (July 2005), pp. 373–400. ISSN: 1532-7906. DOI: 10.1207/s15327906mbr4003_5.

Abstract: Many important research hypotheses concern conditional relations in which the effect of one predictor varies with the value of another. Such relations are commonly evaluated as multi-

plicative interactions and can be tested in both fixed- and random-effects regression. Often, these interactive effects must be further probed to fully explicate the nature of the conditional relation. The most common method for probing interactions is to test simple slopes at specific levels of the predictors. A more general method is the Johnson-Neyman (J-N) technique. This technique is not widely used, however, because it is currently limited to categorical by continuous interactions in fixed-effects regression and has yet to be extended to the broader class of random-effects regression models. The goal of our article is to generalize the J-N technique to allow for tests of a variety of interactions that arise in both fixed- and random-effects regression. We review existing methods for probing interactions, explicate the analytic expressions needed to expand these tests to a wider set of conditions, and demonstrate the advantages of the J-N technique relative to simple slopes with three empirical examples.

D. J. Bauer et al.: Conceptualizing and testing random indirect effects and moderated mediation in multilevel models: New procedures and recommendations

Bauer-Preacher-Gil-2006

Daniel J. Bauer, Kristopher J. Preacher, and Karen M. Gil. "Conceptualizing and testing random indirect effects and moderated mediation in multilevel models: New procedures and recommendations". In: *Psychological Methods* 11.2 (2006), pp. 142–163. DOI: 10.1037/1082-989x.11.2.142.

Bell et al.: Explaining fixed effects: Random effects modeling of time-series cross-sectional and panel data

Bell-Jones-2014

Andrew Bell and Kelvyn Jones. "Explaining fixed effects: Random effects modeling of time-series cross-sectional and panel data". In: *Political Science Research and Methods* 3.1 (May 2014), pp. 133–153. ISSN: 2049-8489. DOI: 10.1017/psrm.2014.7.

Abstract: This article challenges Fixed Effects (FE) modeling as the 'default' for time-series-cross-sectional and panel data. Understanding different within and between effects is crucial when choos-

ing modeling strategies. The downside of Random Effects (RE) modeling—correlated lower-level covariates and higher-level residuals—is omitted-variable bias, solvable with Mundlak's (1978a) formulation. Consequently, RE can provide everything that FE promises and more, as confirmed by Monte-Carlo simulations, which additionally show problems with Plümper and Troeger's FE Vector Decomposition method when data are unbalanced. As well as incorporating time-invariant variables, RE models are readily extendable, with random coefficients, cross-level interactions and complex variance functions. We argue not simply for technical solutions to endogeneity, but for the substantive importance of context/heterogeneity, modeled using RE. The implications extend beyond political science to all multilevel datasets. However, omitted variables could still bias estimated higher-level variable effects; as with any model, care is required in interpretation.

P. M. Bentler et al.: Covariance structures under polynomial constraints: Applications to correlation and alpha-type structural models   Bentler-Lee-1983

Abstract: This paper provides methods for the estimation of covariance structure models under polynomial constraints. Estimation is based on maximum likelihood principles under constraints, and the test statistics, parameter estimates, and standard errors are based on a statistical theory that takes into account the constraints. The approach is illustrated by obtaining statistics for the squared multiple correlation, for predictors in a standardized metric, and in the analysis of longitudinal data via old and new models having constraints that cannot be obtained by standard methods.

Peter M. Bentler: Can scientifically useful hypotheses be tested with correlations?

Peter M. Bentler. "Can scientifically useful hypotheses be tested with correlations?" In: *American Psychologist* 62.8 (2007), pp. 772–782. ISSN: 0003-066X. DOI: 10.1037/0003-066x.62.8.772.

Abstract: Historically, interesting psychological theories have been phrased in terms of correlation coefficients, which are standardized covariances, and various statistics derived from them. Methodological practice over the last 40 years, however, has suggested it is necessary to transform such theories into hypotheses on covariances and statistics derived from them. This complication turns out to be unnecessary, because the methodology now exists to test hypotheses on latent structures of correlations directly. Two examples are given. Limitations of correlation structures are also noted.

Beran: The impact of the bootstrap on statistical algorithms and theory

Rudolf Beran. "The impact of the bootstrap on statistical algorithms and theory". In: *Statistical Science* 18.2 (May 2003). DOI: 10.1214/ss/1063994972.

Abstract: Bootstrap ideas yield remarkably effective algorithms for realizing certain programs in statistics. These include the construction of (possibly simultaneous) confidences sets and tests in classical models for which exact or asymptotic distribution theory is intractable. Success of the bootstrap, in the sense of doing what is expected under a probability model for data, is not universal. Modifications to Efron's definition of the bootstrap are needed to make the idea work for modern procedures that are not classically regular.

Bergstrom: Continuous time stochastic models and issues of aggregation over time

Bergstrom-1984

A. R. Bergstrom. "Continuous time stochastic models and issues of aggregation over time". In: *Handbook of Econometrics*. Ed. by Zvi Griliches and Michael D. Intriligator. Vol. 2. Amsterdam, 1984.

Bernardo et al.: Pathways from collectivist coping to life satisfaction among Chinese: The roles of locus-of-hope

Bernardo-Wang-Pesigan-etal-2017

Abstract: Collectivist coping styles describe approaches to coping in collectivist cultures, but there is not much research on these coping styles and their relationship to well-being. We propose that two collectivistic coping styles (acceptance/reframing/striving and family support) contribute to life satisfaction by drawing from both personal and relational resources, and these resources are reflected in the distinct roles of internal and external-family loci-of-hope. Chinese students completed scales on collectivist coping, locus-of-hope, and life satisfaction. Path analysis of three models indicates that the most parsimonious model describes distinct pathways where acceptance/reframing/striving relates to life satisfaction through internal locus-of-hope and family support relates to life satisfaction through external-family locus-of-hope. The use of specific collectivist coping styles and reports of life satisfaction involves distinct personal and relational resources for coping that are valued in collectivist societies.

Abstract: Theoretical models specifying indirect or mediated effects are common in the social sciences. An indirect effect exists when an independent variable's influence on the dependent variable is mediated through an intervening variable. Classic approaches to assessing such mediational hypotheses (Baron & Kenny, 1986; Sobel, 1982) have in recent years been supplemented by computationally intensive methods such as bootstrapping, the distribution of the product methods, and hierarchical Bayesian Markov chain Monte Carlo (MCMC) methods. These different approaches for assessing mediation are illustrated using data from Dunn, Biesanz, Human, and Finn (2007). However, little is known about how these methods perform relative to each other, particularly in more challenging situations, such as with data that are incomplete and/or nonnormal. This article presents an extensive Monte Carlo simulation evaluating a host of approaches for assessing mediation. We examine Type I error rates, power, and coverage. We study normal and nonnormal data as well as complete and incomplete data. In addition, we adapt a method, recently proposed in statistical literature, that does not rely on confidence intervals (CIs) to test the null hypothesis of no indirect effect. The results suggest that the new inferential method–the partial posterior p value–slightly outperforms existing ones in terms of maintaining Type I error rates while maximizing power, especially with incomplete data. Among confidence interval approaches, the bias-corrected accelerated (BCa) bootstrapping approach often has inflated Type I error rates and inconsistent coverage and is not recommended. In contrast, the bootstrapped percentile confidence interval and the hierarchical Bayesian MCMC method perform best overall, maintaining Type I error rates, exhibiting reasonable power, and producing stable and accurate coverage rates.

Blanca et al.: Skewness and kurtosis in real data samples

Maria J. Blanca et al. "Skewness and kurtosis in real data samples". In: *Methodology* 9.2 (May 2013), pp. 78–84. DOI: 10.1027/1614-2241/a000057.

Abstract: Parametric statistics are based on the assumption of normality. Recent findings suggest that Type I error and power can be adversely affected when data are non-normal. This paper aims to assess the distributional shape of real data by examining the values of the third and fourth central moments as a measurement of skewness and kurtosis in small samples. The analysis concerned 693 distributions with a sample size ranging from 10 to 30. Measures of cognitive ability and of other psychological variables were included. The results showed that skewness ranged between -2.49 and 2.33. The values of kurtosis ranged between -1.92 and 7.41. Considering skewness and kurtosis together the results indicated that only 5.5% of distributions were close to expected values under normality. Although extreme contamination does not seem to be very frequent, the findings are consistent with previous research suggesting that normality is not the rule with real data.

Boettiger et al.: An introduction to Rocker: Docker containers for R

Carl Boettiger and Dirk Eddelbuettel. "An introduction to Rocker: Docker containers for R". In: *The R Journal* 9.2 (2017), p. 527. DOI: 10.32614/rj-2017-065.

Abstract: We describe the Rocker project, which provides a widely-used suite of Docker images with customized R environments for particular tasks. We discuss how this suite is organized, and how these tools can increase portability, scaling, reproducibility, and convenience of R users and developers.

Boker: Consequences of continuity: The hunt for intrinsic properties within parameters of dynamics in psychological processes   Boker-2002

Abstract: Notes that over 300 yrs ago Sir Isaac Newton wrote of a simple set of relations that could be used to predict the motions of objects relative to one another. The main advantage of this insight was that the relationship between the movements of the planets and stars could be predicted much more simply than with the accurate, but cumbersome Ptolemaic calculations. But perhaps the most important consequence of the acceptance of Newton's insight was that intrinsic properties such as mass could be distinguished from measurements such as weight. A similar revolution in thinking appears to be underway in the behavioral sciences. It is likely that intensive longitudinal measurement coupled with dynamical systems analyses will lead to simplified but powerful models of the evolution of psychological processes. In this case, it is reasonable to expect that a set of intrinsic psychological properties may be able to be extracted from the parameters of successful dynamical systems models. The purpose of this article is to issue an invitation to the hunt, to provide a tentative map as to where the game might likely be found, and blow a call on the hunting horn.

Bolger et al.: Diary methods: Capturing life as it is lived   Bolger-Davis-Rafaeli-2003

Abstract: In diary studies, people provide frequent reports on the events and experiences of their daily lives. These reports capture the particulars of experience in a way that is not possible using traditional designs. We review the types of research questions that diary methods are best equipped

to answer, the main designs that can be used, current technology for obtaining diary reports, and appropriate data analysis strategies. Major recent developments include the use of electronic forms of data collection and multilevel models in data analysis. We identify several areas of research opportunities: 1. in technology, combining electronic diary reports with collateral measures such as ambulatory heart rate; 2. in measurement, switching from measures based on between-person differences to those based on within-person changes; and 3. in research questions, using diaries to (a) explain why people differ in variability rather than mean level, (b) study change processes during major events and transitions, and (c) study interpersonal processes using dyadic and group diary methods.

Bolger et al.: Intensive longitudinal methods: An introduction to diary and experience sampling research   Bolger-Laurenceau-2013

Niall Bolger and Jean-Philippe Laurenceau. *Intensive longitudinal methods: An introduction to diary and experience sampling research*. Guilford Publications, 2013. ISBN: 9781462506927.

Abstract: A complete, practical guide to planning and executing an intensive longitudinal study, this book provides the tools for understanding within-subject social, psychological, and physiological processes in everyday contexts. Intensive longitudinal studies involve many repeated measurements taken on individuals, dyads, or groups, and include diary and experience sampling studies. A range of engaging, worked-through research examples with datasets are featured. Coverage includes how to: select the best intensive longitudinal design for a particular research question, model within-subject change processes for continuous and categorical outcomes, distinguish within-subject from between-subjects effects, assess the reliability of within-subject changes, assure sufficient statistical power, and more. Several end-of-chapter write-ups illustrate effective ways to present study findings for publication.

Abstract: Decomposing the total effects of one variable on another into direct and indirect effects has long been of interest to researchers who use path analysis. In this paper, I review the decomposition of effects in general structural equation models with latent and observed variables. I present the two approaches to defining total effects. One is based on sums of powers of coefficient matrices. The other defines total effects as reducedform coefficients. I show the conditions under which these two definitions are equivalent. I also compare the different types of specific indirect effects. These are the influences that are transmitted through particular variables in a model. Finally, I propose a more general definition of specific effects that includes the effects transmitted by any path or combination of paths. I also include a section on computing standard errors for all types of effects.

Bollen et al.: A general panel model with random and fixed effects: A structural equations approach   Bollen-Brand-2010

Abstract: Fixed- and random-effects models for longitudinal data are common in sociology. Their primary advantage is that they control for time-invariant omitted variables. However, analysts face several issues when they employ these models. One is the choice of which to apply; another is that FEM and REM models as usually implemented might be insufficiently flexible. For example, the effects of variables, including the latent time-invariant variable, might change over time. The latent time-invariant variable might correlate with some variables and not others. Lagged endogenous variables might be necessary. Alternatives that move beyond the classic FEM and REM models

are known, but they involve estimators and software that make these extended models difficult to implement and to compare. This article presents a general panel model that includes the standard FEM and REM as special cases. In addition, it provides a sequence of nested models that provide a richer range of models that researchers can easily compare with likelihood ratio tests and fit statistics. Furthermore, researchers can implement our general panel model and its special cases in widely available structural equation models software.

Bollen et al.: Direct and indirect effects: Classical and bootstrap estimates of variability
Bollen-Stine-1990

Abstract: The decomposition of effects in structural equation models has been of considerable interest to social scientists. Finite-sample or asymptotic results for the sampling distribution of estimators of direct effects are widely available. Statistical inferences about indirect effects have relied exclusively on asymptotic methods which assume that the limiting distribution of the estimator is normal, with a standard error derived from the delta method. We examine bootstrap procedures as another way to generate standard errors and confidence intervals and to estimate the sampling distributions of estimators of direct and indirect effects. We illustrate the classical and the bootstrap methods with three empirical examples. We find that in a moderately large sample, the bootstrap distribution of an estimator is close to that assumed with the classical and delta methods but that in small samples, there are some differences. Bootstrap methods provide a check on the classical and delta methods when the latter are applied under less than ideal conditions.

Bond et al.: Adjustments for drink size and ethanol content: New results from a self-report diary and transdermal sensor validation study

Abstract: Background: Prior studies adjusting self-reported measures of alcohol intake for drink size and ethanol (EtOH) content have relied on single-point assessments. Methods: A prospective 28-day diary study investigated magnitudes of drink-EtOH adjustments and factors associated with these adjustments. Transdermal alcohol sensor (TAS) readings and prediction of alcohol-related problems by number of drinks versus EtOH-adjusted intake were used to validate drink-EtOH adjustments. Self-completed event diaries listed up to 4 beverage types and 4 drinking events/d. Eligible volunteers had $\geq$ weekly drinking and $\geq$ 3+ drinks per occasion with $\geq$ 26 reported days and pre- and postsummary measures ($n = 220$). Event reports included drink types, sizes, brands or spirits contents, venues, drinks consumed, and drinking duration. Results: Wine drinks averaged 1.19, beer 1.09, and spirits 1.54 U.S. standard drinks (14 g EtOH). Mean-adjusted alcohol intake was 22% larger using drink size and strength (brand/EtOH concentration) data. Adjusted drink levels were larger than "raw" drinks in all quantity ranges. Individual-level drink-EtOH adjustment ratios (EtOH adjusted/unadjusted amounts) averaged across all days drinking ranged from 0.73 to 3.33 (mean 1.22). Adjustment ratio was only marginally (and not significantly) positively related to usual quantity, frequency, and heavy drinking (all $ps < 0.10$), independent of gender, age, employment, and education, but those with lower incomes (both $p < 0.01$) drank stronger/bigger drinks. Controlling for raw number of drinks and other covariates, degree of adjustment independently predicted alcohol dependence symptoms ($p < 0.01$) and number of consequences ($p < 0.05$). In 30 respondents with sufficiently high-quality TAS readings, higher correlations ($p = 0.04$) were found between the adjusted versus the raw drinks/event and TAS areas under the curve. Conclusions: Absent drink size and strength data, intake assessments are downward biased by at least 20%.

Between-subject variation in typical drink content and pour sizes should be addressed in treatment and epidemiological research.

Boos: Introduction to the bootstrap world   Boos-2003

Abstract: The bootstrap has made a fundamental impact on how we carry out statistical inference in problems without analytic solutions. This fact is illustrated with examples and comments that emphasize the parametric bootstrap and hypothesis testing.

Bou et al.: Univariate versus multivariate modeling of panel data: Model specification and goodness-of-fit testing   Bou-Satorra-2017

Abstract: Two approaches are commonly in use for analyzing panel data: the univariate, which arranges data in long format and estimates just one regression equation; and the multivariate, which arranges data in wide format, and simultaneously estimates a set of regression equations. Although technical articles relating the two approaches exist, they do not seem to have had an impact in organizational research. This article revisits the connection between the univariate and multivariate approaches, elucidating conditions under which they yield the same—or similar—results, and discusses their complementariness. The article is addressed to applied researchers. For those familiar only with the univariate approach, it contributes with conceptual simplicity on goodness-of-fit testing and a variety of tests for misspecification (Hausman test, heteroscedasticity, autocorrelation, etc.), and simplifies expanding the model to time-varying parameters, dynamics, measurement error, and so on. For all practitioners, the comparative and side-by-side analyses of

the two approaches on two data sets—demonstration data and empirical data with missing values—contributes to broadening their perspective of panel data modeling and expanding their tools for analyses. Both univariate and multivariate analyses are performed in Stata and R.

Bradley: Robustness?   Bradley-1978

Abstract: The actual behaviour of the probability of a Type I error under assumption violation is quite complex, depending upon a wide variety of interacting factors. Yet allegations of robustness tend to ignore its highly particularistic nature and neglect to mention important qualifying conditions. The result is often a vast overgeneralization which nevertheless is difficult to refute since a standard quantitative definition of what constitutes robustness does not exist. Yet under any halfway reasonable quantitative definition, many of the most prevalent claims of robustness would be demonstrably false. Therefore robustness is a highly questionable concept.

Bringmann et al.: Assessing temporal emotion dynamics using networks
Bringmann-Pe-Vissers-etal-2016

Abstract: Multivariate psychological processes have recently been studied, visualized, and analyzed as networks. In this network approach, psychological constructs are represented as complex systems of interacting components. In addition to insightful visualization of dynamics, a network perspective leads to a new way of thinking about the nature of psychological phenomena by offering new tools for studying dynamical processes in psychology. In this article, we explain the rationale of the network approach, the associated methods and visualization, and illustrate it using an empirical example focusing on the relation between the daily fluctuations of emotions and neuroticism. The results

suggest that individuals with high levels of neuroticism had a denser emotion network compared with their less neurotic peers. This effect is especially pronounced for the negative emotion network, which is in line with previous studies that found a denser network in depressed subjects than in healthy subjects. In sum, we show how the network approach may offer new tools for studying dynamical processes in psychology.

Bringmann et al.: A network approach to psychopathology: New insights into clinical longitudinal data   Bringmann-Vissers-Wichers-etal-2013

Abstract: In the network approach to psychopathology, disorders are conceptualized as networks of mutually interacting symptoms (e.g., depressed mood) and transdiagnostic factors (e.g., rumination). This suggests that it is necessary to study how symptoms dynamically interact over time in a network architecture. In the present paper, we show how such an architecture can be constructed on the basis of time-series data obtained through Experience Sampling Methodology (ESM). The proposed methodology determines the parameters for the interaction between nodes in the network by estimating a multilevel vector autoregression (VAR) model on the data. The methodology allows combining between-subject and within-subject information in a multilevel framework. The resulting network architecture can subsequently be analyzed through network analysis techniques. In the present study, we apply the method to a set of items that assess mood-related factors. We show that the analysis generates a plausible and replicable network architecture, the structure of which is related to variables such as neuroticism; that is, for subjects who score high on neuroticism, worrying plays a more central role in the network. Implications and extensions of the methodology are discussed.

Abstract: This edition contains a large number of additions and corrections scattered throughout the text, including the incorporation of a new chapter on state-space models. The companion diskette for the IBM PC has expanded into the software package ITSM: An Interactive Time Series Modelling Package for the PC, which includes a manual and can be ordered from Springer-Verlag. * We are indebted to many readers who have used the book and programs and made suggestions for improvements. Unfortunately there is not enough space to acknowledge all who have contributed in this way; however, special mention must be made of our prize-winning fault-finders, Sid Resnick and F. Pukelsheim. Special mention should also be made of Anthony Brockwell, whose advice and support on computing matters was invaluable in the preparation of the new diskettes. We have been fortunate to work on the new edition in the excellent environments provided by the University of Melbourne and Colorado State University. We thank Duane Boes particularly for his support and encouragement throughout, and the Australian Research Council and National Science Foundation for their support of research related to the new material. We are also indebted to Springer-Verlag for their constant support and assistance in preparing the second edition.

Abstract: Many African American youth may develop high levels of allostatic load, a measure of physiological wear and tear on the body, by developing psychosocial competence under conditions of high risk related to socioeconomic status (SES). The current study was designed to test this hypothesis, which is based on John Henryism theory. In a representative sample of 489 African American youth living in the rural South, cumulative SES-related risks and teacher-reported competence were assessed at ages 11 to 13; depressive symptoms, externalizing behavior, and allostatic load were assessed at age 19. The data revealed that rural African American preadolescents who evinced high psychosocial competence under conditions of high cumulative SES-related risk displayed low levels of adjustment problems along with high allostatic load at age 19. These results suggest that, for many rural African Americans, resilience may indeed be only "skin deep."

M. W. Browne: Asymptotically distribution-free methods for the analysis of covariance structures — Browne-1984

Michael W. Browne. "Asymptotically distribution-free methods for the analysis of covariance structures". In: *British Journal of Mathematical and Statistical Psychology* 37.1 (May 1984), pp. 62–83. DOI: 10.1111/j.2044-8317.1984.tb00789.x.

Abstract: Methods for obtaining tests of fit of structural models for covariance matrices and estimator standard error which are asymptotically distribution free are derived. Modifications to standard normal theory tests and standard errors which make them applicable to the wider class of elliptical distributions are provided. A random sampling experiment to investigate some of the proposed methods is described.

Bühlmann: Bootstraps for time series — Buehlmann-2002

Peter Bühlmann. "Bootstraps for time series". In: *Statistical Science* 17.1 (May 2002). ISSN: 0883-4237. DOI: 10.1214/ss/1023798998.

Abstract: We review and compare block, sieve and local bootstraps for time series and thereby illuminate theoretical aspects of the procedures as well as their performance on finite-sample data. Our view is selective with the intention of providing a new and fair picture of some particular aspects of bootstrapping time series. The generality of the block bootstrap is contrasted with sieve bootstraps. We discuss implementational advantages and disadvantages. We argue that two types of sieve often outperform the block method, each of them in its own important niche, namely linear and categorical processes. Local bootstraps, designed for nonparametric smoothing problems, are easy to use and implement but exhibit in some cases low performance.

Casella: Introduction to the silver anniversary of the bootstrap — Casella-2003

George Casella. "Introduction to the silver anniversary of the bootstrap". In: *Statistical Science* 18.2 (May 2003). DOI: 10.1214/ss/1063994967.

Casella et al.: Statistical inference — Casella-Berger-2002

George Casella and Robert L. Berger. *Statistical inference*. Pacific Grove, CA: Thomson Learning, 2002. ISBN: 9780534243128.

Abstract: This book builds theoretical statistics from the first principles of probability theory. Starting from the basics of probability, the authors develop the theory of statistical inference using techniques, definitions, and concepts that are statistical and are natural extensions and consequences of previous concepts. This book can be used for readers who have a solid mathematics background. It can also be used in a way that stresses the more practical uses of statistical theory, being more concerned with understanding basic statistical concepts and deriving reasonable statistical procedures for a variety of situations, and less concerned with formal optimality investigations.

Casey et al.: The adolescent brain — Casey-Jones-Hare-2008

B.J. Casey, Rebecca M. Jones, and Todd A. Hare. "The adolescent brain". In: *Annals of the New*

Abstract: Adolescence is a developmental period characterized by suboptimal decisions and actions that are associated with an increased incidence of unintentional injuries, violence, substance abuse, unintended pregnancy, and sexually transmitted diseases. Traditional neurobiological and cognitive explanations for adolescent behavior have failed to account for the nonlinear changes in behavior observed during adolescence, relative to both childhood and adulthood. This review provides a biologically plausible model of the neural mechanisms underlying these nonlinear changes in behavior. We provide evidence from recent human brain imaging and animal studies that there is a heightened responsiveness to incentives and socioemotional contexts during this time, when impulse control is still relatively immature. These findings suggest differential development of bottom-up limbic systems, implicated in incentive and emotional processing, to top-down control systems during adolescence as compared to childhood and adulthood. This developmental pattern may be exacerbated in those adolescents prone to emotional reactivity, increasing the likelihood of poor outcomes.

Castro-Schilo et al.: Comparison of nomothetic versus idiographic-oriented methods for making predictions about distal outcomes from time series data

CastroSchilo-Ferrer-2013

Abstract: We illustrate the idiographic/nomothetic debate by comparing 3 approaches to using daily self-report data on affect for predicting relationship quality and breakup. The 3 approaches included (a) the first day in the series of daily data; (b) the mean and variability of the daily series; and (c) parameters from dynamic factor analysis, a statistical model that uses all measurement

occasions to estimate the structure and dynamics of the data. Our results indicated that data from the first measurement occasion does not provide information about the couples' relationship quality or breakup 1 to 2 years later. The mean and variability of the time series, however, were more informative: females' average positive and negative affect across time was related to relationship quality, whereas males' variability in negative affect across time was predictive of breakup. The dynamic factor analysis, in turn, allowed us to extract information central to the dyadic dynamics. This information proved useful to predict relationship quality but not breakup. The importance of examining intraindividual variability and couple dynamics is highlighted.

Celeux et al.: An entropy criterion for assessing the number of clusters in a mixture model
Celeux-Soromenho-1996

Abstract: In this paper, we consider an entropy criterion to estimate the number of clusters arising from a mixture model. This criterion is derived from a relation linking the likelihood and the classification likelihood of a mixture. Its performance is investigated through Monte Carlo experiments, and it shows favorable results compared to other classical criteria.

Chassin et al.: Adolescent Substance Use
Chassin-Hussong-Beltran-2009

Abstract: This chapter will describe the prevalence and predictors of adolescent substance use and substance use problems, with particular emphasis on their relation to the developmental issues of

adolescence. Given the wide scope of this field, we do not attempt a comprehensive review. For example, we do not consider issues of substance abuse treatment (see, e.g., Brown & D'Amico, 2003; Morral, McCaffrey, Ridgeway, Mukherji, & Beighley, 2006; Waldron & Kaminer, 2004; Rowe & Liddle, 2006) or prevention (Caulkins, Liccardo, Pacula, Paddock, & Chiesa, 2002; Faggiano et al., 2005; Gates, McCambridge, Smith, & Foxcroft, 2006). Rather, we will selectively emphasize recent empirical work, as well as studies that illustrate important general themes in adolescent substance use research.

Chatfield: The analysis of time series: An introduction   Chatfield-2003

Chris Chatfield. *The analysis of time series: An introduction*. 6th ed. Chapman and Hall/CRC, July 2003. ISBN: 9780203491683. DOI: 10.4324/9780203491683.

Chen et al.: Vector autoregression, structural equation modeling, and their synthesis in neuroimaging data analysis   Chen-Daniel-Ziad-etal-2011

Gang Chen et al. "Vector autoregression, structural equation modeling, and their synthesis in neuroimaging data analysis". In: *Computers in Biology and Medicine* 41.12 (Dec. 2011), pp. 1142–1155. DOI: 10.1016/j.compbiomed.2011.09.004.

Abstract: Vector autoregression (VAR) and structural equation modeling (SEM) are two popular brain-network modeling tools. VAR, which is a data-driven approach, assumes that connected regions exert time-lagged influences on one another. In contrast, the hypothesis-driven SEM is used to validate an existing connectivity model where connected regions have contemporaneous interactions among them. We present the two models in detail and discuss their applicability to FMRI data, and their interpretational limits. We also propose a unified approach that models both lagged and contemporaneous effects. The unifying model, structural vector autoregression (SVAR), may improve statistical and explanatory power, and avoid some prevalent pitfalls that can occur when VAR and SEM are utilized separately.

Cheong et al.: Investigation of mediational processes using parallel process latent growth curve modeling

JeeWon Cheong, David P. MacKinnon, and Siek Toon Khoo. "Investigation of mediational processes using parallel process latent growth curve modeling". In: *Structural Equation Modeling: A Multidisciplinary Journal* 10.2 (Apr. 2003), pp. 238–262. DOI: 10.1207/s15328007sem1002_5.

Abstract: This study investigated a method to evaluate mediational processes using latent growth curve modeling. The mediator and the outcome measured across multiple time points were viewed as 2 separate parallel processes. The mediational process was defined as the independent variable influencing the growth of the mediator, which, in turn, affected the growth of the outcome. To illustrate modeling procedures, empirical data from a longitudinal drug prevention program, Adolescents Training and Learning to Avoid Steroids, were used. The program effects on the growth of the mediator and the growth of the outcome were examined first in a 2-group structural equation model. The mediational process was then modeled and tested in a parallel process latent growth curve model by relating the prevention program condition, the growth rate factor of the mediator, and the growth rate factor of the outcome.

Chesher et al.: The bias of a heteroskedasticity consistent covariance matrix estimator

Andrew Chesher and Ian Jewitt. "The bias of a heteroskedasticity consistent covariance matrix estimator". In: *Econometrica* 55.5 (Sept. 1987), p. 1217. DOI: 10.2307/1911269.

G. W. Cheung et al.: Testing mediation and suppression effects of latent variables

Gordon W. Cheung and Rebecca S. Lau. "Testing mediation and suppression effects of latent variables". In: *Organizational Research Methods* 11.2 (July 2007), pp. 296–325. DOI: 10.1177/1094428107300343.

Abstract: Because of the importance of mediation studies, researchers have been continuously searching for the best statistical test for mediation effect. The approaches that have been most commonly employed include those that use zero-order and partial correlation, hierarchical regression models, and structural equation modeling (SEM). This study extends MacKinnon and colleagues (MacKinnon, Lockwood, Hoffmann, West, & Sheets, 2002; MacKinnon, Lockwood, & Williams, 2004, MacKinnon, Warsi, & Dwyer, 1995) works by conducting a simulation that examines the distribution of mediation and suppression effects of latent variables with SEM, and the properties of confidence intervals developed from eight different methods. Results show that SEM provides unbiased estimates of mediation and suppression effects, and that the bias-corrected bootstrap confidence intervals perform best in testing for mediation and suppression effects. Steps to implement the recommended procedures with Amos are presented.

M. W.-L. Cheung: Comparison of approaches to constructing confidence intervals for mediating effects using structural equation models   Cheung-2007

Abstract: Mediators are variables that explain the association between an independent variable and a dependent variable. Structural equation modeling (SEM) is widely used to test models with mediating effects. This article illustrates how to construct confidence intervals (CIs) of the mediating effects for a variety of models in SEM. Specifically, mediating models with 1 mediator, 2 intermediate mediators, 2 specific mediators, and 1 mediator in 2 independent groups are illustrated. By using phantom variables (Rindskopf, 1984), a Wald CI, percentile bootstrap CI, bias-corrected bootstrap CI, and a likelihood-based CI on the mediating effect are easily constructed with some existing SEM packages, such as LISREL, Mplus, and Mx. Monte Carlo simulation studies are used to compare the coverage probabilities of these CIs. The results show that the coverage probabilities of these CIs are comparable when the mediating effect is large or when the sample size is large. However, when

the mediating effect and the sample size are both small, the bootstrap CI and likelihood-based CI are preferred over the Wald CI. Extensions of this SEM approach for future research are discussed.

M. W.-L. Cheung: A model for integrating fixed-, random-, and mixed-effects meta-analyses into structural equation modeling Cheung-2008

Mike W.-L. Cheung. "A model for integrating fixed-, random-, and mixed-effects meta-analyses into structural equation modeling". In: *Psychological Methods* 13.3 (2008), pp. 182–202. ISSN: 1082-989X. DOI: 10.1037/a0013163.

Abstract: Meta-analysis and structural equation modeling (SEM) are two important statistical methods in the behavioral, social, and medical sciences. They are generally treated as two unrelated topics in the literature. The present article proposes a model to integrate fixed-, random-, and mixed-effects meta-analyses into the SEM framework. By applying an appropriate transformation on the data, studies in a meta-analysis can be analyzed as subjects in a structural equation model. This article also highlights some practical benefits of using the SEM approach to conduct a meta-analysis. Specifically, the SEM-based meta-analysis can be used to handle missing covariates, to quantify the heterogeneity of effect sizes, and to address the heterogeneity of effect sizes with mixture models. Examples are used to illustrate the equivalence between the conventional meta-analysis and the SEM-based meta-analysis. Future directions on and issues related to the SEM-based meta-analysis are discussed.

M. W.-L. Cheung: Comparison of methods for constructing confidence intervals of standardized indirect effects Cheung-2009a

Mike W.-L. Cheung. "Comparison of methods for constructing confidence intervals of standardized indirect effects". In: *Behavior Research Methods* 41.2 (May 2009), pp. 425–438. DOI: 10.3758/brm.41.2.425.

Abstract: Mediation models are often used as a means to explain the psychological mechanisms between an independent and a dependent variable in the behavioral and social sciences. A major limitation of the unstandardized indirect effect calculated from raw scores is that it cannot be interpreted as an effect-size measure. In contrast, the standardized indirect effect calculated from standardized scores can be a good candidate as a measure of effect size because it is scale invariant. In the present article, 11 methods for constructing the confidence intervals (CIs) of the standardized indirect effects were evaluated via a computer simulation. These included six Wald CIs, three bootstrap CIs, one likelihood-based CI, and the PRODCLIN CI. The results consistently showed that the percentile bootstrap, the bias-corrected bootstrap, and the likelihood-based approaches had the best coverage probability. Mplus, LISREL, and Mx syntax were included to facilitate the use of these preferred methods in applied settings. Future issues on the use of the standardized indirect effects are discussed.

M. W.-L. Cheung: Constructing approximate confidence intervals for parameters with structural equation models   Cheung-2009b

Abstract: Confidence intervals (CIs) for parameters are usually constructed based on the estimated standard errors. These are known as Wald CIs. This article argues that likelihood-based CIs (CIs based on likelihood ratio statistics) are often preferred to Wald CIs. It shows how the likelihood-based CIs and the Wald CIs for many statistics and psychometric indexes can be constructed with the use of phantom variables (Rindskopf, 1984) in some of the current structural equation modeling (SEM) packages. The procedures to form CIs for the differences in correlation coefficients, squared multiple correlations, indirect effects, coefficient alphas, and reliability estimates are illustrated. A simulation study on the Pearson correlation is used to demonstrate the advantages of the likelihood-

based CI over the Wald CI. Issues arising from this SEM approach and extensions of this approach are discussed.

M. W.-L. Cheung: Multivariate meta-analysis as structural equation models

Abstract: Multivariate meta-analysis has become increasingly popular in the educational, social, and medical sciences. It is because the outcome measures in a meta-analysis can involve more than one effect size. This article proposes 2 mathematically equivalent models to implement multivariate meta-analysis in structural equation modeling (SEM). Specifically, this article shows how multivariate fixed-, random- and mixed-effects meta-analyses can be formulated as structural equation models. metaSEM (a free R package based on OpenMx) and Mplus are used to implement the proposed procedures. A real data set is used to illustrate the procedures. Formulating multivariate meta-analysis as structural equation models provides many new research opportunities for methodological development in both meta-analysis and SEM. Issues related to and extensions on the SEM-based meta-analysis are discussed.

M. W.-L. Cheung: Meta-analysis: A structural equation modeling approach

Abstract: Presents a novel approach to conducting meta-analysis using structural equation modeling. Structural equation modeling (SEM) and meta-analysis are two powerful statistical methods in the educational, social, behavioral, and medical sciences. They are often treated as two unre-

lated topics in the literature. This book presents a unified framework on analyzing meta-analytic data within the SEM framework, and illustrates how to conduct meta-analysis using the metaSEM package in the R statistical environment. Meta-Analysis: A Structural Equation Modeling Approach begins by introducing the importance of SEM and meta-analysis in answering research questions. Key ideas in meta-analysis and SEM are briefly reviewed, and various meta-analytic models are then introduced and linked to the SEM framework. Fixed-, random-, and mixed-effects models in univariate and multivariate meta-analyses, three-level meta-analysis, and meta-analytic structural equation modeling, are introduced. Advanced topics, such as using restricted maximum likelihood estimation method and handling missing covariates, are also covered. Readers will learn a single framework to apply both meta-analysis and SEM. Examples in R and in Mplus are included. This book will be a valuable resource for statistical and academic researchers and graduate students carrying out meta-analyses, and will also be useful to researchers and statisticians using SEM in biostatistics. Basic knowledge of either SEM or meta-analysis will be helpful in understanding the materials in this book.

M. W.-L. Cheung: Synthesizing indirect effects in mediation models with meta-analytic methods   Cheung-2021

Abstract: Aims: A mediator is a variable that explains the underlying mechanism between an independent variable and a dependent variable. The indirect effect indicates the effect from the predictor to the outcome variable via the mediator. In contrast, the direct effect represents the predictor's effort on the outcome variable after controlling for the mediator. Methods: A single study rarely provides enough evidence to answer research questions in a particular domain. Replications are generally recommended as the gold standard to conduct scientific research. When a sufficient number of studies have been conducted addressing similar research questions, a meta-analysis can be used to synthesize those studies' findings. Results: The main objective of this paper is to introduce two

frameworks to integrating studies using mediation analysis. The first framework involves calculating standardized indirect effects and direct effects and conducting a multivariate meta-analysis on those effect sizes. The second one uses meta-analytic structural equation modeling to synthesize correlation matrices and fit mediation models on the average correlation matrix. We illustrate these procedures on a real dataset using the R statistical platform. Conclusion: This paper closes with some further directions for future studies.

S. F. Cheung et al.: Improving an old way to measure moderation effect in standardized units. Cheung-Cheung-Lau-etal-2022

Abstract: Moderation effects in multiple regression, tested usually by the inclusion of a product term, are frequently investigated in health psychology. However, several issues in presenting the moderation effects in standardized units and their associated confidence intervals are commonly observed. While an old method had been proposed to standardize variables in moderated regression before fitting a moderated regression model, this method was rarely used due to inconvenience and even when used, the confidence intervals derived were biased. Here, we attempt to solve these two problems by providing a tool to conveniently conduct standardization in moderated regression without the step of standardizing the variables beforehand and to accurately form the nonparametric bootstrapping confidence intervals for this standardized measure of moderation effects. Health psychology researchers are now equipped with a tool that can be used to report and interpret standardized moderation effects correctly.

S. F. Cheung et al.: FINDOUT: Using either SPSS commands or graphical user interface to identify influential cases in structural equation modeling in AMOS

Shu Fai Cheung and Ivan Jacob Agaloos Pesigan. "FINDOUT: Using either SPSS commands or graphical user interface to identify influential cases in structural equation modeling in AMOS". In: *Multivariate Behavioral Research* 58.5 (Jan. 2023), pp. 964–968. DOI: 10.1080/00273171.2022.2148089.

Abstract: The results in a structural equation modeling (SEM) analysis can be influenced by just a few observations, called influential cases. Tools have been developed for users of R to identify them. However, similar tools are not available for AMOS, which is also a popular SEM software package. We introduce the FINDOUT toolset, a group of SPSS extension commands, and an AMOS plugin, to identify influential cases and examine how these cases influence the results. The SPSS commands can be used either as syntax commands or as custom dialogs from pull-down menus, and the AMOS plugin can be run from AMOS pull-down menu. We believe these tools can help researchers to examine the robustness of their findings to influential cases.

S. F. Cheung et al.: semlbci: An R package for forming likelihood-based confidence intervals for parameter estimates, correlations, indirect effects, and other derived parameters

Shu Fai Cheung and Ivan Jacob Agaloos Pesigan. "semlbci: An R package for forming likelihood-based confidence intervals for parameter estimates, correlations, indirect effects, and other derived parameters". In: *Structural Equation Modeling: A Multidisciplinary Journal* 30.6 (May 2023), pp. 985–999. DOI: 10.1080/10705511.2023.2183860.

Abstract: There are three common types of confidence interval (CI) in structural equation modeling (SEM): Wald-type CI, bootstrapping CI, and likelihood-based CI (LBCI). LBCI has the following advantages: (1) it has better coverage probabilities and Type I error rate compared to Wald-type

CI when the sample size is finite; (2) it correctly tests the null hypothesis of a parameter based on likelihood ratio chi-square difference test; (3) it is less computationally intensive than bootstrapping CI; and (4) it is invariant to transformations. However, LBCI is not available in many popular SEM software packages. We developed an R package, semlbci, for forming LBCI for parameters in models fitted by lavaan, a popular open-source SEM package, such that researchers have more options in forming CIs for parameters in SEM. The package supports both unstandardized and standardized estimates, derived parameters such as indirect effect, multisample models, and the robust LBCI proposed by Falk.

S. F. Cheung et al.: DIY bootstrapping: Getting the nonparametric bootstrap confidence interval in SPSS for any statistics or function of statistics (when this bootstrapping is appropriate) Cheung-Pesigan-Vong-2023

Abstract: Researchers can generate bootstrap confidence intervals for some statistics in SPSS using the BOOTSTRAP command. However, this command can only be applied to selected procedures, and only to selected statistics in these procedures. We developed an extension command and prepared some sample syntax files based on existing approaches from the Internet to illustrate how researchers can (a) generate a large number of nonparametric bootstrap samples, (b) do desired analysis on all these samples, and (c) form the bootstrap confidence intervals for selected statistics using the OMS commands. We developed these tools to help researchers apply nonparametric bootstrapping to any statistics for which this method is appropriate, including statistics derived from other statistics, such as standardized effect size measures computed from the t test results. We also discussed how researchers can extend the tools for other statistics and scenarios they encounter.

Abstract: The ability to maintain the separation between positive emotion and negative emotion in times of stress has been construed as a resilience mechanism. Emotional resiliency is particularly relevant in old age given concomitant declines in cognitive performance. In the present study, the authors examined the dynamical linkages among positive emotion, negative emotion, and cognition as individuals performed a complex cognitive task. Comparisons were made between younger (n = 63) and older (n = 52) age groups. Older adults manifested significant unidirectional coupling from negative emotion to cognitive performance; younger adults manifested significant unidirectional coupling from negative emotion to positive emotion and from cognitive performance to both positive and negative emotions. Implications for age differences in emotion regulatory strategies are discussed.

Abstract: State-space modeling techniques have been compared to structural equation modeling (SEM) techniques in various contexts but their unique strengths have often been overshadowed by their similarities to SEM. In this article, we provide a comprehensive discussion of these 2 approaches' similarities and differences through analytic comparisons and numerical simulations,

with a focus on their use in representing intraindividual dynamics and interindividual differences. To demonstrate the respective strengths and weaknesses of the 2 approaches in representing these 2 aspects, we simulated data under (a) a cross-sectional common factor model, (b) a latent difference score model with random effects in intercept and slope, and (c) a bivariate dynamic factor analysis model with auto- and cross-regression parameters. Possible ways in which SEM and state-space modeling can be utilized as complementary tools in representing human developmental and other related processes are discussed.

Chow et al.: Continuous-time dynamic models: Connections to structural equation models and other discrete-time models   Chow-Losardo-Park-etal-2023

Sy-Miin Chow, Diane Losardo, et al. "Continuous-time dynamic models: Connections to structural equation models and other discrete-time models". In: *Handbook of structural equation modeling.* Ed. by Rick H. Hoyle. 2nd ed. New York: The Guilford Press, 2023. ISBN: 9781462550722.

Chow et al.: The cusp catastrophe model as cross-sectional and longitudinal mixture structural equation models   Chow-Witkiewitz-Grasman-etal-2015

Sy-Miin Chow, Katie Witkiewitz, et al. "The cusp catastrophe model as cross-sectional and longitudinal mixture structural equation models". In: *Psychological Methods* 20.1 (Mar. 2015), pp. 142–164. ISSN: 1082-989X. DOI: 10.1037/a0038962.

Abstract: Catastrophe theory (Thom, 1972, 1993) is the study of the many ways in which continuous changes in a system's parameters can result in discontinuous changes in 1 or several outcome variables of interest. Catastrophe theory—inspired models have been used to represent a variety of change phenomena in the realm of social and behavioral sciences. Despite their promise, widespread applications of catastrophe models have been impeded, in part, by difficulties in performing model fitting and model comparison procedures. We propose a new modeling framework for testing 1 kind of catastrophe model—the cusp catastrophe model—as a mixture structural equa-

tion model (MSEM) when cross-sectional data are available; or alternatively, as an MSEM with regime-switching (MSEM-RS) when longitudinal panel data are available. The proposed models and the advantages offered by this alternative modeling framework are illustrated using 2 empirical examples and a simulation study.

Chow et al.: Dynamic factor analysis models with time-varying parameters

<div align="right">**Chow-Zu-Shifren-etal-2011**</div>

Abstract: Dynamic factor analysis models with time-varying parameters offer a valuable tool for evaluating multivariate time series data with time-varying dynamics and/or measurement properties. We use the Dynamic Model of Activation proposed by Zautra and colleagues (Zautra, Potter, & Reich, 1997) as a motivating example to construct a dynamic factor model with vector autoregressive relations and time-varying cross-regression parameters at the factor level. Using techniques drawn from the state-space literature, the model was fitted to a set of daily affect data (over 71 days) from 10 participants who had been diagnosed with Parkinson's disease. Our empirical results lend partial support and some potential refinement to the Dynamic Model of Activation with regard to how the time dependencies between positive and negative affects change over time. A simulation study is conducted to examine the performance of the proposed techniques when (a) changes in the time-varying parameters are represented using the true model of change, (b) supposedly time-invariant parameters are represented as time-varying, and (c) the time-varying parameters show discrete shifts that are approximated using an autoregressive model of differences.

John D. Clapp et al. "Examining the social ecology of a bar-crawl: An exploratory pilot study". In: *PLOS ONE* 12.9 (Sept. 2017). Ed. by Etsuro Ito, pp. 1–27. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0185238.

Abstract: Many of the problems associated with alcohol occur after a single drinking event (e.g. drink driving, assault). These acute alcohol problems have a huge global impact and account for a large percentage of unintentional and intentional injuries in the world. Nonetheless, alcohol research and preventive interventions rarely focus on drinking at the event-level since drinking events are complex, dynamic, and methodologically challenging to observe. This exploratory study provides an example of how event-level data may be collected, analyzed, and interpreted. The drinking behavior of twenty undergraduate students enrolled at a large Midwestern public university was observed during a single bar crawl event that is organized by students annually. Alcohol use was monitored with transdermal alcohol devices coupled with ecological momentary assessments and geospatial data. "Small N, Big Data" studies have the potential to advance health behavior theory and to guide real-time interventions. However, such studies generate large amounts of within subject data that can be challenging to analyze and present. This study examined how to visually display event-level data and also explored the relationship between some basic indicators and alcohol consumption.

C. Robert Cloninger. "Neurogenetic adaptive mechanisms in alcoholism". In: *Science* 236.4800 (Apr. 1987), pp. 410–416. ISSN: 1095-9203. DOI: 10.1126/science.2882604.

Abstract: Clinical, genetic, and neuropsychopharmacological studies of developmental factors in alcoholism are providing a better understanding of the neurobiological bases of personality and learning. Studies of the adopted-away children of alcoholics show that the predisposition to initiate

alcohol-seeking behavior is genetically different from susceptibility to loss of control after drinking begins. Alcohol-seeking behavior is a special case of exploratory appetitive behavior and involves different neurogenetic processes than do susceptibility to behavioral tolerance and dependence on the antianxiety or sedative effects of alcohol. Three dimensions of personality have been described that may reflect individual differences in brain systems modulating the activation, maintenance, and inhibition of behavioral responses to the effects of alcohol and other environmental stimuli. These personality traits distinguish alcoholics with different patterns of behavioral, neurophysiological, and neuropharmacological responses to alcohol.

Cochran: The χ^2 test of goodness of fit   Cochran-1952

William G. Cochran. "The χ^2 test of goodness of fit". In: *The Annals of Mathematical Statistics* 23.3 (Sept. 1952), pp. 315–345. DOI: 10.1214/aoms/1177729380.

Abstract: This paper contains an expository discussion of the chi square test of goodness of fit, intended for the student and user of statistical theory rather than for the expert. Part I describes the historical development of the distribution theory on which the test rests. Research bearing on the practical application of the test–in particular on the minimum expected number per class and the construction of classes–is discussed in Part II. Some varied opinions about the extent to which the test actually is useful to the scientist are presented in Part III. Part IV outlines a number of tests that have been proposed as substitutes for the chi square test (the ω^2 test, the smooth test, the likelihood ratio test) and Part V a number of supplementary tests (the run test, tests based on low moments, subdivision of chi square into components).

Cohen: Statistical power analysis for the behavioral sciences   Cohen-1988

Jacob Cohen. *Statistical power analysis for the behavioral sciences*. 2nd ed. Routledge, 1988. ISBN: 9780203771587. DOI: 10.4324/9780203771587.

Abstract: Statistical Power Analysis is a nontechnical guide to power analysis in research planning

that provides users of applied statistics with the tools they need for more effective analysis. The Second Edition includes:

- a chapter covering power analysis in set correlation and multivariate methods;

- a chapter considering effect size, psychometric reliability, and the efficacy of "qualifying" dependent variables and;

- expanded power and sample size tables for multiple regression/correlation.

.

Cole et al.: Empirical and conceptual problems with longitudinal trait-state models: Introducing a trait-state-occasion model   Cole-Martin-Steiger-2005

Abstract: The latent trait-state-error model (TSE) and the latent state-trait model with autoregression (LST-AR) represent creative structural equation methods for examining the longitudinal structure of psychological constructs. Application of these models has been somewhat limited by empirical or conceptual problems. In the present study, Monte Carlo analysis revealed that TSE models tend to generate improper solutions when N is too small, when waves are too few, and when occasion factor stability is either too large or too small. Mathematical analysis of the LST-AR model revealed its limitation to constructs that become more highly auto-correlated over time. The trait-state-occasion model has fewer empirical problems than does the TSE model and is more broadly applicable than is the LST-AR model.

Cole et al.: Testing mediational models with longitudinal data: Questions and tips in the use of structural equation modeling. Cole-Maxwell-2003

Abstract: R. M. Baron and D. A. Kenny (1986; see record 1987-13085-001) provided clarion conceptual and methodological guidelines for testing mediational models with cross-sectional data. Graduating from cross-sectional to longitudinal designs enables researchers to make more rigorous inferences about the causal relations implied by such models. In this transition, misconceptions and erroneous assumptions are the norm. First, we describe some of the questions that arise (and misconceptions that sometimes emerge) in longitudinal tests of mediational models. We also provide a collection of tips for structural equation modeling (SEM) of mediational processes. Finally, we suggest a series of 5 steps when using SEM to test mediational processes in longitudinal designs: testing the measurement model, testing for added components, testing for omitted paths, testing the stationarity assumption, and estimating the mediational effects.

Collins et al.: Best methods for the analysis of change: Recent advances, unanswered questions, future directions Collins-Horn-1991

Abstract: The chapters making up this book represent a rich offering of current research on the analysis of change.

Linda M. Collins and Aline Sayer, eds. *New methods for the analysis of change. [based on a conference held in 1998 at The Pennsylvania State University, a follow-up to the Los Angeles conference Best Methods for the Analysis of Change].* 2nd ed. Decade of behavior. Washington, DC: American Psychological Association, 2002. 442 pp. ISBN: 1557987548.

Collins et al.: Latent class models for stage-sequential dynamic latent variables

Linda M. Collins and Stuart E. Wugalter. "Latent class models for stage-sequential dynamic latent variables". In: *Multivariate Behavioral Research* 27.1 (Jan. 1992), pp. 131–157. ISSN: 1532-7906. DOI: 10.1207/s15327906mbr2701_8.

Abstract: Stage-sequential dynamic latent variables are of interest in many longitudinal studies. Measurement theory for these latent variables, called Latent Transition Analysis (LTA), can be found in recent generalizations of latent class theory. LTA expands the latent Markov model to allow applications to more complex latent variables and the use of multiple indicators. Because complex latent class models result in sparse contingency tables, that may lead to poor parameter estimation, a simulation study was conducted in order to determine whether model parameters are recovered adequately by LTA, and whether additional indicators result in better measurement or in impossibly sparse tables. The results indicated that parameter recovery was satisfactory overall, although as expected the standard errors were large in some conditions with few subjects. The simulation also indicated that at least within the conditions examined here, the benefits of adding indicators outweigh the costs. Additional indicators improved standard errors, even in conditions producing extremely sparse tables. An example of LTA analysis of empirical data on math skill development is presented.

Cooper et al.: Drinking to regulate positive and negative emotions: A motivational model of alcohol use Cooper-Frone-Russell-etal-1995

Abstract: The present study proposed and tested a motivational model of alcohol use in which people are hypothesized to use alcohol to regulate both positive and negative emotions. Two central premises underpin this model: (a) that enhancement and coping motives for alcohol use are proximal determinants of alcohol use and abuse through which the influence of expectancies, emotions, and other individual differences are mediated and (b) that enhancement and coping motives represent phenomenologically distinct behaviors having both unique antecedents and consequences. This model was tested in 2 random samples (1 of adults, 1 of adolescents) using a combination of moderated regression and path analysis corrected for measurement error. Results revealed strong support for the hypothesized model in both samples and indicate the importance of distinguishing psychological motives for alcohol use.

Courtney et al.: Testing affect regulation models of drinking prior to and after drinking initiation using ecological momentary assessment Courtney-Russell-2021

Abstract: Objective: Affect regulation models of drinking state that affect motivates and reinforces drinking. Few studies have been able to elucidate the timing of these associations in natural settings. We tested positive affect (PA) and negative affect (NA) as predictors of drinking behavior, both prior to and during drinking episodes, and whether drinking predicted changes in affect during episodes. Method: Two hundred twenty two regularly drinking young adults (21-29 years, 84% un-

dergraduates), completed an ecological momentary assessment (EMA) protocol for five consecutive 24-hr periods stretching across 6 days (Wednesday-Monday). Participants provided PA and NA reports three times daily and every half hour during drinking episodes. Alcohol consumption reports were provided each morning and every half hour during drinking episodes. Results: Multi-level models showed that greater pre-drinking PA predicted higher odds of drinking and greater number of drinks consumed. Pre-drinking NA did not predict same day odds of drinking or drinks consumed. Episode-level results revealed different associations for PA and NA with drinking. Current PA did not predict drinks consumed over the next half hour; however, increased drinking was associated with greater increases in PA over the next half hour. Higher NA predicted fewer drinks consumed in the next half hour and higher odds of the end of a drinking episode; however, increased drinking was not associated with changes in NA. Conclusions: PA increased following drinking during episodes. Our results suggest that a focus on PA prior to episodes and a focus on NA during episodes may interrupt processes leading to heavy drinking, and may therefore aid prevention efforts.

Cox et al.: A motivational model of alcohol use Cox-Klinger-1988

Abstract: The final, common pathway to alcohol use is motivational. A person decides consciously or unconsciously to consume or not to consume any particular drink of alcohol according to whether or not he or she expects that the positive affective consequences of drinking will outweigh those of not drinking. Various factors (e.g., past experiences with drinking, current life situation) help to form expectations of affective change from drinking, these factors always modulated by a person's neurochemical reactivity to alcohol. Such major influences include the person's current nonchemical incentives and the prospect of acquiring new positive incentives and removing current negative incentives. Our motivational counseling technique uses nonchemical goals and incentives to help the alcoholic develop a satisfying life without the necessity of alcohol. The technique first assesses the

alcoholic's motivational structure and then seeks to modify it through a multicomponent counseling procedure. The counseling technique is one example of the heuristic value of the motivational model.

Craig: On the frequency function of xy Craig-1936

Cecil C. Craig. "On the frequency function of xy". In: *The Annals of Mathematical Statistics* 7.1 (Mar. 1936), pp. 1–15. DOI: 10.1214/aoms/1177732541.

Cribari-Neto: Asymptotic inference under heteroskedasticity of unknown form
CribariNeto-2004

Francisco Cribari-Neto. "Asymptotic inference under heteroskedasticity of unknown form". In: *Computational Statistics & Data Analysis* 45.2 (Mar. 2004), pp. 215–233. DOI: 10.1016/s0167-9473(02)00366-3.

Abstract: We focus on the finite-sample behavior of heteroskedasticity-consistent covariance matrix estimators and associated quasi-t tests. The estimator most commonly used is that proposed by Halbert White. Its finite-sample behavior under both homoskedasticity and heteroskedasticity is analyzed using Monte Carlo methods. We also consider two other consistent estimators, namely: the HC3 estimator, which is an approximation to the jackknife estimator, and the weighted bootstrap estimator. Additionally, we evaluate the finite-sample behavior of two bootstrap quasi-t tests: the test based on a single bootstrapping scheme and the test based on a double, nested bootstrapping scheme. The latter is very computer-intensive, but proves to work well in small samples. Finally, we propose a new estimator, which we call HC4; it is tailored to take into account the effect of leverage points in the design matrix on associated quasi-t tests.

Cribari-Neto et al.: A new heteroskedasticity-consistent covariance matrix estimator for the linear regression model   CribariNeto-daSilva-2010

Francisco Cribari-Neto and Wilton Bernardino da Silva. "A new heteroskedasticity-consistent co-

variance matrix estimator for the linear regression model". In: *AStA Advances in Statistical Analysis* 95.2 (Nov. 2010), pp. 129–146. DOI: 10.1007/s10182-010-0141-2.

Abstract: The assumption that all random errors in the linear regression model share the same variance (homoskedasticity) is often violated in practice. The ordinary least squares estimator of the vector of regression parameters remains unbiased, consistent and asymptotically normal under unequal error variances. Many practitioners then choose to base their inferences on such an estimator. The usual practice is to couple it with an asymptotically valid estimation of its covariance matrix, and then carry out hypothesis tests that are valid under heteroskedasticity of unknown form. We use numerical integration methods to compute the exact null distributions of some quasi-t test statistics, and propose a new covariance matrix estimator. The numerical results favor testing inference based on the estimator we propose.

Cribari-Neto et al.: Inference under heteroskedasticity and leveraged data

CribariNeto-Souza-Vasconcellos-2007

Francisco Cribari-Neto, Tatiene C. Souza, and Klaus L. P. Vasconcellos. "Inference under heteroskedasticity and leveraged data". In: *Communications in Statistics - Theory and Methods* 36.10 (Aug. 2007), pp. 1877–1888. DOI: 10.1080/03610920601126589.

Abstract: We evaluate the finite-sample behavior of different heteros-ke-das-ticity-consistent covariance matrix estimators, under both constant and unequal error variances. We consider the estimator proposed by Halbert White (HC0), and also its variants known as HC2, HC3, and HC4; the latter was recently proposed by Cribari-Neto (2004). We propose a new covariance matrix estimator: HC5. It is the first consistent estimator to explicitly take into account the effect that the maximal leverage has on the associated inference. Our numerical results show that quasi-t inference based on HC5 is typically more reliable than inference based on other covariance matrix estimators.

Cribari-Neto et al.: Errata: Inference under heteroskedasticity and leveraged data, Communications in Statistics, Theory and Methods, 36, 1877–1888, 2007

Francisco Cribari-Neto, Tatiene C. Souza, and Klaus L. P. Vasconcellos. "Errata: Inference under heteroskedasticity and leveraged data, Communications in Statistics, Theory and Methods, 36, 1877–1888, 2007". In: *Communications in Statistics - Theory and Methods* 37.20 (Sept. 2008), pp. 3329–3330. DOI: 10.1080/03610920802109210.

Crocetti et al.: Generalized anxiety symptoms and identity processes in cross-cultural samples of adolescents from the general population

Elisabetta Crocetti et al. "Generalized anxiety symptoms and identity processes in cross-cultural samples of adolescents from the general population". In: *Child & Youth Care Forum* 44.2 (2015), pp. 159–174. ISSN: 1573-3319. DOI: 10.1007/s10566-014-9275-9.

Abstract: Background: Approximately 20% of adolescents around the world experience mental health problems, most commonly depression or anxiety. High levels of anxiety disorder symptoms can hinder adolescent development, persist into adulthood, and predict negative mental outcomes, such as suicidal ideation and attempts. Objectives: We analyzed generalized anxiety disorder (GAD) symptoms in cross-cultural samples from the general population. We sought to examine cultural and gender differences, and correlates of GAD symptoms in samples of adolescents from six countries located in three different continents (Europe: Bulgaria, Italy, the Netherlands; Africa: Kenya; Asia: China and Philippines). Methods: Participants were 3,445 (51% male) adolescents aged between 14 and 18 years old. They filled self-report measures of GAD symptoms and identity. Results: First, it was found that the scores on GAD symptoms varied significantly across countries, with Dutch respondents reporting the lowest levels whereas Filipino participants exhibited the highest levels of GAD symptoms. Second, gender differences (i.e., girls reported more GAD symptoms than boys)

were significant in each country (as well as in the total sample), with the only exception being that of Kenya. Third, GAD symptoms were significantly related to identity processes and similarities and differences across countries were examined.

Cronbach et al.: How we should measure "change": Or should we?

Lee J. Cronbach and Lita Furby. "How we should measure "change": Or should we?" In: *Psychological Bulletin* 74.1 (July 1970), pp. 68–80. DOI: 10.1037/h0029382.

Abstract: Examines procedures previously recommended by various authors for the estimation of "change" scores, "residual," or "basefree" measures of change, and other kinds of difference scores. A procedure proposed by F. M. Lord is extended to obtain more precise estimates, and an alternative to the L. R. Tucker, F. Damarin, and S. A. Messick (see 41:3) procedure is offered. A consideration of the purposes for which change measures have been sought in the past leads to a series of recommended procedures which solve research and personnel-decision problems without estimation of change scores for individuals.

Cudeck: Analysis of correlation matrices using covariance structure models

Robert Cudeck. "Analysis of correlation matrices using covariance structure models". In: *Psychological Bulletin* 105.2 (Mar. 1989), pp. 317–327. ISSN: 0033-2909. DOI: 10.1037/0033-2909.105.2.317.

Abstract: It is often assumed that covariance structure models can be arbitrarily applied to sample correlation matrices as readily as to sample covariance matrices. Although this is true in many cases and leads to an analysis that is mostly correct, it is not permissible for all structures. This article reviews three interrelated problems associated with the analysis of structural models using a matrix of sample correlations. Depending upon the model, applying a covariance structure to a matrix of correlations may (a) modify the model being studied, (b) produce incorrect values of the omnibus

test statistic, or (c) yield incorrect standard errors. An important class of models are those that are scale invariant (Browne, 1982), for then Errors a and b cannot occur when a correlation matrix is analyzed. A number of examples based on restricted factor analysis are presented to illustrate the concepts described in the article.

Curran et al.: The disaggregation of within-person and between-person effects in longitudinal models of change   Curran-Bauer-2011

Abstract: Longitudinal models are becoming increasingly prevalent in the behavioral sciences, with key advantages including increased power, more comprehensive measurement, and establishment of temporal precedence. One particularly salient strength offered by longitudinal data is the ability to disaggregate between-person and within-person effects in the regression of an outcome on a time-varying covariate. However, the ability to disaggregate these effects has not been fully capitalized upon in many social science research applications. Two likely reasons for this omission are the general lack of discussion of disaggregating effects in the substantive literature and the need to overcome several remaining analytic challenges that limit existing quantitative methods used to isolate these effects in practice. This review explores both substantive and quantitative issues related to the disaggregation of effects over time, with a particular emphasis placed on the multilevel model. Existing analytic methods are reviewed, a general approach to the problem is proposed, and both the existing and proposed methods are demonstrated using several artificial data sets. Potential limitations and directions for future research are discussed, and recommendations for the disaggregation of effects in practice are offered.

Curran et al.: The separation of between-person and within-person components of individual change over time: A latent curve model with structured residuals.

Abstract: Objective: Although recent statistical and computational developments allow for the empirical testing of psychological theories in ways not previously possible, one particularly vexing challenge remains: how to optimally model the prospective, reciprocal relations between 2 constructs as they developmentally unfold over time. Several analytic methods currently exist that attempt to model these types of relations, and each approach is successful to varying degrees. However, none provide the unambiguous separation over time of between-person and within-person components of stability and change, components that are often hypothesized to exist in the psychological sciences. Our goal in this article is to propose and demonstrate a novel extension of the multivariate latent curve model to allow for the disaggregation of these effects. Method: We begin with a review of the standard latent curve models and describe how these primarily capture between-person differences in change. We then extend this model to allow for regression structures among the time-specific residuals to capture within-person differences in change. Results: We demonstrate this model using an artificial data set generated to mimic the developmental relation between alcohol use and depressive symptomatology spanning 5 repeated measures. Conclusions: We obtain a specificity of results from the proposed analytic strategy that is not available from other existing methodologies. We conclude with potential limitations of our approach and directions for future research.

Davidson et al.: Estimation and inference in econometrics   Davidson-MacKinnon-1993

Russell Davidson and James G. MacKinnon. *Estimation and inference in econometrics.* New York, NY: Oxford University Press, 1993. ISBN: 9780195060119.

Library: HB139 .D368 1993.

Davison et al.: Bootstrap methods and their application   Davison-Hinkley-1997

Anthony Christopher Davison and David Victor Hinkley. *Bootstrap methods and their application.* Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge and New York, NY, USA: Cambridge University Press, 1997. ISBN: 9780521573917. DOI: 10.1017/CBO9780511802843.

Abstract: Bootstrap methods are computer-intensive methods of statistical analysis, which use simulation to calculate standard errors, confidence intervals, and significance tests. The methods apply for any level of modelling, and so can be used for fully parametric, semiparametric, and completely nonparametric analysis. This 1997 book gives a broad and up-to-date coverage of bootstrap methods, with numerous applied examples, developed in a coherent way with the necessary theoretical basis. Applications include stratified data; finite populations; censored and missing data; linear, nonlinear, and smooth regression models; classification; time series and spatial problems. Special features of the book include: extensive discussion of significance tests and confidence intervals; material on various diagnostic methods; and methods for efficient computation, including improved Monte Carlo simulation. Each chapter includes both practical and theoretical exercises. S-Plus programs for implementing the methods described in the text are available from the supporting website.

Library: QA276.8 .D38 1997.

Davison et al.: Recent developments in bootstrap methodology

Abstract: Ever since its introduction, the bootstrap has provided both a powerful set of solutions for practical statisticians, and a rich source of theoretical and methodological problems for statistics. In this article, some recent developments in bootstrap methodology are reviewed and discussed. After a brief introduction to the bootstrap, we consider the following topics at varying levels of detail: the use of bootstrapping for highly accurate parametric inference; theoretical properties of nonparametric bootstrapping with unequal probabilities; subsampling and the m out of n bootstrap; bootstrap failures and remedies for superefficient estimators; recent topics in significance testing; bootstrap improvements of unstable classifiers and resampling for dependent data. The treatment is telegraphic rather than exhaustive.

de Haan-Rietdijk et al.: Discrete- vs. continuous-time modeling of unequally spaced experience sampling method data

---

Abstract: The Experience Sampling Method is a common approach in psychological research for collecting intensive longitudinal data with high ecological validity. One characteristic of ESM data is that it is often unequally spaced, because the measurement intervals within a day are deliberately varied, and measurement continues over several days. This poses a problem for discrete-time (DT) modeling approaches, which are based on the assumption that all measurements are equally

spaced. Nevertheless, DT approaches such as (vector) autoregressive modeling are often used to analyze ESM data, for instance in the context of affective dynamics research. There are equivalent continuous-time (CT) models, but they are more difficult to implement. In this paper we take a pragmatic approach and evaluate the practical relevance of the violated model assumption in DT AR(1) and VAR(1) models, for the $N = 1$ case. We use simulated data under an ESM measurement design to investigate the bias in the parameters of interest under four different model implementations, ranging from the true CT model that accounts for all the exact measurement times, to the crudest possible DT model implementation, where even the nighttime is treated as a regular interval. An analysis of empirical affect data illustrates how the differences between DT and CT modeling can play out in practice. We find that the size and the direction of the bias in DT (V)AR models for unequally spaced ESM data depend quite strongly on the true parameter in addition to data characteristics. Our recommendation is to use CT modeling whenever possible, especially now that new software implementations have become available.

Deboeck et al.: Integration of stochastic differential equations using structural equation modeling: A method to facilitate model fitting and pedagogy   Deboeck-Boulton-2016

Abstract: Stochastic differential equation (SDE) models are a promising method for modeling intraindividual change and variability. Applications of SDEs in the social sciences are relatively limited, as these models present conceptual and programming challenges. This article presents a novel method for conceptualizing SDEs. This method uses structural equation modeling (SEM) conventions to simplify SDE specification, the flexibility of SEM to expand the range of SDEs that can be fit, and SEM diagram conventions to facilitate the teaching of SDE concepts. This method is a variation of latent difference scores (McArdle, 2009; McArdle & Hamagami, 2001) and the oversampling

approach (Singer, 2012), and approximates the advantages of analytic methods such as the exact discrete model (Oud & Jansen, 2000) while retaining the modeling flexibility of methods such as latent differential equation modeling (Boker, Neale, & Rausch, 2004). A simulation and empirical example are presented to illustrate that this method can be implemented on current computing hardware, produces good approximations of analytic solutions, and can flexibly accommodate novel models.

Deboeck et al.: No need to be discrete: A method for continuous time mediation analysis — Deboeck-Preacher-2015

Abstract: Mediation is one concept that has shaped numerous theories. The list of problems associated with mediation models, however, has been growing. Mediation models based on cross-sectional data can produce unexpected estimates, so much so that making longitudinal or causal inferences is inadvisable. Even longitudinal mediation models have faults, as parameter estimates produced by these models are specific to the lag between observations, leading to much debate over appropriate lag selection. Using continuous time models (CTMs) rather than commonly employed discrete time models, one can estimate lag-independent parameters. We demonstrate methodology that allows for continuous time mediation analyses, with attention to concepts such as indirect and direct effects, partial mediation, the effect of lag, and the lags at which relations become maximal. A simulation compares common longitudinal mediation methods with CTMs. Reanalysis of a published covariance matrix demonstrates that CTMs can be fit to data used in longitudinal mediation studies.

Pascal R. Deboeck, Kristopher J. Preacher, and David A. Cole. "Mediation modeling: Differing perspectives on time alter mediation inferences". In: *Continuous time modeling in the behavioral and related sciences*. Springer International Publishing, 2018, pp. 179–203. ISBN: 9783319772196. DOI: 10.1007/978-3-319-77219-6_8.

Abstract: Time is unlike any other variable collected in the social, behavioral, and medical sciences. Research participants who are sampled, and variables that are measured, come in distinct, discrete units. Although time is often recorded in such discrete units (e.g., wave 1, grade 3, day 5), time is markedly different from either participants or variables. Sampling time points is unlike sampling people or variables, as there are an arbitrary number of additional samples that can be collected between any two occasions of measurement. These interstitial samples are ignored by many longitudinal modeling paradigms. These observations that occur between sampling occasions form the basis for the perspectives on mediation explored in this chapter. We focus on the difference in perspectives offered by discrete time approaches commonly utilized in mediation research versus models that conceptualize time as a continuous variable. The differences in how one conceptualizes time have the potential to alter such core mediation concepts as direct and indirect effect, complete and partial mediation, and even what constitutes a "mediation" model.

DeMartini et al.: Dynamic structural equation modeling of the relationship between alcohol habit and drinking variability   DeMartini-Gueorguieva-Taylor-etal-2022

Kelly S. DeMartini et al. "Dynamic structural equation modeling of the relationship between alcohol habit and drinking variability". In: *Drug and Alcohol Dependence* 233 (Apr. 2022), p. 109202. ISSN: 0376-8716. DOI: 10.1016/j.drugalcdep.2021.109202.

Abstract: Background: A hyper-engaged habit system may be common in alcohol use disorders (AUDs). Regarding drinking patterns, habit may be expressed as higher levels of drinking autore-

gression, where previous day drinking is correlated with next day drinking. This study utilized dynamic structural equation models (DSEM) with intensive longitudinal data to understand whether alcohol habit relates to drinking autoregression and variable levels of alcohol consumption. Methods: Participants were adult drinkers ($N = 313$) who completed baseline self-report assessments of past 30-day alcohol consumption and alcohol habit. Alcohol habit was measured by the Self Report Habit Index (SRHI). Thirty-day coding of the Timeline Followback assessed total daily drinking and any daily heavy drinking. Results: The DSEM model for daily drinking found a weak but significant autoregressive data structure. Alcohol habit was related to increased mean drinking but did not strengthen the autoregressive effect of drinks per day. Higher alcohol habit was associated with higher levels of drinks per day person-specific variability. This pattern was replicated with the DSEM model for heavy drinking. Alcohol habit did not impact the autoregressive effect of heavy drinking but was associated with higher levels of heavy drinking. Conclusions: While both drinks per day and heavy drinking showed a significant autoregressive structure, evidence of alcohol habit did not strengthen this effect. Alcohol habit did impact drinking variability; higher alcohol habit is associated with greater levels of drinking variability and higher mean drinking. Strategies to regulate drinking variability, including heavier drinking occasions, could target AUD habit.

Demeshko et al.: A novel continuous and structural VAR modeling approach and its application to reactor noise analysis   Demeshko-Washio-Kawahara-etal-2015

Abstract: A vector autoregressive model in discrete time domain (DVAR) is often used to analyze continuous time, multivariate, linear Markov systems through their observed time series data sampled at discrete timesteps. Based on previous studies, the DVAR model is supposed to be a non-canonical representation of the system, that is, it does not correspond to a unique system bijectively. However, in this article, we characterize the relations of the DVAR model with its corresponding

Structural Vector AR (SVAR) and Continuous Time Vector AR (CTVAR) models through a finite difference method across continuous and discrete time domain. We further clarify that the DVAR model of a continuous time, multivariate, linear Markov system is canonical under a highly generic condition. Our analysis shows that we can uniquely reproduce its SVAR and CTVAR models from the DVAR model. Based on these results, we propose a novel Continuous and Structural Vector Autoregressive (CSVAR) modeling approach to derive the SVAR and the CTVAR models from their DVAR model empirically derived from the observed time series of continuous time linear Markov systems. We demonstrate its superior performance through some numerical experiments on both artificial and real-world data.

Didier et al.: Signal processing and machine learning with transdermal alcohol concentration to predict natural environment alcohol consumption.

Abstract: Wrist-worn alcohol biosensors continuously and discreetly record transdermal alcohol concentration (TAC) and may allow alcohol researchers to monitor alcohol consumption in participants' natural environments. However, the field lacks established methods for signal processing and detecting alcohol events using these devices. We developed software that streamlines analysis of raw data (TAC, temperature, and motion) from a wrist-worn alcohol biosensor (BACtrack Skyn) through a signal processing and machine learning pipeline: biologically implausible skin surface temperature readings (¡ 28C) were screened for potential device removal and TAC artifacts were corrected, features that describe TAC (e.g., rise duration) were calculated and used to train models (random forest and logistic regression) that predict self-reported alcohol consumption, and model performances were measured and summarized in autogenerated reports. The software was tested using 60 Skyn data sets recorded during 30 alcohol drinking episodes and 30 nonalcohol drinking

episodes. Participants (N = 36; 13 with alcohol use disorder) wore the Skyn during one alcohol drinking episode and one nonalcohol drinking episode in their natural environment. In terms of distinguishing alcohol from nonalcohol drinking, correcting artifacts in the data resulted in 10% improvement in model accuracy relative to using raw data. Random forest and logistic regression models were both accurate, correctly predicting 97% (58/60; AUC-ROCs = 0.98, 0.96) of episodes. Area under TAC curve, rise duration of TAC curve, and peak TAC were the most important features for predictive accuracy. With promising model performance, this protocol will enhance the efficiency and reliability of TAC sensors for future alcohol monitoring research.

Dishion et al.: Parental monitoring and the prevention of child and adolescent problem behavior: A conceptual and empirical formulation   Dishion-McMahon-1998

Thomas J. Dishion and Robert J. McMahon. "Parental monitoring and the prevention of child and adolescent problem behavior: A conceptual and empirical formulation". In: *Clinical Child and Family Psychology Review* 1.1 (Mar. 1998), pp. 61–75. ISSN: 1573-2827. DOI: 10.1023/a:1021800432380.

Abstract: The present report accomplishes three goals. First, to provide an empirical rationale for placing parental monitoring of children's adaptations as a key construct in development and prevention research. Second, to stimulate more research on parental monitoring and provide an integrative framework for various research traditions as well as developmental periods of interest. Third, to discuss current methodological issues that are developmentally and culturally sensitive and based on sound measurement. Possible intervention and prevention strategies that specifically target parental monitoring are discussed.

Doñamayor et al.: The involuntary nature of binge drinking: Goal directedness and awareness of intention   Donamayor-Strelchuk-Baek-etal-2017

Nuria Doñamayor et al. "The involuntary nature of binge drinking: Goal directedness and awareness

of intention". In: *Addiction Biology* 23.1 (Apr. 2017), pp. 515–526. ISSN: 1369-1600. DOI: 10.1111/adb.12505.

Abstract: Binge drinking represents a public health issue and is a known risk factor in the development of alcohol use disorders. Previous studies have shown behavioural as well as neuroanatomical alterations associated with binge drinking. Here, we address the question of the automaticity or involuntary nature of the behaviour by assessing goal-directed behaviour and intentionality. In this study, we used a computational two-step task, designed to discern between model-based/goal-directed and model-free/habitual behaviours, and the classic Libet clock task, to study intention awareness, in a sample of 31 severe binge drinkers (BD) and 35 matched healthy volunteers. We observed that BD had impaired goal-directed behaviour in the two-step task compared with healthy volunteers. In the Libet clock task, BD showed delayed intention awareness. Further, we demonstrated that alcohol use severity, as reflected by the alcohol use disorders identification test, correlated with decreased conscious awareness of volitional intention in BD, although it was unrelated to performance on the two-step task. However, the time elapsed since the last drinking binge influenced the model-free scores, with BD showing less habitual behaviour after longer abstinence. Our findings suggest that the implementation of goal-directed strategies and the awareness of volitional intention are affected in current heavy alcohol users. However, the modulation of these impairments by alcohol use severity and abstinence suggests a state effect of alcohol use in these measures and that top-down volitional control might be ameliorated with alcohol use cessation.

Dora et al.: The daily association between affect and alcohol use: A meta-analysis of individual participant data   **Dora-Piccirillo-Foster-etal-2023**

Jonas Dora et al. "The daily association between affect and alcohol use: A meta-analysis of individual participant data". In: *Psychological Bulletin* 149.1-2 (Jan. 2023), pp. 1–24. ISSN: 0033-2909. DOI: 10.1037/bul0000387.

Abstract: Influential psychological theories hypothesize that people consume alcohol in response to the experience of both negative and positive emotions. Despite two decades of daily diary and ecological momentary assessment research, it remains unclear whether people consume more alcohol on days they experience higher negative and positive affects in everyday life. In this preregistered meta-analysis, we synthesized the evidence for these daily associations between affect and alcohol use. We included individual participant data from 69 studies (N = 12,394), which used daily and momentary surveys to assess the affect and the number of alcoholic drinks consumed. Results indicate that people are not more likely to drink on days they experience high negative affect but are more likely to drink and drink heavily on days high in positive affect. People self-reporting a motivational tendency to drink-to-cope and drink-to-enhance consumed more alcohol but not on days they experienced higher negative and positive affects. Results were robust across different operationalizations of affect, study designs, study populations, and individual characteristics. These findings challenge the long-held belief that people drink more alcohol following increase in negative affect. Integrating these findings under different theoretical models and limitations of this field of research, we collectively propose an agenda for future research to explore open questions surrounding affect and alcohol use.

Driver: Inference with cross-lagged effects—Problems in time   Driver-2025

Abstract: The interpretation of cross-effects from vector autoregressive models to infer structure and causality among constructs is widespread and sometimes problematic. I describe problems in the interpretation of cross-effects when processes that are thought to fluctuate continuously in time are, as is typically done, modeled as changing only in discrete steps (as in e.g., structural equation modeling)–zeroes in a discrete-time temporal matrix do not necessarily correspond to zero effects in the underlying continuous processes, and vice versa. This has implications for the common case when the presence or absence of cross-effects is used for inference about underlying causal processes.

I demonstrate these problems via simulation, and also show that when an underlying set of processes
are continuous in time, even relatively few direct causal links can result in much denser temporal
effect matrices in discrete-time. I demonstrate one solution to these issues, namely parameteriz-
ing the system as a stochastic differential equation and focusing inference on the continuous-time
temporal effects. I follow this with some discussion of issues regarding the switch to continuous-
time, specifically regularization, appropriate measurement time lag, and model order. An empirical
example using intensive longitudinal data highlights some of the complexities of applying such ap-
proaches to real data, particularly with respect to model specification, examining misspecification,
and parameter interpretation.

Driver et al.: Continuous time structural equation modeling with R package ctsem

Driver-Oud-Voelkle-2017

Abstract: We introduce ctsem, an R package for continuous time structural equation modeling of
panel ($N > 1$) and time series ($N = 1$) data, using full information maximum likelihood. Most
dynamic models (e.g., cross-lagged panel models) in the social and behavioural sciences are discrete
time models. An assumption of discrete time models is that time intervals between measurements
are equal, and that all subjects were assessed at the same intervals. Violations of this assumption
are often ignored due to the difficulty of accounting for varying time intervals, therefore parameter
estimates can be biased and the time course of effects becomes ambiguous. By using stochastic
differential equations to estimate an underlying continuous process, continuous time models allow
for any pattern of measurement occasions. By interfacing to OpenMx, ctsem combines the flexible
specification of structural equation models with the enhanced data gathering opportunities and
improved estimation of continuous time models. ctsem can estimate relationships over time for
multiple latent processes, measured by multiple noisy indicators with varying time intervals between

observations. Within and between effects are estimated simultaneously by modeling both observed covariates and unobserved heterogeneity. Exogenous shocks with different shapes, group differences, higher order diffusion effects and oscillating processes can all be simply modeled. We first introduce and define continuous time models, then show how to specify and estimate a range of continuous time models using ctsem.

Driver et al.: Hierarchical Bayesian continuous time dynamic modeling.

Abstract: Continuous time dynamic models are similar to popular discrete time models such as autoregressive cross-lagged models, but through use of stochastic differential equations can accurately account for differences in time intervals between measurements, and more parsimoniously specify complex dynamics. As such they offer powerful and flexible approaches to understand ongoing psychological processes and interventions, and allow for measurements to be taken a variable number of times, and at irregular intervals. However, limited developments have taken place regarding the use of continuous time models in a fully hierarchical context, in which all model parameters are allowed to vary over individuals. This has meant that questions regarding individual differences in parameters have had to rely on single-subject time series approaches, which require far more measurement occasions per individual. We present a hierarchical Bayesian approach to estimating continuous time dynamic models, allowing for individual variation in all model parameters. We also describe an extension to the ctsem package for R, which interfaces to the Stan software and allows simple specification and fitting of such models. To demonstrate the approach, we use a subsample from the German socioeconomic panel and relate overall life satisfaction and satisfaction with health.

Dudgeon: Some improvements in confidence intervals for standardized regression coefficients Dudgeon-2017

Abstract: Yuan and Chan (Psychometrika 76:670-690, 2011. doi:10.1007/S11336-011-9224-6) derived consistent confidence intervals for standardized regression coefficients under fixed and random score assumptions. Jones and Waller (Psychometrika 80:365-378, 2015. doi:10.1007/S11336-013-9380-Y) extended these developments to circumstances where data are non-normal by examining confidence intervals based on Browne's (Br J Math Stat Psychol 37:62-83, 1984. doi:10.1111/j.2044-8317.1984.tb00789.x) asymptotic distribution-free (ADF) theory. Seven different heteroscedastic-consistent (HC) estimators were investigated in the current study as potentially better solutions for constructing confidence intervals on standardized regression coefficients under non-normality. Normal theory, ADF, and HC estimators were evaluated in a Monte Carlo simulation. Findings confirmed the superiority of the HC3 (MacKinnon and White, J Econ 35:305-325, 1985. doi:10.1016/0304-4076(85)90158-7) and HC5 (Cribari-Neto and Da Silva, Adv Stat Anal 95:129-146, 2011. doi:10.1007/s10182-010-0141-2) interval estimators over Jones and Waller's ADF estimator under all conditions investigated, as well as over the normal theory method. The HC5 estimator was more robust in a restricted set of conditions over the HC3 estimator. Some possible extensions of HC estimators to other effect size measures are considered for future developments.

Dumenci et al.: A latent trait-state model of adolescent depression using the Center for Epidemiologic Studies-Depression Scale Dumenci-Windle-1996

Abstract: Utilized the latent trait-state model for estimating stable and changing components of depressive symptomology in adolescents. The factorial structure of the Center for Epidemiologic Studies-Depression Scale (CES-D) was assessed separately for males and females at 4 measurement occasions, at 6-mo intervals. The variance decomposition of general trait, state, specific trait, and random error parameters for the CES-D scores was estimated simultaneously and tested statistically. Parameter estimates indicated that the CES-D measured both trait- and state-depression about equally well, and that the trait-specific variance parameter was statistically significant, but substantially smaller than those associated with general trait- and state-depression. Findings are discussed with regard to depressive mood fluctuations among adolescents and the potential usefulness of the latent trait-state model to capture such dynamic features of development.

O. D. Duncan: Some linear models for two-wave, two-variable panel analysis

Duncan-1969

Otis D. Duncan. "Some linear models for two-wave, two-variable panel analysis". In: *Psychological Bulletin* 72.3 (Sept. 1969), pp. 177–182. ISSN: 0033-2909. DOI: 10.1037/h0027876.

Abstract: In the absence of a sufficient number of a priori substantive assumptions ruling out certain conceivable causal linkages among variables, neither cross-lagged correlation nor any other technique for analyzing 2-wave, 2-variable panel data will yield a unique causal inference. A wide variety of distinct linear causal models will always be compatible with a given set of panel data.

Eddelbuettel: Seamless R and C++ integration with Rcpp   Eddelbuettel-2013

Dirk Eddelbuettel. *Seamless R and C++ integration with Rcpp*. Springer New York, 2013. ISBN: 978-1-4614-6868-4. DOI: 10.1007/978-1-4614-6868-4.

Abstract: Illustrates a range of statistical computations in R using the Rcpp package. Provides a general introduction to extending R with C++ code. Features an appendix for R users new to the

C++ programming language Rcpp packages are presented in the context of useful application case studies.

Eddelbuettel et al.: Extending R with C++: A brief introduction to Rcpp

Eddelbuettel-Balamuta-2017

Dirk Eddelbuettel and James Joseph Balamuta. "Extending R with C++: A brief introduction to Rcpp". In: *PeerJ Preprints* 3188v1.3 (Aug. 2017). DOI: 10.7287/peerj.preprints.3188v1.

Abstract: R has always provided an application programming interface (API) for extensions. Based on the C language, it uses a number of macros and other low-level constructs to exchange data structures between the R process and any dynamically-loaded component modules authors added to it. With the introduction of the Rcpp package, and its later refinements, this process has become considerably easier yet also more robust. By now, Rcpp has become the most popular extension mechanism for R. This article introduces Rcpp, and illustrates with several examples how the Rcpp Attributes mechanism in particular eases the transition of objects between R and C++ code.

Eddelbuettel et al.: Rcpp: Seamless R and C++ integration

Eddelbuettel-Francois-2011

Dirk Eddelbuettel and Romain François. "Rcpp: Seamless R and C++ integration". In: *Journal of Statistical Software* 40.8 (2011). DOI: 10.18637/jss.v040.i08.

Abstract: The Rcpp package simplifies integrating C++ code with R. It provides a consistent C++ class hierarchy that maps various types of R objects (vectors, matrices, functions, environments, ...) to dedicated C++ classes. Object interchange between R and C++ is managed by simple, flexible and extensible concepts which include broad support for C++ Standard Template Library idioms. C++ code can both be compiled, linked and loaded on the fly, or added via packages. Flexible error and exception code handling is provided. Rcpp substantially lowers the barrier for programmers wanting to combine C++ code with R.

Eddelbuettel et al.: Rcpp: Seamless R and C++ Integration

Eddelbuettel-Francois-Allaire-etal-2023

Dirk Eddelbuettel, Romain Francois, et al. *Rcpp: Seamless R and C++ Integration*. 2023. URL: https://CRAN.R-project.org/package=Rcpp.

Eddelbuettel et al.: RcppArmadillo: Accelerating R with high-performance C++ linear algebra

Eddelbuettel-Sanderson-2014

Dirk Eddelbuettel and Conrad Sanderson. "RcppArmadillo: Accelerating R with high-performance C++ linear algebra". In: *Computational Statistics & Data Analysis* 71 (Mar. 2014), pp. 1054–1063. DOI: 10.1016/j.csda.2013.02.005.

Abstract: The R statistical environment and language has demonstrated particular strengths for interactive development of statistical algorithms, as well as data modelling and visualisation. Its current implementation has an interpreter at its core which may result in a performance penalty in comparison to directly executing user algorithms in the native machine code of the host CPU. In contrast, the C++ language has no built-in visualisation capabilities, handling of linear algebra or even basic statistical algorithms; however, user programs are converted to high-performance machine code, ahead of execution. A new method avoids possible speed penalties in R by using the Rcpp extension package in conjunction with the Armadillo C++ matrix library. In addition to the inherent performance advantages of compiled code, Armadillo provides an easy-to-use template-based meta-programming framework, allowing the automatic pooling of several linear algebra operations into one, which in turn can lead to further speedups. With the aid of Rcpp and Armadillo, conversion of linear algebra centred algorithms from R to C++ becomes straightforward. The algorithms retain the overall structure as well as readability, all while maintaining a bidirectional link with the host R environment. Empirical timing comparisons of R and C++ implementations of a Kalman filtering algorithm indicate a speedup of several orders of magnitude.

Abstract: We discuss the following problem: given a random sample $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ from an unknown probability distribution F, estimate the sampling distribution of some prespecified random variable $R(\boldsymbol{X}, F)$, on the basis of the observed data $\boldsymbol{x}$. (Standard jackknife theory gives an approximate mean and variance in the case $R(\boldsymbol{X}, F) = \theta\left(\hat{F}\right) - \theta(F)$, θ some parameter of interest.) A general method, called the "bootstrap" is introduced, and shown to work satisfactorily on a variety of estimation problems. The jackknife is shown to be a linear approximation method for the bootstrap. The exposition proceeds by a series of examples: variance of the sample median, error rates in a linear discriminant analysis, ratio estimation, estimating regression parameters, etc.

Abstract: This is a survey article concerning recent advances in certain areas of statistical theory, written for a mathematical audience with no background in statistics. The topics are chosen to illustrate a special point: how the advent of the high-speed computer has affected the development of statistical theory. The topics discussed include nonparametric methods, the jackknife, the bootstrap, cross-validation, error-rate estimation in discriminant analysis, robust estimation, the influence function, censored data, the EM algorithm, and Cox's likelihood function. The exposition is mainly by example, with only a little offered in the way of theoretical development.

Abstract: We investigate several nonparametric methods; the bootstrap, the jackknife, the delta method, and other related techniques. The first and simplest goal is the assignment of nonparametric standard errors to a real-valued statistic. More ambitiously, we consider setting nonparametric confidence intervals for a real-valued parameter. Building on the well understood case of confidence intervals for the median, some hopeful evidence is presented that such a theory may be possible.

Abstract: We consider the problem of setting approximate confidence intervals for a single parameter θ in a multiparameter family. The standard approximate intervals based on maximum likelihood theory, $\hat{\theta} \pm \hat{\sigma} z^{(\alpha)}$, can be quite misleading. In practice, tricks based on transformations, bias corrections, and so forth, are often used to improve their accuracy. The bootstrap confidence intervals discussed in this article automatically incorporate such tricks without requiring the statistician to think them through for each new application, at the price of a considerable increase in computational effort. The new intervals incorporate an improvement over previously suggested methods,

which results in second-order correctness in a wide variety of problems. In addition to parametric families, bootstrap intervals are also developed for nonparametric situations.

Efron: Bootstrap confidence intervals: Good or bad? Efron-1988

Abstract: The bootstrap is a nonparametric technique for estimating standard errors and approximate confidence intervals. Rasmussen has used a simulation experiment to suggest that bootstrap confidence intervals perform very poorly in the estimation of a correlation coefficient. Part of Rasmussen's simulation is repeated. A careful look at the results shows the bootstrap intervals performing quite well. Some remarks are made concerning the virtues and defects of bootstrap intervals in general.

Efron: Second thoughts on the bootstrap Efron-2003

Abstract: This brief review article is appearing in the issue of Statistical Science that marks the 25th anniversary of the bootstrap. It concerns some of the theoretical and methodological aspects of the bootstrap and how they might influence future work in statistics.

Efron: Bayesian inference and the parametric bootstrap Efron-2012

Abstract: The parametric bootstrap can be used for the efficient computation of Bayes posterior distributions. Importance sampling formulas take on an easy form relating to the deviance in ex-

ponential families and are particularly simple starting from Jeffreys invariant prior. Because of the i.i.d. nature of bootstrap sampling, familiar formulas describe the computational accuracy of the Bayes estimates. Besides computational methods, the theory provides a connection between Bayesian and frequentist analysis. Efficient algorithms for the frequentist accuracy of Bayesian inferences are developed and demonstrated in a model selection example.

Efron et al.: An introduction to the bootstrap   Efron-Tibshirani-1993

Bradley Efron and Robert J. Tibshirani. *An introduction to the bootstrap*. Monographs on statistics and applied probability ; 57. New York: Chapman & Hall, 1993. ISBN: 9780412042317. DOI: 10.1201/9780429246593.

Abstract: Statistics is a subject of many uses and surprisingly few effective practitioners. The traditional road to statistical knowledge is blocked, for most, by a formidable wall of mathematics. The approach in An Introduction to the Bootstrap avoids that wall. It arms scientists and engineers, as well as statisticians, with the computational techniques they need to analyze and understand complicated data sets.

Eilers et al.: Flexible smoothing with B-splines and penalties   Eilers-Marx-1996

Paul H. C. Eilers and Brian D. Marx. "Flexible smoothing with B-splines and penalties". In: *Statistical Science* 11.2 (May 1996). ISSN: 0883-4237. DOI: 10.1214/ss/1038425655.

Abstract: B-splines are attractive for nonparametric modelling, but choosing the optimal number and positions of knots is a complex task. Equidistant knots can be used, but their small and discrete number allows only limited control over smoothness and fit. We propose to use a relatively large number of knots and a difference penalty on coefficients of adjacent B-splines. We show connections to the familiar spline penalty on the integral of the squared second derivative. A short

overview of B-splines, of their construction and of penalized likelihood is presented. We discuss properties of penalized B-splines and propose various criteria for the choice of an optimal penalty parameter. Nonparametric logistic regression, density estimation and scatterplot smoothing are used as examples. Some details of the computations are presented.

Eisenberg et al.: Emotion-related self-regulation and its relation to children's maladjustment
<div align="right">**Eisenberg-Spinrad-Eggum-2010**</div>

Abstract: The development of children's emotion-related self-regulation appears to be related to, and likely involved in, many aspects of children's development. In this review, the distinction between effortful self-regulatory processes and those that are somewhat less voluntary is discussed, and literature on the former capacities is reviewed. Emotion-related self-regulation develops rapidly in the early years of life and improves more slowly into adulthood. Individual differences in children's self-regulation are fairly stable after the first year or two of life. Such individual differences are inversely related to at least some types of externalizing problems. Findings for internalizing problems are less consistent and robust, although emotion-related self-regulation appears to be inversely related to internalizing problems after the early years. Self-regulatory capacities have been related to both genetic and environmental factors and their interaction. Some interventions designed to foster self-regulation and, hence, reduce maladjustment, have proved to be at least partially effective.

Enders: Applied missing data analysis
<div align="right">**Enders-2010**</div>

Abstract: Walking readers step by step through complex concepts, this book translates missing

data techniques into something that applied researchers and graduate students can understand and utilize in their own research. Enders explains the rationale and procedural details for maximum likelihood estimation, Bayesian estimation, multiple imputation, and models for handling missing not at random (MNAR) data. Easy-to-follow examples and small simulated data sets illustrate the techniques and clarify the underlying principles. The companion website (www.appliedmissingdata.com) includes data files and syntax for the examples in the book as well as up-to-date information on software. The book is accessible to substantive researchers while providing a level of detail that will satisfy quantitative specialists.

Enders et al.: A Bayesian approach for estimating mediation effects with missing data
Enders-Fairchild-MacKinnon-2013

Abstract: Methodologists have developed mediation analysis techniques for a broad range of substantive applications, yet methods for estimating mediating mechanisms with missing data have been understudied. This study outlined a general Bayesian missing data handling approach that can accommodate mediation analyses with any number of manifest variables. Computer simulation studies showed that the Bayesian approach produced frequentist coverage rates and power estimates that were comparable to those of maximum likelihood with the bias-corrected bootstrap. We share a SAS macro that implements Bayesian estimation and use two data analysis examples to demonstrate its use.

George L. Engel. "The need for a new medical model: A challenge for biomedicine". In: *Science* 196.4286 (Apr. 1977), pp. 129–136. ISSN: 1095-9203. DOI: 10.1126/science.847460.

Abstract: The dominant model of disease today is biomedical, and it leaves no room within tis framework for the social, psychological, and behavioral dimensions of illness. A biopsychosocial model is proposed that provides a blueprint for research, a framework for teaching, and a design for action in the real world of health care.

Epskamp et al.: Estimating psychological networks and their accuracy: A tutorial paper

Epskamp-Borsboom-Fried-2017

Sacha Epskamp, Denny Borsboom, and Eiko I. Fried. "Estimating psychological networks and their accuracy: A tutorial paper". In: *Behavior Research Methods* 50.1 (Mar. 2017), pp. 195–212. ISSN: 1554-3528. DOI: 10.3758/s13428-017-0862-1.

Abstract: The usage of psychological networks that conceptualize behavior as a complex interplay of psychological and other components has gained increasing popularity in various research fields. While prior publications have tackled the topics of estimating and interpreting such networks, little work has been conducted to check how accurate (i.e., prone to sampling variation) networks are estimated, and how stable (i.e., interpretation remains similar with less observations) inferences from the network structure (such as centrality indices) are. In this tutorial paper, we aim to introduce the reader to this field and tackle the problem of accuracy under sampling variation. We first introduce the current state-of-the-art of network estimation. Second, we provide a rationale why researchers should investigate the accuracy of psychological networks. Third, we describe how bootstrap routines can be used to (A) assess the accuracy of estimated network connections, (B) investigate the stability of centrality indices, and (C) test whether network connections and centrality estimates for different variables differ from each other. We introduce two novel statistical methods: for (B) the correlation

stability coefficient, and for (C) the bootstrapped difference test for edge-weights and centrality indices. We conducted and present simulation studies to assess the performance of both methods. Finally, we developed the free R-package bootnet that allows for estimating psychological networks in a generalized framework in addition to the proposed bootstrap methods. We showcase bootnet in a tutorial, accompanied by R syntax, in which we analyze a dataset of 359 women with posttraumatic stress disorder available online.

Epskamp et al.: qgraph: Network visualizations of relationships in psychometric data
Epskamp-Cramer-Waldorp-etal-2012

Abstract: We present the qgraph package for R, which provides an interface to visualize data through network modeling techniques. For instance, a correlation matrix can be represented as a network in which each variable is a node and each correlation an edge; by varying the width of the edges according to the magnitude of the correlation, the structure of the correlation matrix can be visualized. A wide variety of matrices that are used in statistics can be represented in this fashion, for example matrices that contain (implied) covariances, factor loadings, regression parameters and p values. qgraph can also be used as a psychometric tool, as it performs exploratory and confirmatory factor analysis, using sem and lavaan; the output of these packages is automatically visualized in qgraph, which may aid the interpretation of results. In this article, we introduce qgraph by applying the package functions to data from the NEO-PI-R, a widely used personality questionnaire.

Epskamp et al.: The Gaussian graphical model in cross-sectional and time-series data
Epskamp-Lourens-Mottus-etal-2018

Sacha Epskamp, Lourens J. Waldorp, et al. "The Gaussian graphical model in cross-sectional and

time-series data". In: *Multivariate Behavioral Research* 53.4 (Apr. 2018), pp. 453–480. DOI: 10.1080/00273171.2018.1454823.

Abstract: We discuss the Gaussian graphical model (GGM; an undirected network of partial correlation coefficients) and detail its utility as an exploratory data analysis tool. The GGM shows which variables predict one-another, allows for sparse modeling of covariance structures, and may highlight potential causal relationships between observed variables. We describe the utility in three kinds of psychological data sets: data sets in which consecutive cases are assumed independent (e.g., cross-sectional data), temporally ordered data sets (e.g., n = 1 time series), and a mixture of the 2 (e.g., n ¿ 1 time series). In time-series analysis, the GGM can be used to model the residual structure of a vector-autoregression analysis (VAR), also termed graphical VAR. Two network models can then be obtained: a temporal network and a contemporaneous network. When analyzing data from multiple subjects, a GGM can also be formed on the covariance structure of stationary means-the between-subjects network. We discuss the interpretation of these models and propose estimation methods to obtain these networks, which we implement in the R packages graphicalVAR and mlVAR. The methods are showcased in two empirical examples, and simulation studies on these methods are included in the supplementary materials.

Ernst et al.: Utilizing a quantile function approach to obtain exact bootstrap solutions

Ernst-Hutson-2003

Michael D. Ernst and Alan D. Hutson. "Utilizing a quantile function approach to obtain exact bootstrap solutions". In: *Statistical Science* 18.2 (May 2003). DOI: 10.1214/ss/1063994978.

Abstract: The popularity of the bootstrap is due in part to its wide applicability and the ease of implementing resampling procedures on modern computers. But careful reading of Efron (1979) will show that at its heart, the bootstrap is a "plug-in" procedure that involves calculating a functional $\theta\left(\hat{F}\right)$ from an estimate of the c.d.f. F. Resampling becomes invaluable when, as is often the case,

$\theta\left(\hat{F}\right)$ cannot be calculated explicitly. We discuss some situations where working with the sample quantile function, $\hat{Q}$, rather than $\hat{F}$, can lead to explicit (exact) solutions to $\theta\left(\hat{F}\right)$.

Evans et al.: Childhood poverty, chronic stress, self-regulation, and coping

Abstract: Poverty is a powerful factor that can alter lifetime developmental trajectories in cognitive, socioemotional, and physical health outcomes. Most explanatory work on the underlying psychological processes of how poverty affects development has focused on parental investment and parenting practices, principally responsiveness. Our primary objective in this article was to describe a third, complementary pathway—chronic stress and coping—that may also prove helpful in understanding the developmental impacts of early childhood poverty throughout life. Disadvantaged children are more likely than their wealthier peers to confront a wide array of physical stressors (e.g., substandard housing, chaotic environments) and psychosocial stressors (e.g., family turmoil, separation from adult caregivers). As exposure to stressors accumulates, physiological response systems that are designed to handle relatively infrequent, acute environmental demands are overwhelmed. Chronic cumulative stressors also disrupt the self-regulatory processes that help children cope with external demands.

Fairchild et al.: Using mediation and moderation analyses to enhance prevention research

Abstract: Integrating mediating and moderating variables into prevention research can refine interventions and guide program evaluation by demonstrating how and for whom programs work, as well as lending insight into the construct validity of an intervention. In this way, program development and evaluation strategies that incorporate mediation and moderation analyses contribute to our ability to affect behavioral change. This chapter aims to illustrate how mediation and moderation analyses enhance and inform prevention and intervention work. To that end we define and differentiate the models, discuss their application to prevention programming and research, and provide information on their estimation for individuals seeking to implement these analyses.

Fairchild et al.: R^2 effect-size measures for mediation analysis

Fairchild-MacKinnon-Taborga-etal-2009

Amanda J. Fairchild, David P. MacKinnon, et al. "R^2 effect-size measures for mediation analysis". In: *Behavior Research Methods* 41.2 (May 2009), pp. 486–498. ISSN: 1554-3528. DOI: 10.3758/brm.41.2.486.

Abstract: R^2 effect-size measures are presented to assess variance accounted for in mediation models. The measures offer a means to evaluate both component paths and the overall mediated effect in mediation models. Statistical simulation results indicate acceptable bias across varying parameter and sample-size combinations. The measures are applied to a real-world example using data from a team-based health promotion program to improve the nutrition and exercise habits of firefighters. SAS and SPSS computer code are also provided for researchers to compute the measures in their own data.

Feitelson et al.: Job Scheduling Strategies for Parallel Processing

Feitelson-Rudolph-Schwiegelshohn-2003

Dror Feitelson, Larry Rudolph, and Uwe Schwiegelshohn, eds. *Job Scheduling Strategies for Parallel Processing*. Springer Berlin Heidelberg, 2003. ISBN: 9783540397274. DOI: 10.1007/10968987.

Ephrem Fernandez. *Anxiety, depression, and anger in pain: Research findings and clinical options.*
Dallas, TX: Advanced Psychological Resources, 2002. ISBN: 978-0972316408.

Abstract: This book is about the many ways in which anxiety, depression, and anger can predispose
a person to pain, trigger the pain, aggravate it, maintain it, in addition to being correlates or
consequences of pain. These interactions are clearly illustrated and embellished with examples.
Pain is described in terms of neurological signals, sensation, perception, cognition, and behavior,
but with special reference to emotions, moods, and affective disorders. In each of the chapters on
anxiety, depression, and anger, the author conveys the significance of emotional problems while also
providing data on their prevalence and relationship to demographic factors. Underlying mechanisms
are explored with keen attention to psychosocial and biochemical processes. Then, options are
discussed for assessment and treatment. Psychological tests for anxiety, depression, and anger are
pitted against one another to allow the selection of the best. Treatment strategies of both the
psychological and pharmacological varieties are evaluated for effectiveness and side effects. Thus for
anxiety, information is provided on tranquilizers as well as attention-diversion, thought-stopping,
reappraisal, respiratory regulation, muscle relaxation, biofeedback, music, hypnosis, and massage.
Depression treatment is described with reference to psychodynamic and cognitive therapies but with
an in-depth analysis of whether antidepressant medications actually relieve pain or depression. For
anger, a case is made for the novel integration of cognitive, behavioral, and experiential strategies.
The final chapter succinctly summarizes all the main findings while also suggesting ideas for future
study. The book is practical in its objectives to the very end. What gives it particular strength
is the heavy reliance on empirical evidence and theory. In short, this book unravels the complex
interactions among pain, anxiety, depression, and anger – consistently sounding its relevance to
pain sufferers, pain clinicians, scholars, and students in this field.

Abstract: Structural equation models are presented as alternative models for examining longitudinal data. The models include (a) a cross-lagged regression model, (b) a factor model based on latent growth curves, and (c) a dynamic model based on latent difference scores. The illustrative data are on motivation and perceived competence of students during their first semester in high school. The 3 models yielded different results and such differences were discussed in terms of the conceptualization of change underlying each model. The last model was defended as the most reasonable for these data because it captured the dynamic interrelations between the examined constructs and, at the same time, identified potential growth in the variables.

Abstract: Panel data, which consist of information gathered from the same individuals or units at several different points in time, are commonly used in the social sciences to test theories of individual and social change. This book provides an overview of models that are appropriate for the analysis of panel data, focusing specifically on the area where panels offer major advantages over cross-sectional research designs: the analysis of causal interrelationships among variables. Without "painting" panel data as a cure all for the problems of causal inference in nonexperimental research, the author shows how panel data offer multiple ways of strengthening the causal inference process. In addition, he shows how to estimate models that contain a variety of lag specifications, reciprocal

effects, and imperfectly measured variables. Appropriate for readers who are familiar with multiple regression analysis and causal modeling, this book will offer readers the highlights of developments in this technique from diverse disciplines to analytic traditions.

Finlay et al.: Leisure activities, the social weekend, and alcohol use: Evidence from a daily study of first-year college students   Finlay-Ram-Maggs-etal-2012

Abstract: Objective: The aim of this study was to document within-person and between-persons associations between the duration of day-to-day activities (volunteering, spiritual activities, media use, socializing, entertainment/campus events and clubs, athletics, classes, working for pay) and alcohol use (quantity and heavy drinking) and to examine whether these associations differed by gender and the time of week. Method: First-semester college students ($N = 717$ persons; 51.6% female) provided up to 14 consecutive days of data ($N = 9,431$ days) via daily web-based surveys. Multilevel analyses tested whether alcohol use was associated with activity duration, gender, and time of week. Results: Between-persons associations indicated that alcohol use was higher among individuals who spent more time involved in athletics and socializing and lower among students who spent more time in spiritual and volunteer activities. Within-person associations indicated that students consumed more alcohol and were more likely to drink heavily on weekends, on days they spent more time than usual socializing, and on days they spent less time than usual in spiritual activities and using media. Conclusions: Select activities and days were linked with less alcohol use at both the between- and within-person levels, suggesting that attention should be paid to both selection effects and social context to understand the mechanisms linking activity duration and student drinking.

Abstract: Individual variation is increasingly recognized as important to psychopathology research. Concurrently, new methods of analysis based on network models are bringing new perspectives on mental (dys)function. This current work analyzed idiographic multivariate time series data using a novel network methodology that incorporates contemporaneous and lagged associations in mood and anxiety symptomatology. Data were taken from 40 individuals with generalized anxiety disorder (GAD), major depressive disorder (MDD), or comorbid GAD and MDD, who answered questions about 21 descriptors of mood and anxiety symptomatology 4 times a day over a period of approximately 30 days. The model provided an excellent fit to the intraindividual symptom dynamics of all 40 individuals. The most central symptoms in contemporaneous systems were those related to positive and negative mood. The temporal networks highlighted the importance of anger to symptomatology, while also finding that depressed mood and worry—the principal diagnostic criteria for GAD and MDD—were the least influential nodes across the sample. The method's potential for analysis of individual symptom patterns is demonstrated by 3 exemplar participants. Idiographic network-based analysis may fundamentally alter the way psychopathology is assessed, classified, and treated, allowing researchers and clinicians to better understand individual symptom dynamics.

Abstract: This volume provides a psychobiological perspective on people who experience chronic pain and describes a comprehensive approach to their treatment. The text focuses on the interaction of psychosocial (psychological, behavioral, and social) and physiological processes in people with chronic pain and the implications that follow. Our basic hypothesis is that chronic pain is a learned response, whereby "pain memories" rather than current nociceptive input determine much of the pain experienced. Moreover, interdisciplinary approaches that integrate psychological principles and approaches with traditional biomedical knowledge in the assessment and treatment of people with chronic pain are more fruitful than any single modalities, be they physical (surgery, medication, regional anesthesia, or neuroaugmentive interventions) or psychological (biofeedback, counseling, or psychotherapy). Although our emphasis is on the role of psychological and social factors in chronic pain states, we attempt to integrate these aspects with the current biological understanding of the neurophysiology of nociception.

Flora et al.: An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. Flora-Curran-2004

Abstract: Confirmatory factor analysis (CFA) is widely used for examining hypothesized relations among ordinal variables (e.g., Likert-type items). A theoretically appropriate method fits the CFA model to polychoric correlations using either weighted least squares (WLS) or robust WLS. Importantly, this approach assumes that a continuous, normal latent process determines each observed variable. The extent to which violations of this assumption undermine CFA estimation is not well-known. In this article, the authors empirically study this issue using a computer simulation study. The results suggest that estimation of polychoric correlations is robust to modest violations of underlying normality. Further, WLS performed adequately only at the largest sample size but led to

substantial estimation difficulties with smaller samples. Finally, robust WLS performed well across all conditions.

Fridberg et al.: Examining features of transdermal alcohol biosensor readings: A promising approach with implications for research and intervention

Fridberg-Wang-Porges-2022

Daniel J. Fridberg, Yan Wang, and Eric Porges. "Examining features of transdermal alcohol biosensor readings: A promising approach with implications for research and intervention". In: *Alcoholism: Clinical and Experimental Research* 46.4 (Feb. 2022), pp. 514–516. ISSN: 1530-0277. DOI: 10.1111/acer.14794.

Fritz et al.: Required sample size to detect the mediated effect Fritz-MacKinnon-2007

Matthew S. Fritz and David P. MacKinnon. "Required sample size to detect the mediated effect". In: *Psychological Science* 18.3 (Mar. 2007), pp. 233–239. DOI: 10.1111/j.1467-9280.2007.01882.x.

Abstract: Mediation models are widely used, and there are many tests of the mediated effect. One of the most common questions that researchers have when planning mediation studies is, "How many subjects do I need to achieve adequate power when testing for mediation?" This article presents the necessary sample sizes for six of the most common and the most recommended tests of mediation for various combinations of parameters, to provide a guide for researchers when designing studies or applying for grants.

Fritz et al.: Explanation of two anomalous results in statistical mediation analysis

Fritz-Taylor-MacKinnon-2012

Matthew S. Fritz, Aaron B. Taylor, and David P. MacKinnon. "Explanation of two anomalous results in statistical mediation analysis". In: *Multivariate Behavioral Research* 47.1 (Feb. 2012), pp. 61–87. DOI: 10.1080/00273171.2012.640596.

Abstract: Previous studies of different methods of testing mediation models have consistently found two anomalous results. The first result is elevated Type I error rates for the bias-corrected and accelerated bias-corrected bootstrap tests not found in nonresampling tests or in resampling tests that did not include a bias correction. This is of special concern as the bias-corrected bootstrap is often recommended and used due to its higher statistical power compared with other tests. The second result is statistical power reaching an asymptote far below 1.0 and in some conditions even declining slightly as the size of the relationship between X and M, a, increased. Two computer simulations were conducted to examine these findings in greater detail. Results from the first simulation found that the increased Type I error rates for the bias-corrected and accelerated bias-corrected bootstrap are a function of an interaction between the size of the individual paths making up the mediated effect and the sample size, such that elevated Type I error rates occur when the sample size is small and the effect size of the nonzero path is medium or larger. Results from the second simulation found that stagnation and decreases in statistical power as a function of the effect size of the a path occurred primarily when the path between M and Y, b, was small. Two empirical mediation examples are provided using data from a steroid prevention and health promotion program aimed at high school football players (Athletes Training and Learning to Avoid Steroids; Goldberg et al., 1996), one to illustrate a possible Type I error for the bias-corrected bootstrap test and a second to illustrate a loss in power related to the size of a. Implications of these findings are discussed.

Gatchel et al.: The biopsychosocial approach to chronic pain: Scientific advances and future directions.   Gatchel-Peng-Peters-etal-2007

Abstract: The prevalence and cost of chronic pain is a major physical and mental health care problem in the United States today. As a result, there has been a recent explosion of research on chronic pain, with significant advances in better understanding its etiology, assessment, and treatment. The

purpose of the present article is to provide a review of the most noteworthy developments in the field. The biopsychosocial model is now widely accepted as the most heuristic approach to chronic pain. With this model in mind, a review of the basic neuroscience processes of pain (the bio part of biopsychosocial), as well as the psychosocial factors, is presented. This spans research on how psychological and social factors can interact with brain processes to influence health and illness as well as on the development of new technologies, such as brain imaging, that provide new insights into brain-pain mechanisms.

Gates et al.: Intensive longitudinal analysis of human processes

<div align="right">

Gates-Chow-Molenaar-2023

</div>

Abstract: This book focuses on a span of statistical topics relevant to researchers who seek to conduct person-specific analysis of human data. Our purpose is to provide one consolidated resource that includes techniques from disciplines such as engineering, physics, statistics, and quantitative psychology and outlines their application to data often seen in human research. The book balances mathematical concepts with information needed for using these statistical approaches in applied settings, such as interpretative caveats and issues to consider when selecting an approach. The statistical topics covered here include foundational material as well as state-of-the-art methods. These analytic approaches can be applied to a range of data types such as psychophysiological, self-report, and passively collected measures such as those obtained from smartphones. We provide examples using varied data sources including functional MRI (fMRI), daily diary, and ecological momentary assessment data. Features: Description of time series, measurement, model building, and network methods for person-specific analysis Discussion of the statistical methods in the context of human research Empirical and simulated data examples used throughout the book R code for analyses and recorded lectures for each chapter available at the book website: https://www.personspecific.com/

Across various disciplines of human study, researchers are increasingly seeking to conduct person-specific analysis. This book provides comprehensive information, so no prior knowledge of these methods is required. We aim to reach active researchers who already have some understanding of basic statistical testing. Our book provides a comprehensive resource for those who are just beginning to learn about person-specific analysis as well as those who already conduct such analysis but seek to further deepen their knowledge and learn new tools.

Gates et al.: Automatic search for fMRI connectivity mapping: An alternative to Granger causality testing using formal equivalences among SEM path modeling, VAR, and unified SEM   Gates-Molenaar-Hillary-etal-2010

Abstract: Modeling the relationships among brain regions of interest (ROIs) carries unique potential to explicate how the brain orchestrates information processing. However, hurdles arise when using functional MRI data. Variation in ROI activity contains sequential dependencies and shared influences on synchronized activation. Consequently, both lagged and contemporaneous relationships must be considered for unbiased statistical parameter estimation. Identifying these relationships using a data-driven approach could guide theory-building regarding integrated processing. The present paper demonstrates how the unified SEM attends to both lagged and contemporaneous influences on ROI activity. Additionally, this paper offers an approach akin to Granger causality testing, Lagrange multiplier testing, for statistically identifying directional influence among ROIs and employs this approach using an automatic search procedure to arrive at the optimal model. Rationale for this equivalence is offered by explicating the formal relationships among path modeling, vector autoregression, and unified SEM. When applied to simulated data, biases in estimates

which do not consider both lagged and contemporaneous paths become apparent. Finally, the use of unified SEM with the automatic search procedure is applied to an empirical data example.

George et al.: Concurrent and subsequent associations between daily digital technology use and high-risk adolescents' mental health symptoms

George-Russell-Piontak-etal-2017

Madeleine J. George et al. "Concurrent and subsequent associations between daily digital technology use and high-risk adolescents' mental health symptoms". In: *Child Development* 89.1 (May 2017), pp. 78–88. ISSN: 1467-8624. DOI: 10.1111/cdev.12819.

Abstract: Adolescents are spending an unprecedented amount of time using digital technologies (especially mobile technologies), and there are concerns that adolescents' constant connectivity is associated with poor mental health, particularly among at-risk adolescents. Participants included 151 adolescents at risk for mental health problems ($M_{\mathrm{age}} = 13.1$) who completed a baseline assessment, 30-day ecological momentary assessment, and 18 month follow-up assessment. Results from multilevel regression models showed that daily reports of both time spent using digital technologies and the number of text messages sent were associated with increased same-day attention deficit hyperactivity disorder (ADHD) and conduct disorder (CD) symptoms. Adolescents' reported digital technology usage and text messaging across the ecological momentary assessment (EMA) period was also associated with poorer self-regulation and increases in conduct problem symptoms between the baseline and follow-up assessments.

Georgeson et al.: A sensitivity analysis for temporal bias in cross-sectional mediation

Georgeson-AlvarezBartolo-MacKinnon-2025

A. R. Georgeson, Diana Alvarez-Bartolo, and David P. MacKinnon. "A sensitivity analysis for temporal bias in cross-sectional mediation". In: *Psychological Methods* 30.6 (Dec. 2025), pp. 1326–1344. ISSN: 1082-989X. DOI: 10.1037/met0000628.

Abstract: For over three decades, methodologists have cautioned against the use of cross-sectional mediation analyses because they yield biased parameter estimates. Yet, cross-sectional mediation models persist in practice and sometimes represent the only analytic option. We propose a sensitivity analysis procedure to encourage a more principled use of cross-sectional mediation analysis, drawing inspiration from Gollob and Reichardt (1987, 1991). The procedure is based on the two-wave longitudinal mediation model and uses phantom variables for the baseline data. After a researcher provides ranges of possible values for cross-lagged, autoregressive, and baseline Y and M correlations among the phantom and observed variables, they can use the sensitivity analysis to identify longitudinal conditions in which conclusions from a cross-sectional model would differ most from a longitudinal model. To support the procedure, we first show that differences in sign and effect size of the b-path occur most often when the cross-sectional effect size of the b-path is small and the cross-lagged and the autoregressive correlations are equal or similar in magnitude. We then apply the procedure to cross-sectional analyses from real studies and compare the sensitivity analysis results to actual results from a longitudinal mediation analysis. While no statistical procedure can replace longitudinal data, these examples demonstrate that the sensitivity analysis can recover the effect that was actually observed in the longitudinal data if provided with the correct input information. Implications of the routine application of sensitivity analysis to temporal bias are discussed. R code for the procedure is provided in the online supplementary materials.

Gollob et al.: Taking account of time lags in causal models   Gollob-Reichardt-1987

Abstract: Although it takes time for a cause to exert an effect, causal models often fail to allow adequately for time lags. In particular, causal models that contain cross-sectional relations (i. e., relations between values of 2 variables at the same time) are unsatisfactory because (a) they omit the values of variables at prior times, (b) they omit effects that variables can have on themselves, and (c) they fail to specify the length of the causal interval that is being studied. These omissions

can produce severe biases in estimates of the size of causal effects. Longitudinal models also can fail to take account of time lags properly, and this too can lead to severely biased estimates. The discussion illustrates the biases that can occur in both cross-sectional and longitudinal models, introduces the latent longitudinal approach to causal modeling, and shows how latent longitudinal models can be used to reduce bias by taking account of time lags even when data are available for only 1 point in time.

Gollob et al.: Interpreting and estimating indirect effects assuming time lags really matter   Gollob-Reichardt-1991

Harry F. Gollob and Charles S. Reichardt. "Interpreting and estimating indirect effects assuming time lags really matter". In: *Best methods for the analysis of change: Recent advances, unanswered questions, future directions.* Ed. by Linda M. Collins and John L. Horn. Washington, DC: American Psychological Association, 1991, pp. 243–259. ISBN: 978-1-55798-113-4. DOI: 10.1037/10099-015.

Goodman: On the exact variance of products   Goodman-1960

Leo A. Goodman. "On the exact variance of products". In: *Journal of the American Statistical Association* 55.292 (Dec. 1960), pp. 708–713. DOI: 10.1080/01621459.1960.10483369.

Abstract: A simple exact formula for the variance of the product of two random variables, say, x and y, is given as a function of the means and central product-moments of x and y. The usual approximate variance formula for xy is compared with this exact formula; e.g., we note, in the special case where x and y are independent, that the "variance" computed by the approximate formula is less than the exact variance, and that the accuracy of the approximation depends on the sum of the reciprocals of the squared coefficients of variation of x and y. The case where x and y need not be independent is also studied, and exact variance formulas are presented for several different "product estimates." (The usefulness of exact formulas becomes apparent when the variances of

these estimates are compared.) When x and y are independent, simple unbiased estimates of these exact variances are suggested; in the more general case, consistent estimates are presented.

J. W. Graham et al.: How many imputations are really needed? Some practical clarifications of multiple imputation theory   Graham-Olchowski-Gilreath-2007

Abstract: Multiple imputation (MI) and full information maximum likelihood (FIML) are the two most common approaches to missing data analysis. In theory, MI and FIML are equivalent when identical models are tested using the same variables, and when m, the number of imputations performed with MI, approaches infinity. However, it is important to know how many imputations are necessary before MI and FIML are sufficiently equivalent in ways that are important to prevention scientists. MI theory suggests that small values of m, even on the order of three to five imputations, yield excellent results. Previous guidelines for sufficient m are based on relative efficiency, which involves the fraction of missing information (γ) for the parameter being estimated, and m. In the present study, we used a Monte Carlo simulation to test MI models across several scenarios in which γ and m were varied. Standard errors and p-values for the regression coefficient of interest varied as a function of m, but not at the same rate as relative efficiency. Most importantly, statistical power for small effect sizes diminished as m became smaller, and the rate of this power falloff was much greater than predicted by changes in relative efficiency. Based our findings, we recommend that researchers using MI should perform many more imputations than previously considered sufficient. These recommendations are based on γ, and take into consideration one's tolerance for a preventable power falloff (compared to FIML) due to using too few imputations.

Granger: Investigating causal relations by econometric models and cross-spectral methods   Granger-1969

Abstract: There occurs on some occasions a difficulty in deciding the direction of causality between two related variables and also whether or not feedback is occurring. Testable definitions of causality and feedback are proposed and illustrated by use of simple two-variable models. The important problem of apparent instantaneous causality is discussed and it is suggested that the problem often arises due to slowness in recording information or because a sufficiently wide class of possible causal variables has not been used. It can be shown that the cross spectrum between two variables can be decomposed into two parts, each relating to a single causal arm of a feedback situation. Measures of causal lag and causal strength can then be constructed. A generalisation of this result with the partial cross spectrum is suggested.

Greenfield et al.: Risks of alcohol use disorders related to drinking patterns in the U.S. general population   Greenfield-Ye-Bond-etal-2014

Abstract: Objective: The purpose of this study was to examine the relations between drinking (mean quantity and heavy drinking patterns) and alcohol use disorders (AUDs) in the U.S. general population. Method: Data from three telephone National Alcohol Surveys (in 2000, 2005, and 2010) were pooled, with separate analyses for men and women restricted to current drinkers ($ns = 5,922$ men, 6,270 women). Predictors were 12-month volume (mean drinks per day), rates of heavy drinking ($5+/4+$ drinks in a day for men/women), and very heavy drinking ($8+$, $12+$, and $24+$ drinks in a day). Outcomes were negative alcohol-related consequences constituting abuse ($1+$ of 4

DSM-IV-based domains assessed by 13 items) and alcohol dependence (symptoms in 3+ of 7 DSM-IV-based domains), together taken to indicate an AUD. Segmentation analyses were used to model risks of problem outcomes from drinking patterns separately by gender. Results: In the general population, men and women who consumed ≤ 1 drink/day on average with no heavy drinking days did not incur substantial risks of an AUD ($< 10\%$). Men who drank from 1 to 2 drinks/day on average but never 5+ incurred a 16% risk of reporting an AUD (3.5% alcohol dependence). At higher volumes, men and women who indicated higher rates of drinking larger amounts per day and/or involving 8+ and 12+ drinks/day (and even 24+ drinks/day for men) showed much higher risks of experiencing AUDs. Conclusions: The findings provide quantitative guidance for primary care practitioners who wish to make population-based recommendations to patients who might benefit by reducing both overall intake and amounts per occasion in an effort to lower their risks of developing AUDs.

Grundy et al.: Marital conflict and preadolescent behavioral competence: Maternal knowledge as a longitudinal mediator   Grundy-Gondoli-BlodgettSalafia-2007

Abstract: The present study considered whether maternal knowledge mediated the relation between overt marital conflict and preadolescent behavioral competence. Four years of self-report data were collected from 133 mothers and their preadolescents, beginning when the preadolescents were in 4th grade. Marital conflict, maternal knowledge, and preadolescent behavioral competence were assessed at all 4 time points in order to apply a stringent methodology for assessing longitudinal mediating patterns. The results indicated that maternal knowledge mediated the relation between marital conflict and preadolescent behavioral competence. Thus, the present study identified one possible process through which marital conflict may affect preadolescent behavior.

Abstract: Mediation is a causal process that evolves over time. Thus, a study of mediation requires data collected throughout the process. However, most applications of mediation analysis use cross-sectional rather than longitudinal data. Another implicit assumption commonly made in longitudinal designs for mediation analysis is that the same mediation process universally applies to all members of the population under investigation. This assumption ignores the important issue of ergodicity before aggregating the data across subjects. We first argue that there exists a discrepancy between the concept of mediation and the research designs that are typically used to investigate it. Second, based on the concept of ergodicity, we argue that a given mediation process probably is not equally valid for all individuals in a population. Therefore, the purpose of this article is to propose a two-faceted solution. The first facet of the solution is that we advocate a single-subject time-series design that aligns data collection with researchers' conceptual understanding of mediation. The second facet is to introduce a flexible statistical method—the state space model—as an ideal technique to analyze single-subject time series data in mediation studies. We provide an overview of the state space method and illustrative applications using both simulated and real time series data. Finally, we discuss additional issues related to research design and modeling.

Gunn et al.: Characterising patterns of alcohol use among heavy drinkers: A cluster analysis utilising alcohol biosensor data   **Gunn-Steingrimsson-Merrill-etal-2021**

Abstract: Introduction: Previous research has predominately relied on person-level or single characteristics of drinking episodes to characterise patterns of drinking that may confer risk. This research often relies on self-report measures. Advancements in wearable alcohol biosensors provide a multifaceted objective measure of drinking. The current study aimed to characterise drinking episodes using data derived from a wearable alcohol biosensor. Methods: Participants ($n = 45$) were adult heavy drinkers who wore the Secure Continuous Remote Alcohol Monitoring (SCRAM) bracelet and reported on their drinking behaviours. Cluster analysis was used to evaluate unique combinations of alcohol episode characteristics. Associations between clusters and self-reported person and event-level factors were also examined in univariable and multivariable models. Results: Results suggested three unique clusters: Cluster 1 (most common, slowest rate of rise to and decline from peak), Cluster 2 (highest peak transdermal alcohol concentration and area under the curve) and Cluster 3 (fastest rate of decline from peak). Univariable analyses distinguished Cluster 1 as having fewer self-reported drinks and fewer episodes that occurred on weekends relative to Cluster 2. The effect for number of drinks remained in multivariable analyses. Discussion and Conclusions: This is the first study to characterise drinking patterns at the event-level using objective data. Results suggest that it is possible to distinguish drinking episodes based on several characteristics derived from wearable alcohol biosensors. This examination lays the groundwork for future studies to characterise patterns of drinking and their association with consequences of drinking behaviour.

Hall: A short prehistory of the bootstrap   Hall-2003

Abstract: The contemporary development of bootstrap methods, from the time of Efron's early articles to the present day, is well documented and widely appreciated. Likewise, the relationship of bootstrap techniques to certain early work on permutation testing, the jackknife and cross-validation is well understood. Less known, however, are the connections of the bootstrap to research on survey sampling for spatial data in the first half of the last century or to work from the 1940s to the 1970s

on subsampling and resampling. In a selective way, some of these early linkages will be explored, giving emphasis to developments with which the statistics community tends to be less familiar. Particular attention will be paid to the work of P. C. Mahalanobis, whose development in the 1930s and 1940s of moving-block sampling methods for spatial data has a range of interesting features, and to contributions of other scientists who, during the next 40 years, developed half-sampling, subsampling and resampling methods.

Hamaker et al.: At the frontiers of modeling intensive longitudinal data: Dynamic structural equation models for the affective measurements from the COGITO study

<div align="right">Hamaker-Asparouhov-Brose-etal-2018</div>

Abstract: With the growing popularity of intensive longitudinal research, the modeling techniques and software options for such data are also expanding rapidly. Here we use dynamic multilevel modeling, as it is incorporated in the new dynamic structural equation modeling (DSEM) toolbox in Mplus, to analyze the affective data from the COGITO study. These data consist of two samples of over 100 individuals each who were measured for about 100 days. We use composite scores of positive and negative affect and apply a multilevel vector autoregressive model to allow for individual differences in means, autoregressions, and cross-lagged effects. Then we extend the model to include random residual variances and covariance, and finally we investigate whether prior depression affects later depression scores through the random effects of the daily diary measures. We end with discussing several urgent—but mostly unresolved—issues in the area of dynamic multilevel modeling.

Abstract: The current article aims to provide an up-to-date synopsis of available techniques to study affect dynamics using intensive longitudinal data (ILD). We do so by introducing the following eight dichotomies that help elucidate what kind of data one has, what process aspects are of interest, and what research questions are being considered: (1) single- versus multiple-person data; (2) univariate versus multivariate models; (3) stationary versus nonstationary models; (4) linear versus nonlinear models; (5) discrete time versus continuous time models; (6) discrete versus continuous variables; (7) time versus frequency domain; and (8) modeling the process versus computing descriptives. In addition, we discuss what we believe to be the most urging future challenges regarding the modeling of affect dynamics.

Abstract: Results obtained with interindividual techniques in a representative sample of a population are not necessarily generalizable to the individual members of this population. In this article the specific condition is presented that must be satisfied to generalize from the interindividual level to the intraindividual level. A way to investigate whether this condition is satisfied is by means

of multivariate time series analysis. More generally, time series analysis can be used to investigate psychological processes situated within individuals. In this article we consider a well established class of multivariate stationary time series models that may be used to study the intraindividual covariance structure. We demonstrate the application of some of these models with an empirical example consisting of state measurements of behavior associated with the Five Factor Model of Personality. We illustrate how one can investigate whether individuals are similar with respect to their intraindividual structure of variation, and whether this structure is similar to the structure of interindividual variation.

Hamaker et al.: To center or not to center? Investigating inertia with a multilevel autoregressive model   Hamaker-Grasman-2015

Abstract: Whether level 1 predictors should be centered per cluster has received considerable attention in the multilevel literature. While most agree that there is no one preferred approach, it has also been argued that cluster mean centering is desirable when the within-cluster slope and the between-cluster slope are expected to deviate, and the main interest is in the within-cluster slope. However, we show in a series of simulations that if one has a multilevel autoregressive model in which the level 1 predictor is the lagged outcome variable (i.e., the outcome variable at the previous occasion), cluster mean centering will in general lead to a downward bias in the parameter estimate of the within-cluster slope (i.e., the autoregressive relationship). This is particularly relevant if the main question is whether there is on average an autoregressive effect. Nonetheless, we show that if the main interest is in estimating the effect of a level 2 predictor on the autoregressive parameter (i.e., a cross-level interaction), cluster mean centering should be preferred over other forms of centering. Hence, researchers should be clear on what is considered the main goal of their study, and base their choice of centering method on this when using a multilevel autoregressive model.

Abstract: The cross-lagged panel model (CLPM) is believed by many to overcome the problems associated with the use of cross-lagged correlations as a way to study causal influences in longitudinal panel data. The current article, however, shows that if stability of constructs is to some extent of a trait-like, timeinvariant nature, the autoregressive relationships of the CLPM fail to adequately account for this. As a result, the lagged parameters that are obtained with the CLPM do not represent the actual within-person relationships over time, and this may lead to erroneous conclusions regarding the presence, predominance, and sign of causal influences. In this article we present an alternative model that separates the within-person process from stable between-person differences through the inclusion of random intercepts, and we discuss how this model is related to existing structural equation models that include cross-lagged relationships. We derive the analytical relationship between the cross-lagged parameters from the CLPM and the alternative model, and use simulations to demonstrate the spurious results that may arise when using the CLPM to analyze data that include stable, trait-like individual differences. We also present a modeling strategy to avoid this pitfall and illustrate this using an empirical data set. The implications for both existing and future cross-lagged panel research are discussed.

Abstract: In many disciplines researchers use longitudinal panel data to investigate the potentially causal relationship between 2 variables. However, the conventions and concerns vary widely across disciplines. Here we focus on 2 concerns, that is: (a) the concern about random effects versus fixed effects, which is central in the (micro)econometrics/sociology literature; and (b) the concern about grand mean versus group (or person) mean centering, which is central in the multilevel literature associated with disciplines like psychology and educational sciences. We show that these 2 concerns are actually addressing the same underlying issue. We discuss diverse modeling methods based on either multilevel regression modeling with the data in long format, or structural equation modeling with the data in wide format, and compare these approaches with simulated data. We extend the multilevel model with random slopes and discuss the consequences of this. Subsequently, we provide guidelines on how to choose between the diverse modeling options. We illustrate the use of these guidelines with an empirical example based on intensive longitudinal data, in which we consider both a time-varying and a time-invariant covariate.

Hamaker et al.: The integrated trait-state model

Hamaker-Nesselroade-Molenaar-2007

Abstract: It has been acknowledged that both trait and state contribute to psychological measurements. However, existing structural equation models for disentangling these sources of variability are based on assumptions that are not tenable in the light of empirical results. A new model is presented, termed the integrated trait-state (ITS) model, which both decomposes state and trait variance and allows one to test the assumptions that underly existing approaches. This is illustrated with an empirical example. The relationship between the ITS model and other analytic approaches as well as conceptual models of traits and states are discussed.

Hamaker et al.: Using a few snapshots to distinguish mountains from waves: Weak factorial invariance in the context of trait-state research

Ellen L. Hamaker, Noémi K. Schuurman, and Eva A. O. Zijlmans. "Using a few snapshots to distinguish mountains from waves: Weak factorial invariance in the context of trait-state research". In: *Multivariate Behavioral Research* 52.1 (Nov. 2016), pp. 47–60. ISSN: 1532-7906. DOI: 10.1080/00273171.2016.1251299.

Abstract: In this article, we show that the underlying dimensions obtained when factor analyzing cross-sectional data actually form a mix of within-person state dimensions and between-person trait dimensions. We propose a factor analytical model that distinguishes between four independent sources of variance: common trait, unique trait, common state, and unique state. We show that by testing whether there is weak factorial invariance across the trait and state factor structures, we can tackle the fundamental question first raised by Cattell; that is, are within-person state dimensions qualitatively the same as between-person trait dimensions? Furthermore, we discuss how this model is related to other trait-state factor models, and we illustrate its use with two empirical data sets. We end by discussing the implications for cross-sectional factor analysis and suggest potential future developments.

Hamilton: Time series analysis   **Hamilton-1994**

James D. Hamilton. *Time series analysis.* Princeton, NJ: Princeton University Press, 1994. 1799 pp. ISBN: 9780691218632.

Harvey: Forecasting, structural time series models and the Kalman filter

Harvey-1990

Andrew C. Harvey. *Forecasting, structural time series models and the Kalman filter.* Cambridge University Press, Feb. 1990. DOI: 10.1017/cbo9781107049994.

Abstract: In this book, Andrew Harvey sets out to provide a unified and comprehensive theory of structural time series models. Unlike the traditional ARIMA models, structural time series models consist explicitly of unobserved components, such as trends and seasonals, which have a direct interpretation. As a result the model selection methodology associated with structural models is much closer to econometric methodology. The link with econometrics is made even closer by the natural way in which the models can be extended to include explanatory variables and to cope with multivariate time series. From the technical point of view, state space models and the Kalman filter play a key role in the statistical treatment of structural time series models. The book includes a detailed treatment of the Kalman filter. This technique was originally developed in control engineering, but is becoming increasingly important in fields such as economics and operations research. This book is concerned primarily with modelling economic and social time series, and with addressing the special problems which the treatment of such series poses. The properties of the models and the methodological techniques used to select them are illustrated with various applications. These range from the modellling of trends and cycles in US macroeconomic time series to to an evaluation of the effects of seat belt legislation in the UK.

Hatemi-J: A new method to choose optimal lag order in stable and unstable VAR models   HatemiJ-2003

Abstract: A crucial aspect of empirical research based on the vector autoregressive (VAR) model is the choice of the lag order, since all inference in the VAR model is based on the chosen lag order. Here, a new information criterion is introduced for this purpose. The conducted Monte Carlo simulation experiments show that this new information criterion performs well in picking the true lag order in stable as well as unstable VAR models.

Hatemi-J: Multivariate tests for autocorrelation in the stable and unstable VAR models
<div align="right">HatemiJ-2004</div>

Abstract: This study investigates the size and power properties of three multivariate tests for autocorrelation, namely portmanteau test, Lagrange multiplier (LM) test and Rao F-test, in the stable and unstable vector autoregressive (VAR) models, with and without autoregressive conditional heteroscedasticity (ARCH) using Monte Carlo experiments. Many combinations of parameters are used in the simulations to cover a wide range of situations in order to make the results more representative. The results of conducted simulations show that all three tests perform relatively well in stable VAR models without ARCH. In unstable VAR models the portmanteau test exhibits serious size distortions. LM and Rao tests perform well in unstable VAR models without ARCH. These results are true, irrespective of sample size or order of autocorrelation. Another clear result that the simulations show is that none of the tests have the correct size when ARCH is present irrespective of VAR models being stable or unstable and regardless of the sample size or order of autocorrelation. The portmanteau test appears to have slightly better power properties than the LM test in almost all scenarios.

Hayes: Beyond Baron and Kenny: Statistical mediation analysis in the new millennium
<div align="right">Hayes-2009</div>

Abstract: Understanding communication processes is the goal of most communication researchers. Rarely are we satisfied merely ascertaining whether messages have an effect on some outcome of focus in a specific context. Instead, we seek to understand how such effects come to be. What kinds of

causal sequences does exposure to a message initiate? What are the causal pathways through which a message exerts its effect? And what role does communication play in the transmission of the effects of other variables over time and space? Numerous communication models attempt to describe the mechanism through which messages or other communication-related variables transmit their effects or intervene between two other variables in a causal model. The communication literature is replete with tests of such models. Over the years, methods used to test such process models have grown in sophistication. An example includes the rise of structural equation modeling (SEM), which allows investigators to examine how well a process model that links some focal variable X to some outcome Y through one or more intervening pathways fits the observed data. Yet frequently, the analytical choices communication researchers make when testing intervening variables models are out of step with advances made in the statistical methods literature. My goal here is to update the field on some of these new advances. While at it, I challenge some conventional wisdom and nudge the field toward a more modern way of thinking about the analysis of intervening variable effects.

Hayes: Introduction to mediation, moderation, and conditional process analysis: A regression-based approach   Hayes-2022

Abstract: Lauded for its easy-to-understand, conversational discussion of the fundamentals of mediation, moderation, and conditional process analysis, this book has been fully revised with 50% new content, including sections on working with multicategorical antecedent variables, the use of PROCESS version 3 for SPSS and SAS for model estimation, and annotated PROCESS v3 outputs. Using the principles of ordinary least squares regression, Andrew F. Hayes carefully explains procedures for testing hypotheses about the conditions under and the mechanisms by which causal effects operate, as well as the moderation of such mechanisms. Hayes shows how to estimate and interpret direct, indirect, and conditional effects; probe and visualize interactions; test questions

about moderated mediation; and report different types of analyses. Data for all the examples are available on the companion website (www.afhayes.com) along with links to download PROCESS.

Hayes et al.: Using heteroskedasticity-consistent standard error estimators in OLS regression: An introduction and software implementation   Hayes-Cai-2007

Andrew F. Hayes and Li Cai. "Using heteroskedasticity-consistent standard error estimators in OLS regression: An introduction and software implementation". In: *Behavior Research Methods* 39.4 (Nov. 2007), pp. 709–722. DOI: 10.3758/bf03192961.

Hayes et al.: The relative trustworthiness of inferential tests of the indirect effect in statistical mediation analysis   Hayes-Scharkow-2013

Andrew F. Hayes and Michael Scharkow. "The relative trustworthiness of inferential tests of the indirect effect in statistical mediation analysis". In: *Psychological Science* 24.10 (Aug. 2013), pp. 1918–1927. DOI: 10.1177/0956797613480187.

Abstract: A content analysis of 2 years of Psychological Science articles reveals inconsistencies in how researchers make inferences about indirect effects when conducting a statistical mediation analysis. In this study, we examined the frequency with which popularly used tests disagree, whether the method an investigator uses makes a difference in the conclusion he or she will reach, and whether there is a most trustworthy test that can be recommended to balance practical and performance considerations. We found that tests agree much more frequently than they disagree, but disagreements are more common when an indirect effect exists than when it does not. We recommend the bias-corrected bootstrap confidence interval as the most trustworthy test if power is of utmost concern, although it can be slightly liberal in some circumstances. Investigators concerned about Type I errors should choose the Monte Carlo confidence interval or the distribution-of-the-product

approach, which rarely disagree. The percentile bootstrap confidence interval is a good compromise test.

Hecht et al.: Continuous-time modeling in prevention research: An illustration

<div align="right">**Hecht-Voelkle-2019**</div>

Abstract: The analysis of cross-lagged relationships is a popular approach in prevention research to explore the dynamics between constructs over time. However, a limitation of commonly used cross-lagged models is the requirement of equally spaced measurement occasions that prevents the usage of flexible longitudinal designs and complicates cross-study comparisons. Continuous-time modeling overcomes these limitations. In this article, we illustrate the use of continuous-time models using Bayesian and frequentist approaches to model estimation. As an empirical example, we study the dynamic interplay of physical activity and health, a classic research topic in prevention science, using data from the "Midlife in the United States (MIDUS 2): Daily Stress Project, 2004-2009." To help prevention researchers in adopting the approach, we provide annotated R scripts and a simulated data set based on the results from analyzing the MIDUS 2 data.

Hecht et al.: A computationally more efficient Bayesian approach for estimating continuous-time models

<div align="right">**Hecht-Zitzmann-2020a**</div>

Abstract: Continuous-time modeling is gaining in popularity as more and more intensive longitudinal data need to be analyzed. Current Bayesian software implementations of continuous-time

models suffer from rather high, inadequate run times. Therefore, we apply a model reformulation approach to reduce run time. In a simulation study, we investigate the estimation quality and run time gain. We then illustrate our optimized Bayesian continuous-time model estimation and compare it to established continuous-time modeling software using an empirical example. Parameter estimates and inference statistics were very comparable, while run times were very different. Our approach reduces the run times for Bayesian estimations of continuous-time models from hours to minutes.

Hecht et al.: Sample size recommendations for continuous-time models: Compensating shorter time series with larger numbers of persons and vice versa

Abstract: Autoregressive modeling has traditionally been concerned with time-series data from one unit (N = 1). For short time series (T ¡ 50), estimation performance problems are well studied and documented. Fortunately, in psychological and social science research, besides T, another source of information is often available for model estimation, that is, the persons (N ¿ 1). In this work, we illustrate the N/T compensation effect: With an increasing number of persons N at constant T, the model estimation performance increases, and vice versa, with an increasing number of time points T at constant N, the performance increases as well. Based on these observations, we develop sample size recommendations in the form of easily accessible N/T heatmaps for two popular autoregressive continuous-time models.

Hecht et al.: Exploring the unfolding of dynamic effects with continuous-time models: Recommendations concerning statistical power to detect peak cross-lagged effects
Hecht-Zitzmann-2021

Abstract: Cross-lagged panel models have been commonly applied to investigate the dynamic interplay of variables. In such discrete-time models, the size of the cross-lagged effects depends on the length of the time interval between the measurement occasions. Continuous-time modeling allows to explore this interval dependence of cross-lagged effects and thus to identify the maximal "peak" cross-lagged effects. To detect these peak effects, sufficient statistical power is needed. Based on results from a simulation study, we employed machine learning algorithms to identify a highly accurate prediction model. Results are incorporated into a Shiny App (available at https://psychtools.shinyapps.io/ ContinuousTimePowerCalculation) for easy power calculations. Although limitations apply, our results might be helpful for study planning.

Hektner et al.: Experience sampling method: Measuring the quality of everyday life
Hektner-Schmidt-Csikszentmihalyi-2007

Abstract: Experience Sampling Method: Measuring the Quality of Everyday Life is the first book to bring together the theoretical foundations and practical applications of this indispensable methodology. Authors Joel M. Hektner, Jennifer A. Schmidt, and Mihaly Csikszentmihalyi provide fascinating information for anyone interested in how people go about their daily lives. Key Features:

Provides a step-by-step guide: In nontechnical prose, the book details the logistics of carrying out an Experience Sampling Method (ESM) study and guides the reader through every step of the process, from conceiving the research question to analyzing the data. In addition, a through treatment of the measurement of Csikszentmihalyi s flow describes all of the different ways in which flow can be measured. Includes real-life examples: This book gives readers useful tips to consider before implementing a study, based on real-life examples. It illustrates how the ESM has been used to address a diverse array of topics in social science research including the phenomenology of everyday life, gender differences, family relations, work experiences, cross-cultural differences and similarities, school experiences, and mental health. Offers a complete overview of the foundations for ESM: This is the first source to compile findings from a large and increasingly diverse research literature documenting the use of the ESM. A comprehensive overview is given of issues affecting reliability and validity of the method and empirical evidence of its psychometric properties. Intended Audience: This is a must-have resource for social and behavioral scientists who are studying the human experience in everyday life and need guidelines for how to validate and present their data. It can also be used in various advanced undergraduate and graduate research methods courses in the departments of Education, Educational Psychology, Psychology, Nursing, and Health.

Herring et al.: Utility of the Athlete Drinking Scale for assessing drinking motives among high school athletes   Herring-Zamboanga-Olthuis-etal-2016

Abstract: Research suggests that high school athletes are at greater risk for heavy alcohol use and alcohol-related problems than their non-athlete peers. Drinking motives unique to the athletic experience may contribute to elevated use. The Athlete Drinking Scale (ADS) was designed to assess sport-related motives for alcohol use, but has not yet been validated among high school athletes. The purpose of this study was to examine the psychometric properties of the ADS among

a sample of high school athletes. Participants were 216 high school student-athlete drinkers who completed anonymous self-report surveys. A confirmatory factor analysis resulted in a revised three-factor solution with a satisfactory overall model fit. Path analyses indicated that the Positive Reinforcement motives subscale was the only ADS subscale that was significantly associated with alcohol use and alcohol-related problems when controlling for the effects of the other factors (i.e., age and gender) in this population. The ADS may be a valuable assessment tool for researchers and clinicians involved in alcohol prevention efforts for high school athletes.

Hershberger et al.: Modeling intraindividual variability with repeated measures data: Methods and applications   Hershberger-Moskowitz-2002

Scott L. Hershberger and Debbie S. Moskowitz. *Modeling intraindividual variability with repeated measures data: Methods and applications*. Ed. by Scott L. Hershberger and Debbie S. Moskowitz. Psychology Press, 2002. ISBN: 9781410604477.

Abstract: This book examines how individuals behave across time and to what degree that behavior changes, fluctuates, or remains stable. It features the most current methods on modeling repeated measures data as reported by a distinguished group of experts in the field. The goal is to make the latest techniques used to assess intraindividual variability accessible to a wide range of researchers. Each chapter is written in a "user-friendly" style such that even the "novice" data analyst can easily apply the techniques. Each chapter features: 1) a minimum discussion of mathematical detail; 2) an empirical example applying the technique; and 3) a discussion of the software related to that technique. Content highlights include analysis of mixed, multi-level, structural equation, and categorical data models. It is ideal for researchers, professionals, and students working with repeated measures data from the social and behavioral sciences, business, or biological sciences.

Tim C. Hesterberg. *What teachers should know about the bootstrap: Resampling in the undergraduate statistics curriculum*. 2014. arXiv: 1411.5279 [stat.OT]. URL: https://arxiv.org/abs/1411.5279.

Abstract: I have three goals in this article:

1. To show the enormous potential of bootstrapping and permutation tests to help students understand statistical concepts including sampling distributions, standard errors, bias, confidence intervals, null distributions, and P-values.

2. To dig deeper, understand why these methods work and when they don't, things to watch out for, and how to deal with these issues when teaching.

3. To change statistical practice—by comparing these methods to common t tests and intervals, we see how inaccurate the latter are; we confirm this with asymptotics. $n \geq 30$ isn't enough—think $n \geq 5000$.

Resampling provides diagnostics, and more accurate alternatives. Sadly, the common bootstrap percentile interval badly under-covers in small samples; there are better alternatives. The tone is informal, with a few stories and jokes.

Tim C. Hesterberg. "What teachers should know about the bootstrap: Resampling in the undergraduate statistics curriculum". In: *The American Statistician* 69.4 (Oct. 2015), pp. 371–386. DOI: 10.1080/00031305.2015.1089789.

Abstract: Bootstrapping has enormous potential in statistics education and practice, but there are subtle issues and ways to go wrong. For example, the common combination of nonparametric

bootstrapping and bootstrap percentile confidence intervals is less accurate than using t-intervals for small samples, though more accurate for larger samples. My goals in this article are to provide a deeper understanding of bootstrap methods–how they work, when they work or not, and which methods work better-and to highlight pedagogical issues. Supplementary materials for this article are available online.

Higgins et al.: Quantifying heterogeneity in a meta-analysis   Higgins-Thompson-2002

Abstract: The extent of heterogeneity in a meta-analysis partly determines the difficulty in drawing overall conclusions. This extent may be measured by estimating a between-study variance, but interpretation is then specific to a particular treatment effect metric. A test for the existence of heterogeneity exists, but depends on the number of studies in the meta-analysis. We develop measures of the impact of heterogeneity on a meta-analysis, from mathematical criteria, that are independent of the number of studies and the treatment effect metric. We derive and propose three suitable statistics: H is the square root of the χ^2 heterogeneity statistic divided by its degrees of freedom; R is the ratio of the standard error of the underlying mean from a random effects meta-analysis to the standard error of a fixed effect meta-analytic estimate, and I^2 is a transformation of H that describes the proportion of total variation in study estimates that is due to heterogeneity. We discuss interpretation, interval estimates and other properties of these measures and examine them in five example data sets showing different amounts of heterogeneity. We conclude that H and I^2, which can usually be calculated for published meta-analyses, are particularly useful summaries of the impact of heterogeneity. One or both should be presented in published meta-analyses in preference to the test for heterogeneity.

Hingson et al.: Magnitude and trends in heavy episodic drinking, alcohol-impaired driving, and alcohol-related mortality and overdose hospitalizations among emerging adults of college ages 18-24 in the United States, 1998-2014   Hingson-Zha-Smyth-2017

Abstract: Objective: This article estimates percentages of U.S. emerging adults ages 18-24 engaging in past-month heavy episodic drinking and past-year alcohol-impaired driving, and numbers experiencing alcohol-related unintentional injury deaths and overdose hospitalizations between 1998 and 2014. Method: We analyzed national injury mortality data from coroner, census, and college enrollment statistics, the National Survey on Drug Use and Health, and the Nationwide Inpatient Sample. Results: From 1999 to 2005, percentages of emerging adults ages 18-24 reporting past-month heavy episodic drinking rose from 37.1% to 43.1% and then declined to 38.8% in 2014. Alcohol-impaired driving rose from 24% to 25.5% and then declined to 16.0%. Alcohol-related unintentional injury deaths increased from 4,807 in 1998 to 5,531 in 2005 and then declined to 4,105 in 2014, a reduction of 29% per 100,000 since 1998. Alcohol-related traffic deaths increased from 3,783 in 1998 to 4,114 in 2005 and then declined to 2,614 in 2014, down 43% per 100,000 since 1998. Alcohol-related overdose deaths increased from 207 in 1998 to 891 in 2014, a 254% increase per 100,000. Other types of nontraffic unintentional injury deaths declined. Alcohol-overdose hospitalizations rose 26% per 100,000 from 1998 to 2014, especially from increases in alcohol/other drug overdoses, up 61% (alcohol/opioid overdoses up 197%). Conclusions: Among emerging adults, a trend toward increased alcohol-related unintentional injury deaths, heavy episodic drinking, and alcohol-impaired driving between 1998 and 2005 was reversed by 2014. Persistent high levels of heavy episodic drinking and related problems among emerging adults underscore a need to expand individually oriented interventions, college/community collaborative programs, and evidence-supported policies to reduce their drinking and related problems.

123

David V. Hinkley. "Jackknifing in unbalanced situations". In: *Technometrics* 19.3 (Aug. 1977), pp. 285–292. DOI: 10.1080/00401706.1977.10489550.

Abstract: Both the standard jackknife and a weighted jackknife are investigated in the general linear model situation. Properties of bias reduction and standard error estimation are derived and the weighted jackknife shown to be superior for unbalanced data. There is a preliminary discussion of robust regression fitting using jackknife pseudo-values.

Hollenstein: This time, it's real: Affective flexibility, time scales, feedback loops, and the regulation of emotion   Hollenstein-2015

Tom Hollenstein. "This time, it's real: Affective flexibility, time scales, feedback loops, and the regulation of emotion". In: *Emotion Review* 7.4 (July 2015), pp. 308–315. ISSN: 1754-0747. DOI: 10.1177/1754073915590621.

Abstract: Because both emotional arousal and regulation are continuous, ongoing processes, it is difficult, if not impossible, to separate them. Thus, affective dynamics can reveal the regulation of emotion as it occurs in real time. One way that this can be done is through the examination of intra- and interpersonal flexibility or the transitions into and out of affective states. The present article reviews and then expands upon the Flex3 model of real-time dynamic and reactive flexibility, specifying the ways in which individual differences in emotion regulation manifest as differences in flexibility. The differences in results at the real-time scale versus diurnal variability are also discussed within an emotion regulation framework.

Holmes: Bootstrapping phylogenetic trees: Theory and methods   Holmes-2003a

Susan Holmes. "Bootstrapping phylogenetic trees: Theory and methods". In: *Statistical Science* 18.2 (May 2003). DOI: 10.1214/ss/1063994979.

Abstract: This is a survey of the use of the bootstrap in the area of systematic and evolutionary biology. I present the current usage by biologists of the bootstrap as a tool both for making inferences and for evaluating robustness, and propose a framework for thinking about these problems in terms of mathematical statistics.

Holmes: Bradley Efron: A conversation with good friends   Holmes-2003b

Abstract: Bradley Efron is Professor of Statistics and Biostatistics at Stanford University. He works on a combination of theoretical and applied topics, including empirical Bayes, survival analysis, exponential families, bootstrap and jackknife methods and confidence intervals. Most of his applied work has originated in biomedical consulting projects at the Stanford Medical School, mixed in with a few papers concerning astronomy and physics. Even his theoretical papers usually begin with specific applied problems. All three of the interviewers here have been close scientific collaborators. Brad was born in St. Paul, Minnestora, May 1938, to Esther and Miles Efron, Jewish-Russian immigrants. A Merit Scholarship, in the program's inaugural year, brought him to Caltech, graduating in Mathematics in 1960. He arrived at Stanford that Fall, eventually gaining his Ph.D., under the direction of Rupert Miller and Herb Solomon, in the Statistics Department, whose faculty also included Charles Stein, Herman Chernoff, Manny Parzen, Lincoln Moses and Ingram Olkin. Brad has lived at Stanford since 1960, with sabbaticals at Harvard, Imperial College and Berkeley. He has held several administrative positions in the university: Chair of Statistics, Associate Dean of Science, Chairman of the University Advisory Board and Chair of the Faculty Senate. He is currently Chair of the Undergraduate Program in Applied Mathematics. Honors include doctorates from Chicago, Madrid and Oslo, a MacArthur Prize Fellowship, membership in the National Academy of Sciences and the American Academy of Arts and Sciences, fellowship in the IMS and ASA, the Wilks Medal, Parzen Prize, the newly inaugurated Rao Prize and the outstanding statistician award from the Chicago ASA chapter. He has been the Rietz, Wald, and Fisher lecturers and holds

the Max H. Stein endowed chair as Professor of Humanities and Sciences at Stanford. Professional service includes Theory and Methods Editor of JASA and President of the IMS. Currently he is President-Elect of the American Statistical Association, becoming President in 2004.

Hoogland et al.: Robustness studies in covariance structure modeling: An overview and a meta-analysis — Hoogland-Boomsma-1998

Jeffrey J. Hoogland and Anne Boomsma. "Robustness studies in covariance structure modeling: An overview and a meta-analysis". In: *Sociological Methods & Research* 26.3 (Feb. 1998), pp. 329–367. ISSN: 1552-8294. DOI: 10.1177/0049124198026003003.

Abstract: In covariance structure modeling several estimation methods are available. The robustness of an estimator against specific violations of assumptions can be determined empirically by means of a Monte Carlo study. Many such studies in covariance structure analysis have been published, but the conclusions frequently seem to contradict each other An overview of robustness studies in covariance structure analysis is given, and an attempt is made to generalize their findings. Robustness studies are described and distinguished from each other systematically by means of certain characteristics. These characteristics serve as explanatory variables in a meta-analysis concerning the behavior of parameter estimators, standard error estimators, and goodness-of-fit statistics when the model is correctly specified.

S. D. Horn et al.: Estimating heteroscedastic variances in linear models

Horn-Horn-Duncan-1975

Susan D. Horn, Roger A. Horn, and David B. Duncan. "Estimating heteroscedastic variances in linear models". In: *Journal of the American Statistical Association* 70.350 (June 1975), pp. 380–385. DOI: 10.1080/01621459.1975.10479877.

Abstract: This paper presents examples of problems in estimation and hypothesis testing that demonstrate the use and performance of the bootstrap in econometric settings. The examples are illustrated with two empirical applications. The paper concludes with a discussion of topics on which further research is needed.

Houben et al.: The relation between short-term emotion dynamics and psychological well-being: A meta-analysis Houben-VanDenNoortgate-Kuppens-2015

Abstract: Not only how good or bad people feel on average, but also how their feelings fluctuate across time is crucial for psychological health. The last 2 decades have witnessed a surge in research linking various patterns of short-term emotional change to adaptive or maladaptive psychological functioning, often with conflicting results. A meta-analysis was performed to identify consistent relationships between patterns of short-term emotion dynamics—including patterns reflecting emotional variability (measured in terms of within-person standard deviation of emotions across time), emotional instability (measured in terms of the magnitude of consecutive emotional changes), and emotional inertia of emotions over time (measured in terms of autocorrelation)—and relatively stable indicators of psychological well-being or psychopathology. We determined how such relationships are moderated by the type of emotional change, type of psychological well-being or psychopathology involved, valence of the emotion, and methodological factors. A total of 793 effect sizes were identified from 79 articles ($N = 11,381$) and were subjected to a 3-level meta-analysis.

The results confirmed that overall, low psychological well-being co-occurs with more variable (overall $\hat{\rho} = -.178$), unstable (overall $\hat{\rho} = -.205$), but also more inert (overall $\hat{\rho} = -.151$) emotions. These effect sizes were stronger when involving negative compared with positive emotions. Moreover, the results provided evidence for consistency across different types of psychological well-being and psychopathology in their relation with these dynamical patterns, although specificity was also observed. The findings demonstrate that psychological flourishing is characterized by specific patterns of emotional fluctuations across time, and provide insight into what constitutes optimal and suboptimal emotional functioning.

Hunter: State space modeling in an open source, modular, structural equation modeling environment   Hunter-2017

Abstract: State space models (SSMs) are introduced in the context of structural equation modeling (SEM). In particular, the OpenMx implementation of SSMs using the Kalman filter and prediction error decomposition is discussed. In reflection of modularity, the implementation uses the same full information maximum likelihood missing data procedures for SSMs and SEMs. Similarly, generic OpenMx features such as likelihood ratio tests, profile likelihood confidence intervals, Hessian-based standard errors, definition variables, and the matrix algebra interface are all supported. Example scripts for specification of autoregressive models, multiple lag models (VAR(p)), multiple lag moving average models (VARMA(p, q)), multiple subject models, and latent growth models are provided. Additionally, latent variable calculation based on the Kalman filter and raw data generation based on a model are all included. Finally, future work for extending SSMs to allow for random effects and for presenting them in diagrams is discussed.

A. M. Hussong et al.: An internalizing pathway to alcohol use and disorder.

Andrea M. Hussong et al. "An internalizing pathway to alcohol use and disorder." In: *Psychology of Addictive Behaviors* 25.3 (Sept. 2011), pp. 390–404. ISSN: 0893-164X. DOI: 10.1037/a0024519.

Abstract: Research emanating from the field of developmental science indicates that initial risk factors for alcohol use and disorder can be evident in early childhood. One dominant developmental pathway connecting these initial risk factors with subsequent alcohol involvement focuses on the central role of disinhibited or externalizing behaviors. In the current paper, we delineate a second pathway that focuses on internalizing symptomatology. Several studies indicate that internalizing symptoms in early and middle childhood predict alcohol involvement in adolescence and young adulthood. We use a developmental psychopathology framework to describe a risk model that traces the potential developmental markers of this internalizing pathway and to consider the relation between the internalizing pathway and the more widely researched externalizing pathway. We outline the markers of risk in this pathway and conclude with a discussion of the implications of this model for prevention efforts and future research. In this manner, we strive for a translational goal, linking our existing understanding of internalizing processes and alcohol use and disorder with our efforts to develop effective prevention programs.

Iacus: Simulation and Inference for Stochastic Differential Equations **Iacus-2008**

Stefano M. Iacus. *Simulation and Inference for Stochastic Differential Equations*. Springer New York, 2008. DOI: 10.1007/978-0-387-75839-8.

Jackson et al.: Quantifying the impact of between-study heterogeneity in multivariate meta-analyses **Jackson-White-Riley-2012**

Dan Jackson, Ian R. White, and Richard D. Riley. "Quantifying the impact of between-study

heterogeneity in multivariate meta-analyses". In: *Statistics in Medicine* 31.29 (July 2012), pp. 3805–3820. ISSN: 1097-0258. DOI: 10.1002/sim.5453.

Abstract: Measures that quantify the impact of heterogeneity in univariate meta-analysis, including the very popular I^2 statistic, are now well established. Multivariate meta-analysis, where studies provide multiple outcomes that are pooled in a single analysis, is also becoming more commonly used. The question of how to quantify heterogeneity in the multivariate setting is therefore raised. It is the univariate R^2 statistic, the ratio of the variance of the estimated treatment effect under the random and fixed effects models, that generalises most naturally, so this statistic provides our basis. This statistic is then used to derive a multivariate analogue of I^2, which we call I_R^2. We also provide a multivariate H^2 statistic, the ratio of a generalisation of Cochran's heterogeneity statistic and its associated degrees of freedom, with an accompanying generalisation of the usual I^2 statistic, I_H^2. Our proposed heterogeneity statistics can be used alongside all the usual estimates and inferential procedures used in multivariate meta-analysis. We apply our methods to some real datasets and show how our statistics are equally appropriate in the context of multivariate meta-regression, where study level covariate effects are included in the model. Our heterogeneity statistics may be used when applying any procedure for fitting the multivariate random effects model.

James et al.: Mediators, moderators, and tests for mediation   James-Brett-1984

Lawrence R. James and Jeanne M. Brett. "Mediators, moderators, and tests for mediation". In: *Journal of Applied Psychology* 69.2 (1984), pp. 307–321. DOI: 10.1037/0021-9010.69.2.307.

Abstract: Discusses mediation relations in causal terms. Influences of an antecedent are transmitted to a consequence through an intervening mediator. Mediation relations may assume a number of functional forms, including nonadditive, nonlinear, and nonrecursive forms. Although mediation and moderation are distinguishable processes, with nonadditive forms (moderated mediation) a particular variable may be both a mediator and a moderator within a single set of functional relations. Current models for testing mediation relations in industrial and organizational psychology

often involve an interplay between exploratory (correlational) statistical tests and causal inference. It is suggested that no middle ground exists between exploratory and confirmatory (causal) analysis and that attempts to explain how mediation processes occur require specified causal models.

Jensen et al.: Contributions of psychology to the understanding and treatment of people with chronic pain: Why it matters to ALL psychologists   Jensen-Turk-2014

Abstract: Chronic pain is a prevalent problem with significant costs to individuals, significant others, and society. In this article, which introduces the American Psychologist special issue on chronic pain, we provide an overview of the seminal contributions made by psychologists to our current understanding of this important problem. We also describe the primary treatments that have been developed based on psychological principles and models of pain, many of which have demonstrated efficacy for reducing pain and its impact on psychological and physical functioning. The article ends with an enumeration of directions for future research and clinical practice. We believe that the chronicle of psychology's role in improving our understanding and treatment of pain provides a model for how psychologists can have a significant influence on many fields, and that the models and approaches developed for understanding and treating pain may be of use to psychologists working in other areas. Thus, we think that chronic pain is an important area of study that offers insights about translational research for ALL psychologists.

Johnson et al.: The Johnson-Neyman technique, its theory and application
Johnson-Fay-1950

Abstract: The theoretical basis for the Johnson-Neyman Technique is here presented for the first time in an American journal. In addition, a simplified working procedure is outlined, step-by-step, for an actual problem. The determination of significance is arrived at early in the analysis; and where no significant difference is found, the problem is complete at this point. The plotting of the region of significance where a significant difference does exist has also been simplified by using the procedure of rotation and translation of axes.

Johnson et al.: Tests of certain linear hypotheses and their application to some educational problems   Johnson-Neyman-1936

Palmer O. Johnson and Jerzy Neyman. "Tests of certain linear hypotheses and their application to some educational problems". In: *Statistical Research Memoirs* 1 (1936), pp. 57–93.

Abstract: Beginning with the general ideas of testing hypotheses developed by Neyman and Pearson and using certain recent results of S. Kolodziejczyk, the problem of matched groups is discussed and a numerical illustration given. It is shown that the problem of matched groups may be generalized so that both a more detailed analysis of the experimental data and a greater accuracy of results is obtained. In treating this problem the idea of "region of significance"' is introduced to educational and psychological investigations. The methods proposed, however, are quite general and not limited to problems in these fields.

J. A. Jones et al.: Computing confidence intervals for standardized regression coefficients.   Jones-Waller-2013a

Jeff A. Jones and Niels G. Waller. "Computing confidence intervals for standardized regression coefficients." In: *Psychological Methods* 18.4 (2013), pp. 435–453. DOI: 10.1037/a0033269.

Abstract: With fixed predictors, the standard method (Cohen, Cohen, West, & Aiken, 2003, p. 86; Harris, 2001, p. 80; Hays, 1994, p. 709) for computing confidence intervals (CIs) for standardized regression coefficients fails to account for the sampling variability of the criterion standard deviation.

132

With random predictors, this method also fails to account for the sampling variability of the predictor standard deviations. Nevertheless, under some conditions the standard method will produce CIs with accurate coverage rates. To delineate these conditions, we used a Monte Carlo simulation to compute empirical CI coverage rates in samples drawn from 36 populations with a wide range of data characteristics. We also computed the empirical CI coverage rates for 4 alternative methods that have been discussed in the literature: noncentrality interval estimation, the delta method, the percentile bootstrap, and the bias-corrected and accelerated bootstrap. Our results showed that for many data-parameter configurations–for example, sample size, predictor correlations, coefficient of determination (R^2), orientation of β with respect to the eigenvectors of the predictor correlation matrix, R_X–the standard method produced coverage rates that were close to their expected values. However, when population R^2 was large and when β approached the last eigenvector of R_X, then the standard method coverage rates were frequently below the nominal rate (sometimes by a considerable amount). In these conditions, the delta method and the 2 bootstrap procedures were consistently accurate. Results using noncentrality interval estimation were inconsistent. In light of these findings, we recommend that researchers use the delta method to evaluate the sampling variability of standardized regression coefficients.

J. A. Jones et al.: The normal-theory and asymptotic distribution-free (ADF) covariance matrix of standardized regression coefficients: Theoretical extensions and finite sample behavior   Jones-Waller-2013b

Abstract: Yuan and Chan (2011) recently showed how to compute the covariance matrix of standardized regression coefficients from covariances. In this paper, we describe a new method for computing this covariance matrix from correlations. We then show that Yuan and Chan's origi-

nal equations can also be used when only correlational data are available. Next, we describe an asymptotic distribution-free (ADF; Browne, 1984) method for computing the covariance matrix of standardized regression coefficients. We show that theADF method works well with non-normal data in moderate-to-large samples using both simulated and real-data examples. Finally, we provide R code (R Development Core Team, 2012) in an Appendix to make these methods accessible to applied researchers.

J. A. Jones et al.: The normal-theory and asymptotic distribution-free (ADF) covariance matrix of standardized regression coefficients: Theoretical extensions and finite sample behavior   Jones-Waller-2015

Jeff A. Jones and Niels G. Waller. "The normal-theory and asymptotic distribution-free (ADF) covariance matrix of standardized regression coefficients: Theoretical extensions and finite sample behavior". In: *Psychometrika* 80.2 (June 2015), pp. 365–378. DOI: 10.1007/s11336-013-9380-y.

Abstract: Yuan and Chan (Psychometrika, 76, 670-690, 2011) recently showed how to compute the covariance matrix of standardized regression coefficients from covariances. In this paper, we describe a method for computing this covariance matrix from correlations. Next, we describe an asymptotic distribution-free (ADF; Browne in British Journal of Mathematical and Statistical Psychology, 37, 62-83, 1984) method for computing the covariance matrix of standardized regression coefficients. We show that the ADF method works well with nonnormal data in moderate-to-large samples using both simulated and real-data examples. R code (R Development Core Team, 2012) is available from the authors or through the Psychometrika online repository for supplementary materials.

Jorgensen et al.: semTools: Useful tools for structural equation modeling
Jorgensen-Pornprasertmanit-Schoemann-etal-2022

Terrence D. Jorgensen et al. *semTools: Useful tools for structural equation modeling*. 2022. URL: https://CRAN.R-project.org/package=semTools.

Charles M. Judd and David A. Kenny. "Process analysis". In: *Evaluation Review* 5.5 (Oct. 1981), pp. 602–619. DOI: 10.1177/0193841x8100500502.

Abstract: This article presents the rationale and procedures for conducting a process analysis in evaluation research. Such an analysis attempts to identify the process that mediates the effects of some treatment, by estimating the parameters of a causal chain between the treatment and some outcome variable. Two different procedures for estimating mediation are discussed. In addition we present procedures for examining whether a treatment exerts its effects, in part, by altering the mediating process that produces the outcome. Finally, the benefits of process analysis in evaluation research are underlined.

Kalman: A new approach to linear filtering and prediction problems   Kalman-1960

R. E. Kalman. "A new approach to linear filtering and prediction problems". In: *Journal of Basic Engineering* 82.1 (Mar. 1960), pp. 35–45. DOI: 10.1115/1.3662552.

Abstract: The classical filtering and prediction problem is re-examined using the Bode-Shannon representation of random processes and the "state-transition" method of analysis of dynamic systems. New results are: (1) The formulation and methods of solution of the problem apply without modification to stationary and nonstationary statistics and to growing-memory and infinite-memory filters. (2) A nonlinear difference (or differential) equation is derived for the covariance matrix of the optimal estimation error. From the solution of this equation the co-efficients of the difference (or differential) equation of the optimal linear filter are obtained without further calculations. (3) The filtering problem is shown to be the dual of the noise-free regulator problem. The new method developed here is applied to two well-known problems, confirming and extending earlier results. The discussion is largely self-contained and proceeds from first principles; basic concepts of the theory of random processes are reviewed in the Appendix.

Abstract: We test a model that accounts for the adoption of drug use among adolescents in terms of four explanatory perspectives: self-derogation, peer influence, social control, and early substance use. The data come from a three-wave panel study of junior high school students in Houston (N = 3,052). Using nine variables at Time 1, 10 variables at Time 2, and drug use at Time 3, we operationalize components of all four theoretical perspectives in a path model predicting drug use. Results indicate that the four theoretical perspectives complement each other in predicting subsequent adoption of drug use. Significant primary and intervening roles can be attributed to each of the four perspectives. We discuss these findings in terms of an integrative approach to multivariate models of drug use.

Abstract: The sandwich estimator, also known as robust covariance matrix estimator, heteroscedasticity-consistent covariance matrix estimate, or empirical covariance matrix estimator, has achieved increasing use in the econometric literature as well as with the growing popularity of generalized estimating equations. Its virtue is that it provides consistent estimates of the covariance matrix for parameter estimates even when the fitted parametric model fails to hold or is not even specified. Surprisingly though, there has been little discussion of properties of the sandwich method other than

consistency. We investigate the sandwich estimator in quasi-likelihood models asymptotically, and in the linear case analytically. We show that under certain circumstances when the quasi-likelihood model is correct, the sandwich estimate is often far more variable than the usual parametric variance estimate. The increased variance is a fixed feature of the method and the price that one pays to obtain consistency even when the parametric model fails or when there is heteroscedasticity. We show that the additional variability directly affects the coverage probability of confidence intervals constructed from sandwich variance estimates. In fact, the use of sandwich variance estimates combined with t-distribution quantiles gives confidence intervals with coverage probability falling below the nominal value. We propose an adjustment to compensate for this fact.

Kazak: Editorial: Journal article reporting standards — Kazak-2018

Kelley et al.: On effect size — Kelley-Preacher-2012

Abstract: The call for researchers to report and interpret effect sizes and their corresponding confidence intervals has never been stronger. However, there is confusion in the literature on the definition of effect size, and consequently the term is used inconsistently. We propose a definition for effect size, discuss 3 facets of effect size (dimension, measure/index, and value), outline 10 corollaries that follow from our definition, and review ideal qualities of effect sizes. Our definition of effect size is general and subsumes many existing definitions of effect size. We define effect size as *a quantitative reflection of the magnitude of some phenomenon that is used for the purpose of addressing a question of interest.* Our definition of effect size is purposely more inclusive than the way many have defined and conceptualized effect size, and it is unique with regard to linking effect size to a question

of interest. Additionally, we review some important developments in the effect size literature and discuss the importance of accompanying an effect size with an interval estimate that acknowledges the uncertainty with which the population value of the effect size has been estimated. We hope that this article will facilitate discussion and improve the practice of reporting and interpreting effect sizes.

Kenny et al.: Power anomalies in testing mediation   Kenny-Judd-2013

Abstract: Two rather surprising anomalies relating to statistical power occur in testing mediation. First, in a model with no direct effect for which the total effect and indirect effect are identical, the power for the test of the total effect can be dramatically smaller than the power for the test of the indirect effect. Second, when there is a direct effect of a causal variable on the outcome controlling for the mediator, the power of the test of the indirect effect is often considerably greater than the power of the test of the direct effect, even when the two are of the same magnitude. We try to explain the reasons for these anomalies and how they affect practice.

Kenny et al.: Data analysis in social psychology   Kenny-Kashy-Bolger-1998

Abstract: Focuses on structural equation modeling and multilevel modeling. The chapter begins by discussing nonindependence of observations in group research. After considering ANOVA solutions, multilevel modes that can be used to estimate many forms of grouped data are discussed. Identification in structural equation models and the problem of testing mediation are discussed.

Abstract: Multilevel models are increasingly used to estimate models for hierarchical and repeated measures data. The authors discuss a model in which there is mediation at the lower level and the mediational links vary randomly across upper level units. One repeated measures example is a case in which a person's daily stressors affect his or her coping efforts, which affect his or her mood, and both links vary randomly across persons. Where there is mediation at the lower level and the mediational links vary randomly across upper level units, the formulas for the indirect effect and its standard error must be modified to include the covariance between the random effects. Because no standard method can estimate such a model, the authors developed an ad hoc method that is illustrated with real and simulated data. Limitations of this method and characteristics of an ideal method are discussed.

Abstract: Although researchers in clinical psychology routinely gather data in which many individuals respond at multiple times, there is not a standard way to analyze such data. A new approach for the analysis of such data is described. It is proposed that a person's current standing on a variable is caused by 3 sources of variance: a term that does not change (trait), a term that changes (state), and a random term (error). It is shown how structural equation modeling can be used to estimate

such a model. An extended example is presented in which the correlations between variables are quite different at the trait, state, and error levels.

C.-J. Kim et al.: State-space models with regime switching: Classical and Gibbs-sampling approaches with applications   Kim-Nelson-1999

Chang-Jin Kim and Charles R. Nelson. *State-space models with regime switching: Classical and Gibbs-sampling approaches with applications*. The MIT Press, 1999. ISBN: 9780262277112. DOI: 10.7551/mitpress/6444.001.0001.

Abstract: Both state-space models and Markov switching models have been highly productive paths for empirical research in macroeconomics and finance. This book presents recent advances in econometric methods that make feasible the estimation of models that have both features. One approach, in the classical framework, approximates the likelihood function; the other, in the Bayesian framework, uses Gibbs-sampling to simulate posterior distributions from data. The authors present numerous applications of these approaches in detail: decomposition of time series into trend and cycle, a new index of coincident economic indicators, approaches to modeling monetary policy uncertainty, Friedman's "plucking" model of recessions, the detection of turning points in the business cycle and the question of whether booms and recessions are duration-dependent, state-space models with heteroskedastic disturbances, fads and crashes in financial markets, long-run real exchange rates, and mean reversion in asset returns.

Kisbu-Sakarya et al.: The distribution of the product explains normal theory mediation confidence interval estimation   KisbuSakarya-MacKinnon-Miocevic-2014

Yasemin Kisbu-Sakarya, David P. MacKinnon, and Milica Miočević. "The distribution of the product explains normal theory mediation confidence interval estimation". In: *Multivariate Behavioral Research* 49.3 (May 2014), pp. 261–268. DOI: 10.1080/00273171.2014.903162.

Abstract: The distribution of the product has several useful applications. One of these applications is its use to form confidence intervals for the indirect effect as the product of 2 regression coefficients. The purpose of this article is to investigate how the moments of the distribution of the product explain normal theory mediation confidence interval coverage and imbalance. Values of the critical ratio for each random variable are used to demonstrate how the moments of the distribution of the product change across values of the critical ratio observed in research studies. Results of the simulation study showed that as skewness in absolute value increases, coverage decreases. And as skewness in absolute value and kurtosis increases, imbalance increases. The difference between testing the significance of the indirect effect using the normal theory versus the asymmetric distribution of the product is further illustrated with a real data example. This article is the first study to show the direct link between the distribution of the product and indirect effect confidence intervals and clarifies the results of previous simulation studies by showing why normal theory confidence intervals for indirect effects are often less accurate than those obtained from the asymmetric distribution of the product or from resampling methods.

Koob et al.: Addiction and the brain antireward system   Koob-LeMoal-2008

Abstract: A neurobiological model of the brain emotional systems has been proposed to explain the persistent changes in motivation that are associated with vulnerability to relapse in addiction, and this model may generalize to other psychopathology associated with dysregulated motivational systems. In this framework, addiction is conceptualized as a cycle of decreased function of brain reward systems and recruitment of antireward systems that progressively worsen, resulting in the compulsive use of drugs. Counteradaptive processes, such as opponent process, that are part of the normal homeostatic limitation of reward function fail to return within the normal homeostatic range and are hypothesized to repeatedly drive the allostatic state. Excessive drug taking thus

results in not only the short-term amelioration of the reward deficit but also suppression of the antireward system. However, in the long term, there is worsening of the underlying neurochemical dysregulations that ultimately form an allostatic state (decreased dopamine and opioid peptide function, increased corticotropin-releasing factor activity). This allostatic state is hypothesized to be reflected in a chronic deviation of reward set point that is fueled not only by dysregulation of reward circuits per se but also by recruitment of brain and hormonal stress responses. Vulnerability to addiction may involve genetic comorbidity and developmental factors at the molecular, cellular, or neurocircuitry levels that sensitize the brain antireward systems.

Koopman et al.: Pulling the Sobel test up by its bootstraps

Abstract: In the domain of building and testing theory, mediation relationships are among the most important that can be proposed. Mediation helps to explicate our theoretical models (Leavitt, Mitchell, & Peterson, 2010) and addresses the fundamental question of why two constructs are related (Whetten, 1989). One of the better-known methods for testing mediation is commonly referred to as the "Sobel test," named for the researcher who derived a standard error (Sobel, 1982) to test the significance of the indirect effect. Recently, a number of different research teams (e.g., Preacher & Hayes, 2004; Shrout & Bolger, 2002) have criticized the Sobel test because this standard error requires an assumption of normality for the indirect effect sampling distribution. This distribution tends to be positively skewed (i.e,. not normal), particularly in small samples, and so this assumption can be problematic (Preacher & Hayes, 2004; Stone & Sobel, 1990). As a result, the statistical power of the Sobel test may be lessened in these contexts (Preacher & Hayes 2004; Shrout & Bolger, 2002). In light of this concern, some scholars have advocated instead for the use of bootstrapping

to test the significance of the indirect effect (e.g.. Shrout & Bolger 2002). Bootstrapping requires no a priori assumption about the shape of the sampling distribution because this distribution is empirically estimated using a resampling procedure (Efron & Tibshirani, 1993). As a result, departures from normality are less troublesome when creating a confidence interval for the indirect effect. For this reason, bootstrapping is now widely believed to be inherently superior to the Sobel test when testing the significance of the indirect effect in organizational research. Our position is that this belief constitutes an urban legend. As with all statistical urban legends, there is an underlying kernel of truth to the belief that bootstrapping is superior to the Sobel test. However, as we discuss in this chapter, there are several reasons to be concerned with a broad belief in the superiority of bootstrapping. We begin with a brief overview of mediation testing focusing on the Sobel test and bootstrapping and then explain the underlying kernel of truth that has propelled bootstrapping to the forefront of mediation testing in organizational research. Subsequently, we discuss four areas of concern that cast doubt on the belief of the inherent superiority of bootstrapping. Finally, we conclude with recommendations concerning the future of mediation testing in organizational research.

Koopman et al.: Small sample mediation testing: Misplaced confidence in bootstrapped confidence intervals   Koopman-Howe-Hollenbeck-etal-2015

Abstract: Bootstrapping is an analytical tool commonly used in psychology to test the statistical significance of the indirect effect in mediation models. Bootstrapping proponents have particularly advocated for its use for samples of 20-80 cases. This advocacy has been heeded, especially in the Journal of Applied Psychology, as researchers are increasingly utilizing bootstrapping to test mediation with samples in this range. We discuss reasons to be concerned with this escalation, and in a simulation study focused specifically on this range of sample sizes, we demonstrate not

only that bootstrapping has insufficient statistical power to provide a rigorous hypothesis test in most conditions but also that bootstrapping has a tendency to exhibit an inflated Type I error rate. We then extend our simulations to investigate an alternative empirical resampling method as well as a Bayesian approach and demonstrate that they exhibit comparable statistical power to bootstrapping in small samples without the associated inflated Type I error. Implications for researchers testing mediation hypotheses in small samples are presented. For researchers wishing to use these methods in their own research, we have provided R syntax in the online supplemental materials.

Kossakowski et al.: Data from 'Critical slowing down as a personalized early warning signal for depression'   Kossakowski-Groot-Haslbeck-2017

Abstract: We present a dataset of a single (N = 1) participant diagnosed with major depressive disorder, who completed 1478 measurements over the course of 239 consecutive days in 2012 and 2013. The experiment included a double-blind phase in which the dosage of anti-depressant medication was gradually reduced. The entire study looked at momentary affective states in daily life before, during, and after the double-blind phase. The items, which were asked ten times a day, cover topics like mood, physical condition and social contacts. Also, depressive symptoms were measured on a weekly basis using the Symptom Checklist Revised (SCL-90-R). The data are suitable for various time-series analyses and studies in complex dynamical systems.

Kreiss et al.: Bootstrap methods for time series   Kreiss-Lahiri-2012

Handbook of Statistics. Elsevier, 2012, pp. 3–26. ISBN: 9780444538581. DOI: 10.1016/b978-0-444-53858-1.00001-6.

Abstract: The chapter gives a review of the literature on bootstrap methods for time series data. It describes various possibilities on how the bootstrap method, initially introduced for independent random variables, can be extended to a wide range of dependent variables in discrete time, including parametric or nonparametric time series models, autoregressive and Markov processes, long range dependent time series and nonlinear time series, among others. Relevant bootstrap approaches, namely the intuitive residual bootstrap and Markovian bootstrap methods, the prominent block bootstrap methods as well as frequency domain resampling procedures, are described. Further, conditions for consistent approximations of distributions of parameters of interest by these methods are presented. The presentation is deliberately kept non-technical in order to allow for an easy understanding of the topic, indicating which bootstrap scheme is advantageous under a specific dependence situation and for a given class of parameters of interest. Moreover, the chapter contains an extensive list of relevant references for bootstrap methods for time series.

Krull et al.: Multilevel modeling of individual and group level mediated effects

Krull-MacKinnon-2001

Jennifer L. Krull and David P. MacKinnon. "Multilevel modeling of individual and group level mediated effects". In: *Multivariate Behavioral Research* 36.2 (Apr. 2001), pp. 249–277. ISSN: 1532-7906. DOI: 10.1207/s15327906mbr3602_06.

Abstract: This article combines procedures for single-level mediational analysis with multilevel modeling techniques in order to appropriately test mediational effects in clustered data. A simulation study compared the performance of these multilevel mediational models with that of single-level mediational models in clustered data with individual- or group-level initial independent variables, individual- or group-level mediators, and individual level outcomes. The standard errors of mediated effects from the multilevel solution were generally accurate, while those from the single-level

145

procedure were downwardly biased, often by 20% or more. The multilevel advantage was greatest in those situations involving group-level variables, larger group sizes, and higher intraclass correlations in mediator and outcome variables. Multilevel mediational modeling methods were also applied to data from a preventive intervention designed to reduce intentions to use steroids among players on high school football teams. This example illustrates differences between single-level and multilevel mediational modeling in real-world clustered data and shows how the multilevel technique may lead to more accurate results.

Kuiper et al.: Drawing conclusions from cross-lagged relationships: Re-considering the role of the time-interval — Kuiper-Oisin-2018

Abstract: The cross-lagged panel model (CLPM), a discrete-time (DT) SEM model, is frequently used to gather evidence for (reciprocal) Granger-causal relationships when lacking an experimental design. However, it is well known that CLPMs can lead to different parameter estimates depending on the time-interval of observation. Consequently, this can lead to researchers drawing conflicting conclusions regarding the sign and/or dominance of relationships. Multiple authors have suggested the use of continuous-time models to address this issue. In this article, we demonstrate the exact circumstances under which such conflicting conclusions occur. Specifically, we show that such conflicts are only avoided in general in the case of bivariate, stable, nonoscillating, first-order systems, when comparing models with uniform time-intervals between observations. In addition, we provide a range of tools, proofs, and guidelines regarding the comparison of discrete- and continuous-time parameter estimates.

Kunsch: The jackknife and the bootstrap for general stationary observations

Abstract: We extend the jackknife and the bootstrap method of estimating standard errors to the case where the observations form a general stationary sequence. We do not attempt a reduction to i.i.d. values. The jackknife calculates the sample variance of replicates of the statistic obtained by omitting each block of l consecutive data once. In the case of the arithmetic mean this is shown to be equivalent to a weighted covariance estimate of the spectral density of the observations at zero. Under appropriate conditions consistency is obtained if $l = l(n) \to \infty$ and $l(n)/n \to 0$. General statistics are approximated by an arithmetic mean. In regular cases this approximation determines the asymptotic behavior. Bootstrap replicates are constructed by selecting blocks of length l randomly with replacement among the blocks of observations. The procedures are illustrated by using the sunspot numbers and some simulated data.

Kuppens: It's about time: A special section on affect dynamics

Abstract: The study of affect dynamics aims to discover the patterns and regularities with which emotions and affective experiences and components change across time, the underlying mechanisms involved, and their potential relevance for healthy psychological functioning. The intention of this special section is to serve as a mini handbook covering the contemporary state of research into affect dynamics. Contributions address theoretical viewpoints on the origins and functions of emotional change, methodological and modeling approaches, biological and social perspectives on affect dynamics, and the downstream consequences for well-being and psychopathology.

Abstract: In this article, we examine the concept of emotional inertia as a fundamental property of the emotion dynamics that characterize psychological maladjustment. Emotional inertia refers to the degree to which emotional states are resistant to change. Because psychological maladjustment has been associated with both emotional underreactivity and ineffective emotion-regulation skills, we hypothesized that its overall emotion dynamics would be characterized by high levels of inertia. We provide evidence from two naturalistic studies that, using different methods, showed that the emotional fluctuations of individuals who exhibited low self-esteem (Study 1) and depression (Study 2) were characterized by higher levels of inertia in both positive and negative emotions than the emotional fluctuations of people who did not exhibit low self-esteem and depression. We also discuss the usefulness of the concept of emotional inertia as a hallmark of maladaptive emotion dynamics.

Abstract: People display a remarkable variability in the patterns and trajectories with which their feelings change over time. In this article, we present a theoretical account for the dynamics of affect (DynAffect) that identifies the major processes underlying individual differences in the temporal dynamics of affective experiences. It is hypothesized that individuals are characterized by an affec-

tive home base, a baseline attractor state around which affect fluctuates. These fluctuations vary as the result of internal or external processes to which an individual is more or less sensitive and are regulated and tied back to the home base by the attractor strength. Individual differences in these 3 processes—affective home base, variability, and attractor strength—are proposed to underlie individual differences in affect dynamics. The DynAffect account is empirically evaluated by means of a diffusion modeling approach in 2 extensive experience-sampling studies on people's core affective experiences. The findings show that the model is capable of adequately capturing the observed dynamics in core affect across both large (Study 1) and shorter time scales (Study 2) and illuminate how the key processes are related to personality and emotion dispositions. Implications for the understanding of affect dynamics and affective dysfunctioning in psychopathology are also discussed.

Kurtzer et al.: hpcng/singularity: Singularity 3.7.3   Kurtzer-cclerget-Bauer-etal-2021

Gregory M. Kurtzer, cclerget, et al. *hpcng/singularity: Singularity 3.7.3*. 2021. DOI: `10.5281/ZENODO.1310023`.

Kurtzer et al.: Singularity: Scientific containers for mobility of compute
Kurtzer-Sochat-Bauer-2017

Gregory M. Kurtzer, Vanessa Sochat, and Michael W. Bauer. "Singularity: Scientific containers for mobility of compute". In: *PLOS ONE* 12.5 (May 2017), e0177459. DOI: `10.1371/journal.pone.0177459`.

Abstract: Here we present Singularity, software developed to bring containers and reproducibility to scientific computing. Using Singularity containers, developers can work in reproducible environments of their choosing and design, and these complete environments can easily be copied and executed on other platforms. Singularity is an open source initiative that harnesses the expertise of system and software engineers and researchers alike, and integrates seamlessly into common work-

flows for both of these groups. As its primary use case, Singularity brings mobility of computing to both users and HPC centers, providing a secure means to capture and distribute software and compute environments. This ability to create and deploy reproducible environments across these centers, a previously unmet need, makes Singularity a game changing development for computational science.

Kwan et al.: Comparing standardized coefficients in structural equation modeling: A model reparameterization approach   Kwan-Chan-2011

Abstract: We propose a two-stage method for comparing standardized coefficients in structural equation modeling (SEM). At stage 1, we transform the original model of interest into the standardized model by model reparameterization, so that the model parameters appearing in the standardized model are equivalent to the standardized parameters of the original model. At stage 2, we impose appropriate linear equality constraints on the standardized model and use a likelihood ratio test to make statistical inferences about the equality of standardized coefficients. Unlike other existing methods for comparing standardized coefficients, the proposed method does not require specific modeling features (e.g., specification of nonlinear constraints), which are available only in certain SEM software programs. Moreover, this method allows researchers to compare two or more standardized coefficients simultaneously in a standard and convenient way. Three real examples are given to illustrate the proposed method, using EQS, a popular SEM software program. Results show that the proposed method performs satisfactorily for testing the equality of standardized coefficients.

Abstract: In social science research, a common topic in multiple regression analysis is to compare the squared multiple correlation coefficients in different populations. Existing methods based on asymptotic theories (Olkin & Finn, 1995) and bootstrapping (Chan, 2009) are available but these can only handle a 2-group comparison. Another method based on structural equation modeling (SEM) has been proposed recently. However, this method has three disadvantages. First, it requires the user to explicitly specify the sample R2 as a function in terms of the basic SEM model parameters, which is sometimes troublesome and error prone. Second, it requires the specification of nonlinear constraints, which is not available in some popular SEM software programs. Third, it is for a 2-group comparison primarily. In this article, a 2-stage SEM method is proposed as an alternative. Unlike all other existing methods, the proposed method is simple to use, and it does not require any specific programming features such as the specification of nonlinear constraints. More important, the method allows a simultaneous comparison of 3 or more groups. A real example is given to illustrate the proposed method using EQS, a popular SEM software program.

Lachowicz et al.: A novel measure of effect size for mediation analysis

Abstract: Mediation analysis has become one of the most popular statistical methods in the social sciences. However, many currently available effect size measures for mediation have limitations that

restrict their use to specific mediation models. In this article, we develop a measure of effect size that addresses these limitations. We show how modification of a currently existing effect size measure results in a novel effect size measure with many desirable properties. We also derive an expression for the bias of the sample estimator for the proposed effect size measure and propose an adjusted version of the estimator. We present a Monte Carlo simulation study conducted to examine the finite sampling properties of the adjusted and unadjusted estimators, which shows that the adjusted estimator is effective at recovering the true value it estimates. Finally, we demonstrate the use of the effect size measure with an empirical example. We provide freely available software so that researchers can immediately implement the methods we discuss. Our developments here extend the existing literature on effect sizes and mediation by developing a potentially useful method of communicating the magnitude of mediation.

P. Lahiri: On the impact of bootstrap in survey sampling and small-area estimation

Lahiri-2003

Abstract: Development of valid bootstrap procedures has been a challenging problem for survey samplers for the last two decades. This is due to the fact that in surveys we constantly face various complex issues such as complex correlation structure induced by the survey design, weighting, imputation, small-area estimation, among others. In this paper, we critically review various bootstrap methods developed to deal with these challenging issues. We discuss two applications where the bootstrap has been found to be effective.

S. A. W. Lee et al.: From the individual to the group: Using idiographic analyses and two-stage random effects meta-analysis to obtain population level inferences for within-person processes **Lee-Gates-2023**

Abstract: In psychology, the use of portable technology and wearable devices to ease participant burden in data collection is on the rise. This creates increased interest in collecting real-time or near real-time data from individuals within their natural environments. As a result, vast amounts of observational time series data are generated. Often, motivation for collecting this data hinges on understanding within-person processes that underlie psychological phenomena. Motivated by the body of Dr. Peter Molenaar's life work calling for analytical approaches that consider potential heterogeneity and non-ergodicity, the focus of this paper is on using idiographic analyses to generate population inferences for within-person processes. Meta-analysis techniques using one-stage and two-stage random effects meta-analysis as implemented in single-case experimental designs are presented. The case for preferring a two-stage approach for meta-analysis of single-subject observational time series data is made and demonstrated using an empirical example. This provides a novel implementation of the methodology as prior implementations focus on applications to short time series with experimental designs. Inspired by Dr. Molenaar's work, we describe how an approach, two-stage random effects meta-analysis (2SRE-MA), aligns with recent calls to consider idiographic approaches when making population-level inferences regarding within-person processes.

Leffingwell et al.: Continuous Objective Monitoring of Alcohol Use: Twenty-First Century Measurement Using Transdermal Sensors **Leffingwell-Cooney-Murphy-etal-2012**

Thad R. Leffingwell et al. "Continuous Objective Monitoring of Alcohol Use: Twenty-First Century

Measurement Using Transdermal Sensors". In: *Alcoholism: Clinical and Experimental Research* 37.1 (July 2012), pp. 16–22. ISSN: 1530-0277. DOI: 10.1111/j.1530-0277.2012.01869.x.

Abstract: Transdermal alcohol sensors continuously collect reliable and valid data on alcohol consumption in vivo over the course of hours to weeks. Transdermal alcohol readings are highly correlated with breath alcohol measurements, but transdermal alcohol levels lag behind breath alcohol levels by one or more hours owing to the longer time required for alcohol to be expelled through perspiration. By providing objective information about alcohol consumption, transdermal alcohol sensors can validate self-report and provide important information not previously available. In this article, we describe the development and evaluation of currently available transdermal alcohol sensors, present the strengths and limitations of the technology, and give examples of recent research using the sensors.

Lele: Impact of bootstrap on the estimating functions   Lele-2003

Subhash R. Lele. "Impact of bootstrap on the estimating functions". In: *Statistical Science* 18.2 (May 2003). DOI: 10.1214/ss/1063994973.

Abstract: Estimating functions form an attractive statistical methodology because of their dependence on only a few features of the underlying probabilistic structure. They also put a premium on developing methods that obtain model-robust confidence intervals. Bootstrap and jackknife ideas can be fruitfully used toward this purpose. Another important area in which bootstrap has proved its use is in the context of detecting the problem of multiple roots and searching for the consistent root of an estimating function. In this article, I review, compare and contrast various approaches for bootstrapping estimating functions.

Leventhal et al.: The neighborhoods they live in: The effects of neighborhood residence on child and adolescent outcomes.   Leventhal-BrooksGunn-2000

Tama Leventhal and Jeanne Brooks-Gunn. "The neighborhoods they live in: The effects of neighbor-

hood residence on child and adolescent outcomes." In: *Psychological Bulletin* 126.2 (2000), pp. 309–337. ISSN: 0033-2909. DOI: 10.1037/0033-2909.126.2.309.

Abstract: Provides a comprehensive review of research on the effects of neighborhood residence on child and adolescent well-being. The 1st section reviews key methodological issues. The following section considers links between neighborhood characteristics and child outcomes and suggests the importance of high SES for achievement and low SES and residential instability for behavioral/emotional outcomes. The third section identifies 3 pathways (institutional resources, relationships, and norms/collective efficacy) through which neighborhoods might influence development, and which represent an extension of models identified by C. Jencks and S. Mayer (1990) and R. J. Sampson (1992). The models provide a theoretical base for studying neighborhood mechanisms and specify different levels (individual, family, school, peer, community) at which processes may operate. Implications for an emerging developmental framework for research on neighborhoods are discussed.

Levitt et al.: Journal article reporting standards for qualitative primary, qualitative meta-analytic, and mixed methods research in psychology: The APA Publications and Communications Board task force report   Levitt-Bamberg-Creswell-etal-2018

Heidi M. Levitt et al. "Journal article reporting standards for qualitative primary, qualitative meta-analytic, and mixed methods research in psychology: The APA Publications and Communications Board task force report". In: *American Psychologist* 73.1 (Jan. 2018), pp. 26–46. ISSN: 0003-066X. DOI: 10.1037/amp0000151.

Abstract: The American Psychological Association Publications and Communications Board Working Group on Journal Article Reporting Standards for Qualitative Research (JARS-Qual Working Group) was charged with examining the state of journal article reporting standards as they applied to qualitative research and with generating recommendations for standards that would be appropriate for a wide range of methods within the discipline of psychology. These standards describe what

should be included in a research report to enable and facilitate the review process. This publication marks a historical moment—the first inclusion of qualitative research in APA Style, which is the basis of both the Publication Manual of the American Psychological Association (APA, 2010) and APA Style CENTRAL, an online program to support APA Style. In addition to the general JARS-Qual guidelines, the Working Group has developed standards for both qualitative meta-analysis and mixed methods research. The reporting standards were developed for psychological qualitative research but may hold utility for a broad range of social sciences. They honor a range of qualitative traditions, methods, and reporting styles. The Working Group was composed of a group of researchers with backgrounds in varying methods, research topics, and approaches to inquiry. In this article, they present these standards and their rationale, and they detail the ways that the standards differ from the quantitative research reporting standards. They describe how the standards can be used by authors in the process of writing qualitative research for submission as well as by reviewers and editors in the process of reviewing research.

K. H. Li et al.: Large-sample significance levels from multiply imputed data using moment-based statistics and an F reference distribution

Li-Raghunathan-Rubin-1991

Abstract: We present a procedure for computing significance levels from data sets whose missing values have been multiply imputed data. This procedure uses moment-based statistics, $m \leq 3$ repeated imputations, and an F reference distribution. When $m = \infty$, we show first that our procedure is essentially the same as the ideal procedure in cases of practical importance and, second, that its deviations from the ideal are basically a function of the coefficient of variation of the canonical ratios of complete to observed information. For small m our procedure's performance

is largely governed by this coefficient of variation and the mean of these ratios. Using simulation techniques with small m, we compare our procedure's actual and nominal large-sample significance levels and conclude that it is essentially calibrated and thus represents a definite improvement over previously available procedures. Furthermore, we compare the large-sample power of the procedure as a function of m and other factors, such as the dimensionality of the estimand and fraction of missing information, to provide guidance on the choice of the number of imputations; generally, we find the loss of power due to small m to be quite modest in cases likely to occur in practice.

Y. Li et al.: Bayesian forecasting with a regime-switching zero-inflated multilevel poisson regression model: An application to adolescent alcohol use with spatial covariates
Li-Oravecz-Zhou-etal-2022

Abstract: In this paper, we present and evaluate a novel Bayesian regime-switching zero-inflated multilevel Poisson (RS-ZIMLP) regression model for forecasting alcohol use dynamics. The model partitions individuals' data into two phases, known as regimes, with: (1) a zero-inflation regime that is used to accommodate high instances of zeros (non-drinking) and (2) a multilevel Poisson regression regime in which variations in individuals' log-transformed average rates of alcohol use are captured by means of an autoregressive process with exogenous predictors and a person-specific intercept. The times at which individuals are in each regime are unknown, but may be estimated from the data. We assume that the regime indicator follows a first-order Markov process as related to exogenous predictors of interest. The forecast performance of the proposed model was evaluated using a Monte Carlo simulation study and further demonstrated using substance use and spatial covariate data from the Colorado Online Twin Study (CoTwins). Results showed that the proposed model yielded better forecast performance compared to a baseline model which predicted all cases as non-drinking and a reduced ZIMLP model without the RS structure, as indicated by higher AUC

(the area under the receiver operating characteristic (ROC) curve) scores, and lower mean absolute errors (MAEs) and root-mean-square errors (RMSEs). The improvements in forecast performance were even more pronounced when we limited the comparisons to participants who showed at least one instance of transition to drinking.

Y. Li et al.: Fitting multilevel vector autoregressive models in Stan, JAGS, and Mplus
Li-Wood-Ji-etal-2021

Abstract: The influx of intensive longitudinal data creates a pressing need for complex modeling tools that help enrich our understanding of how individuals change over time. Multilevel vector autoregressive (mlVAR) models allow for simultaneous evaluations of reciprocal linkages between dynamic processes and individual differences, and have gained increased recognition in recent years. High-dimensional and other complex variations of mlVAR models, though often computationally intractable in the frequentist framework, can be readily handled using Markov chain Monte Carlo techniques in a Bayesian framework. However, researchers in social science fields may be unfamiliar with ways to capitalize on recent developments in Bayesian software programs. In this paper, we provide step-by-step illustrations and comparisons of options to fit Bayesian mlVAR models using Stan, JAGS and Mplus, supplemented with a Monte Carlo simulation study. An empirical example is used to demonstrate the utility of mlVAR models in studying intra- and inter-individual variations in affective dynamics.

Little et al.: Statistical analysis with missing data   Little-Rubin-2019

Abstract: An up-to-date, comprehensive treatment of a classic text on missing data in statistics. The topic of missing data has gained considerable attention in recent decades. This new edition by two acknowledged experts on the subject offers an up-to-date account of practical methodology for handling missing data problems. Blending theory and application, authors Roderick Little and Donald Rubin review historical approaches to the subject and describe simple methods for multivariate analysis with missing values. They then provide a coherent theory for analysis of problems based on likelihoods derived from statistical models for the data and the missing data mechanism, and then they apply the theory to a wide range of important missing data problems. Statistical Analysis with Missing Data, Third Edition starts by introducing readers to the subject and approaches toward solving it. It looks at the patterns and mechanisms that create the missing data, as well as a taxonomy of missing data. It then goes on to examine missing data in experiments, before discussing complete-case and available-case analysis, including weighting methods. The new edition expands its coverage to include recent work on topics such as nonresponse in sample surveys, causal inference, diagnostic methods, and sensitivity analysis, among a host of other topics.

- An updated "classic" written by renowned authorities on the subject

- Features over 150 exercises (including many new ones)

- Covers recent work on important methods like multiple imputation, robust alternatives to weighting, and Bayesian methods

- Revises previous topics based on past student feedback and class experience

- Contains an updated and expanded bibliography

The authors were awarded The Karl Pearson Prize in 2017 by the International Statistical Institute, for a research contribution that has had profound influence on statistical theory, methodology or applications. Their work "has been no less than defining and transforming." (ISI) Statistical Analysis with Missing Data, Third Edition is an ideal textbook for upper undergraduate and/or beginning graduate level students of the subject. It is also an excellent source of information for applied statisticians and practitioners in government and industry.

Library: QA276.

159

Lohmann et al.: A primer on continuous-time modeling in educational research: An exemplary application of a continuous-time latent curve model with structured residuals (CT-LCM-SR) to PISA data   **Lohmann-Zitzmann-Voelkle-etal-2022**

Abstract: One major challenge of longitudinal data analysis is to find an appropriate statistical model that corresponds to the theory of change and the research questions at hand. In the present article, we argue that continuous-time models are well suited to study the continuously developing constructs of primary interest in the education sciences and outline key advantages of using this type of model. Furthermore, we propose the continuous-time latent curve model with structured residuals (CT-LCM-SR) as a suitable model for many research questions in the education sciences. The CT-LCM-SR combines growth and dynamic modeling and thus provides descriptions of both trends and process dynamics. We illustrate the application of the CT-LCM-SR with data from PISA reading literacy assessment of 2000 to 2018 and provide a tutorial and annotated code for setting up the CT-LCM-SR model.

Long et al.: Using heteroscedasticity consistent standard errors in the linear regression model   **Long-Ervin-2000**

Abstract: The human affect system is responsible for producing the positive and negative feelings that color and guide our lives. At the same time, when disrupted, its workings lie at the basis of the occurrence of mood disorder. Understanding the functioning and dynamics of the affect system is therefore crucial to understand the feelings that people experience on a daily basis, their dynamics across time, and how they can become dysregulated in mood disorder. In this paper, a nonlinear stochastic model for the dynamics of positive and negative affect is proposed called the Affective Ising Model (AIM). It incorporates principles of statistical mechanics, is inspired by neurophysiological and behavioral evidence about auto-excitation and mutual inhibition of the positive and negative affect dimensions, and is intended to better explain empirical phenomena such as skewness, multimodality, and non-linear relations of positive and negative affect. The AIM is applied to two large experience sampling studies on the occurrence of positive and negative affect in daily life in both normality and mood disorder. It is examined to what extent the model is able to reproduce the aforementioned non-Gaussian features observed in the data, using two sightly different continuous-time vector autoregressive (VAR) models as benchmarks. The predictive performance of the models is also compared by means of leave-one-out cross-validation. The results indicate that the AIM is better at reproducing non-Gaussian features while their performance is comparable for strictly Gaussian features. The predictive performance of the AIM is also shown to be better for the majority of the affect time series. The potential and limitations of the AIM as a computational model approximating the workings of the human affect system are discussed.

Loossens et al.: A comparison of continuous and discrete time modeling of affective processes in terms of predictive accuracy   Loossens-Tuerlinckx-Verdonck-2021

Abstract: Intra-individual processes are thought to continuously unfold across time. For equally spaced time intervals, the discrete-time lag-1 vector autoregressive (VAR(1)) model and the continuous-time Ornstein-Uhlenbeck (OU) model are equivalent. It is expected that by taking into account the unequal spacings of the time intervals in real data between observations will lead to an advantage for the OU in terms of predictive accuracy. In this paper, this is claim is being investigated by comparing the predictive accuracy of the OU model to that of the VAR(1) model on typical ESM data obtained in the context of affect research. It is shown that the VAR(1) model outperforms the OU model for the majority of the time series, even though time intervals in the data are unequally spaced. Accounting for measurement error does not change the result. Deleting large abrupt changes on short time intervals (that may be caused by externally driven events) does however lead to a significant improvement for the OU model. This suggests that processes in psychology may be continuously evolving, but that there are factors, like external events, which can disrupt the continuous flow.

Lüdtke et al.: The multilevel latent covariate model: A new, more reliable approach to group-level effects in contextual studies   Ludtke-Marsh-Robitzsch-etal-2008

Abstract: In multilevel modeling (MLM), group-level (L2) characteristics are often measured by aggregating individual-level (L1) characteristics within each group so as to assess contextual effects

(e.g., group-average effects of socioeconomic status, achievement, climate). Most previous applications have used a multilevel manifest covariate (MMC) approach, in which the observed (manifest) group mean is assumed to be perfectly reliable. This article demonstrates mathematically and with simulation results that this MMC approach can result in substantially biased estimates of contextual effects and can substantially underestimate the associated standard errors, depending on the number of L1 individuals per group, the number of groups, the intraclass correlation, the sampling ratio (the percentage of cases within each group sampled), and the nature of the data. To address this pervasive problem, the authors introduce a new multilevel latent covariate (MLC) approach that corrects for unreliability at L2 and results in unbiased estimates of L2 constructs under appropriate conditions. However, under some circumstances when the sampling ratio approaches 100%, the MMC approach provides more accurate estimates. Based on 3 simulations and 2 real-data applications, the authors evaluate the MMC and MLC approaches and suggest when researchers should most appropriately use one, the other, or a combination of both approaches.

Lütkepohl: New introduction to multiple time series analysis   Lutkepohl-2005

Abstract: This reference work and graduate level textbook considers a wide range of models and methods for analyzing and forecasting multiple time series. The models covered include vector autoregressive, cointegrated, vector autoregressive moving average, multivariate ARCH and periodic processes as well as dynamic simultaneous equations and state space models. Least squares, maximum likelihood and Bayesian methods are considered for estimating these models. Different procedures for model selection and model specification are treated and a wide range of tests and criteria for model checking are introduced. Causality analysis, impulse response analysis and innovation accounting are presented as tools for structural analysis. The book is accessible to graduate students in business and economics. In addition, multiple time series courses in other fields such as statistics and engineering may be based on it. Applied researchers involved in analyzing multiple

time series may benefit from the book as it provides the background and tools for their tasks. It bridges the gap to the difficult technical literature on the topic.

Lyapunov: The general problem of the stability of motion   Lyapunov-1992

Alexandr Mikhailovich Lyapunov. "The general problem of the stability of motion". In: *International Journal of Control* 55.3 (Mar. 1992), pp. 531–534. ISSN: 1366-5820. DOI: 10 . 1080 / 00207179208934253.

D. P. MacKinnon: Analysis of mediating variables in prevention and intervention research   MacKinnon-1994

David P. MacKinnon. "Analysis of mediating variables in prevention and intervention research". In: *NIDA research monograph* 139 (1994), pp. 127–153.

Abstract: Mediational analysis is one way to test specific hypotheses derived from theory. Although this analysis has been suggested in the prevention literature, mediation analysis rarely is conducted. As the field of prevention matures, more questions about how prevention programs work (or fail to work) will emerge. Studies of mediation can address these questions, thereby reducing the cost and enhancing the impact of prevention programs. The methods outlined here can be applied in the evaluation of primary, secondary, and tertiary prevention programs. Since most prevention studies include measurement of some mediating constructs, mediation effects can be assessed on many existing data sets. Mediation analysis can be used to test ideas about prevention.

D. P. MacKinnon: Introduction to statistical mediation analysis   MacKinnon-2008

David P. MacKinnon. *Introduction to statistical mediation analysis.* Multivariate applications. Hoboken: Erlbaum Psych Press, 2008, p. 488. ISBN: 9780805864298. DOI: 10.4324/9780203809556.

Abstract: This volume introduces the statistical, methodological, and conceptual aspects of media-

tion analysis. Applications from health, social, and developmental psychology, sociology, communication, exercise science, and epidemiology are emphasized throughout. Single-mediator, multilevel, and longitudinal models are reviewed. The author's goal is to help the reader apply mediation analysis to their own data and understand its limitations. Each chapter features an overview, numerous worked examples, a summary, and exercises (with answers to the odd numbered questions). The accompanying downloadable resources contain outputs described in the book from SAS, SPSS, LISREL, EQS, MPLUS, and CALIS, and a program to simulate the model. The notation used is consistent with existing literature on mediation in psychology. The book opens with a review of the types of research questions the mediation model addresses. Part II describes the estimation of mediation effects including assumptions, statistical tests, and the construction of confidence limits. Advanced models including mediation in path analysis, longitudinal models, multilevel data, categorical variables, and mediation in the context of moderation are then described. The book closes with a discussion of the limits of mediation analysis, additional approaches to identifying mediating variables, and future directions. Introduction to Statistical Mediation Analysis is intended for researchers and advanced students in health, social, clinical, and developmental psychology as well as communication, public health, nursing, epidemiology, and sociology. Some exposure to a graduate level research methods or statistics course is assumed. The overview of mediation analysis and the guidelines for conducting a mediation analysis will be appreciated by all readers.

Mackinnon et al.: Estimating mediated effects in prevention studies

Mackinnon-Dwyer-1993

Abstract: The purpose of this article is to describe statistical procedures to assess how prevention and intervention programs achieve their effects. The analyses require the measurement of intervening

or mediating variables hypothesized to represent the causal mechanism by which the prevention program achieves its effects. Methods to estimate mediation are illustrated in the evaluation of a health promotion program designed to reduce dietary cholesterol and a school-based drug prevention program. The methods are relatively easy to apply and the information gained from such analyses should add to our understanding of prevention.

D. P. MacKinnon et al.: Distribution of the product confidence limits for the indirect effect: Program PRODCLIN — MacKinnon-Fritz-Williams-etal-2007

Abstract: This article describes a program, PRODCLIN (distribution of the PRODuct Confidence Limits for INdirect effects), written for SAS, SPSS, and R, that computes confidence limits for the product of two normal random variables. The program is important because it can be used to obtain more accurate confidence limits for the indirect effect, as demonstrated in several recent articles (MacKinnon, Lockwood, & Williams, 2004; Pituch, Whittaker, & Stapleton, 2005). Tests of the significance of and confidence limits for indirect effects based on the distribution of the product method have more accurate Type I error rates and more power than other, more commonly used tests. Values for the two paths involved in the indirect effect and their standard errors are entered in the PRODCLIN program, and distribution of the product confidence limits are computed. Several examples are used to illustrate the PRODCLIN program. The PRODCLIN programs in rich text format may be downloaded from www.psychonomic.org/archive.

D. P. MacKinnon et al.: Equivalence of the mediation, confounding and suppression effect — MacKinnon-Krull-Lockwood-2000

David P. MacKinnon, Jennifer L. Krull, and Chondra M. Lockwood. "Equivalence of the mediation,

confounding and suppression effect". In: *Prevention Science* 1.4 (2000), pp. 173–181. ISSN: 1389-4986. DOI: 10.1023/a:1026595011371.

Abstract: This paper describes the statistical similarities among mediation, confounding, and suppression. Each is quantified by measuring the change in the relationship between an independent and a dependent variable after adding a third variable to the analysis. Mediation and confounding are identical statistically and can be distinguished only on conceptual grounds. Methods to determine the confidence intervals for confounding and suppression effects are proposed based on methods developed for mediated effects. Although the statistical estimation of effects and standard errors is the same, there are important conceptual differences among the three types of effects.

D. P. MacKinnon et al.: A comparison of methods to test mediation and other intervening variable effects   MacKinnon-Lockwood-Hoffman-etal-2002

David P. MacKinnon, Chondra M. Lockwood, Jeanne M. Hoffman, et al. "A comparison of methods to test mediation and other intervening variable effects". In: *Psychological Methods* 7.1 (2002), pp. 83–104. DOI: 10.1037/1082-989x.7.1.83.

Abstract: A Monte Carlo study compared 14 methods to test the statistical significance of the intervening variable effect. An intervening variable (mediator) transmits the effect of an independent variable to a dependent variable. The commonly used R. M. Baron and D. A. Kenny (1986) approach has low statistical power. Two methods based on the distribution of the product and 2 difference-in-coefficients methods have the most accurate Type I error rates and greatest statistical power except in 1 important case in which Type I error rates are too high. The best balance of Type I error and statistical power across all cases is the test of the joint significance of the two effects comprising the intervening variable effect.

David P. MacKinnon, Chondra M. Lockwood, and Jason Williams. "Confidence limits for the indirect effect: Distribution of the product and resampling methods". In: *Multivariate Behavioral Research* 39.1 (Jan. 2004), pp. 99–128. DOI: 10.1207/s15327906mbr3901_4.

Abstract: The most commonly used method to test an indirect effect is to divide the estimate of the indirect effect by its standard error and compare the resulting z statistic with a critical value from the standard normal distribution. Confidence limits for the indirect effect are also typically based on critical values from the standard normal distribution. This article uses a simulation study to demonstrate that confidence limits are imbalanced because the distribution of the indirect effect is normal only in special cases. Two alternatives for improving the performance of confidence limits for the indirect effect are evaluated: (a) a method based on the distribution of the product of two normal random variables, and (b) resampling methods. In Study 1, confidence limits based on the distribution of the product are more accurate than methods based on an assumed normal distribution but confidence limits are still imbalanced. Study 2 demonstrates that more accurate confidence limits are obtained using resampling methods, with the bias-corrected bootstrap the best method overall.

James G. MacKinnon and Halbert White. "Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties". In: *Journal of Econometrics* 29.3 (Sept. 1985), pp. 305–325. DOI: 10.1016/0304-4076(85)90158-7.

Abstract: We examine several modified versions of the heteroskedasticity-consistent covariance matrix estimator of Hinkley (1977) and White (1980). On the basis of sampling experiments which compare the performance of quasi t-statistics, we find that one estimator, based on the jackknife,

performs better in small samples than the rest. We also examine the finite-sample properties of using modified critical values based on Edgeworth approximations, as proposed by Rothenberg (1984). In addition, we compare the power of several tests for heteroskedasticity, and find that it may be wise to employ the jackknife heteroskedasticity-consistent covariance matrix even in the absence of detected heteroskedasticity.

Maggs et al.: Initiation and course of alcohol consumption among adolescents and young adults   Maggs-Schulenberg-2005

Abstract: This chapter takes a normative developmental perspective on the etiology of alcohol use, focusing on the initiation and course of alcohol use (rather than alcohol use disorders) during adolescence and early adulthood. We review evidence regarding the sequelae and meaning of the age of initiation of alcohol use, consider variable- and pattern-centered approaches to modeling trajectories describing the course of alcohol use across adolescence and young adulthood, and offer developmental conceptualizations of risk and protective factors for alcohol use and related problems.

Manthey et al.: What are the economic costs to society attributable to alcohol use? A systematic review and modelling study   Manthey-Hassan-Carr-etal-2021

Abstract: Background: Alcohol-attributable costs to society are captured by cost-of-illness studies, however estimates are often not comparable, e.g. due to the omission of relevant cost components.

In this contribution we (1) summarize the societal costs attributable to alcohol use, and (2) estimate the total costs under the assumption that all cost components are considered. Methods: A systematic review and meta-analyses were conducted for studies reporting costs from alcohol consumption for the years 2000 and later, using the EMBASE and MEDLINE databases. Cost estimates were converted into 2019 international dollars (Int$) per adult and into percentage of gross domestic product (GDP). For each study, weights were calculated to correct for the exclusion of cost indicators. Results: Of 1708 studies identified, 29 were included, and the mean costs of alcohol use amounted to 817.6 Int$ per adult (95% confidence interval [CI] 601.8-1033.4), equivalent to 1.5% of the GDP (95% CI 1.2-1.7%). Adjusting for omission of cost components, the economic costs of alcohol consumption were estimated to amount to 1306 Int$ per adult (95% CI 873-1738), or 2.6% (95% CI 2.0-3.1%) of the GDP. About one-third of costs (38.8%) were incurred through direct costs, while the majority of costs were due to losses in productivity (61.2%). Discussion: The identified cost studies were mainly conducted in high-income settings, with high heterogeneity in the employed methodology. Accounting for some methodological variations, our findings demonstrate that alcohol use continues to incur a high level of cost to many societies.

Maxwell et al.: Bias in cross-sectional analyses of longitudinal mediation

Maxwell-Cole-2007

Abstract: Most empirical tests of mediation utilize cross-sectional data despite the fact that mediation consists of causal processes that unfold over time. The authors considered the possibility that longitudinal mediation might occur under either of two different models of change: (a) an autoregressive model or (b) a random effects model. For both models, the authors demonstrated that cross-sectional approaches to mediation typically generate substantially biased estimates of longitudinal parameters even under the ideal conditions when mediation is complete. In longitudinal models where variable M completely mediates the effect of X on Y, cross-sectional estimates of

the direct effect of X on Y, the indirect effect of X on Y through M, and the proportion of the total effect mediated by M are often highly misleading.

Maxwell et al.: Bias in cross-sectional analyses of longitudinal mediation: Partial and complete mediation under an autoregressive model   Maxwell-Cole-Mitchell-2011

Abstract: Maxwell and Cole (2007) showed that cross-sectional approaches to mediation typically generate substantially biased estimates of longitudinal parameters in the special case of complete mediation. However, their results did not apply to the more typical case of partial mediation. We extend their previous work by showing that substantial bias can also occur with partial mediation. In particular, cross-sectional analyses can imply the existence of a substantial indirect effect even when the true longitudinal indirect effect is zero. Thus, a variable that is found to be a strong mediator in a cross-sectional analysis may not be a mediator at all in a longitudinal analysis. In addition, we show that very different combinations of longitudinal parameter values can lead to essentially identical cross-sectional correlations, raising serious questions about the interpretability of cross-sectional mediation data. More generally, researchers are encouraged to consider a wide variety of possible mediation models beyond simple cross-sectional models, including but not restricted to autoregressive models of change.

J. Jack McArdle et al.: Some algebraic properties of the Reticular Action Model for moment structures   McArdle-McDonald-1984

Abstract: A number of models for the analysis of moment structures, such as linear structural relations, have recently been shown to be capable of being given a particularly simple and economical representation, in terms of the reticular action model (RAM). A formal algebraic treatment is presented that shows that RAM directly incorporates many common structural models, including models describing the structure of means. It is also shown that RAM treats coefficient matrices with patterned inverses simply and generally.

John J. McArdle: Latent variable modeling of differences and changes with longitudinal data   McArdle-2009

Abstract: This review considers a common question in data analysis: What is the most useful way to analyze longitudinal repeated measures data? We discuss some contemporary forms of structural equation models (SEMs) based on the inclusion of latent variables. The specific goals of this review are to clarify basic SEM definitions, consider relations to classical models, focus on testable features of the new models, and provide recent references to more complete presentations. A broader goal is to illustrate why so many researchers are enthusiastic about the SEM approach to data analysis. We first outline some classic problems in longitudinal data analysis, consider definitions of differences and changes, and raise issues about measurement errors. We then present several classic SEMs based on the inclusion of invariant common factors and explain why these are so important. This leads to newer SEMs based on latent change scores, and we explain why these are useful.

McKendrick et al.: Drug-induced conditioned place preference and its practical use in substance use disorder research   McKendrick-Graziane-2020

Greer McKendrick and Nicholas M. Graziane. "Drug-induced conditioned place preference and its

practical use in substance use disorder research". In: *Frontiers in Behavioral Neuroscience* 14 (Sept. 2020). ISSN: 1662-5153. DOI: 10.3389/fnbeh.2020.582147.

Abstract: The conditioned place preference (CPP) paradigm is a well-established model utilized to study the role of context associations in reward-related behaviors, including both natural rewards and drugs of abuse. In this review article, we discuss the basic history, various uses, and considerations that are tied to this technique. There are many potential takeaway implications of this model, including negative affective states, conditioned drug effects, memory, and motivation, which are all considered here. We also discuss the neurobiology of CPP including relevant brain regions, molecular signaling cascades, and neuromodulatory systems. We further examine some of our prior findings and how they integrate CPP with self-administration paradigms. Overall, by describing the fundamentals of CPP, findings from the past few decades, and implications of using CPP as a research paradigm, we have endeavored to support the case that the CPP method is specifically advantageous for studying the role of a form of Pavlovian learning that associates drug use with the surrounding environment.

McNeish et al.: A primer on two-level dynamic structural equation models for intensive longitudinal data in Mplus   McNeish-Hamaker-2020

Daniel McNeish and Ellen L. Hamaker. "A primer on two-level dynamic structural equation models for intensive longitudinal data in Mplus". In: *Psychological Methods* 25.5 (Oct. 2020), pp. 610–635. DOI: [10.1037/met0000250](10.1037/met0000250).

Abstract: Technological advances have led to an increase in intensive longitudinal data and the statistical literature on modeling such data is rapidly expanding, as are software capabilities. Common methods in this area are related to time-series analysis, a framework that historically has received little exposure in psychology. There is a scarcity of psychology-based resources introducing the basic ideas of time-series analysis, especially for data sets featuring multiple people. We begin with basics of N = 1 time-series analysis and build up to complex dynamic structural equation

models available in the newest release of Mplus Version 8. The goal is to provide readers with a basic conceptual understanding of common models, template code, and result interpretation. We provide short descriptions of some advanced issues, but our main priority is to supply readers with a solid knowledge base so that the more advanced literature on the topic is more readily digestible to a larger group of researchers.

McNeish et al.: Intensive longitudinal mediation in Mplus   McNeish-MacKinnon-2025

Abstract: Much of the existing longitudinal mediation literature focuses on panel data where relatively few repeated measures are collected over a relatively broad timespan. However, technological advances in data collection (e.g., smartphones, wearables) have led to a proliferation of short duration, densely collected longitudinal data in behavioral research. These intensive longitudinal data differ in structure and focus relative to traditionally collected panel data. As a result, existing methodological resources do not necessarily extend to nuances present in the recent influx of intensive longitudinal data and designs. In this tutorial, we first cover potential limitations of traditional longitudinal mediation models to accommodate unique characteristics of intensive longitudinal data. Then, we discuss how recently developed dynamic structural equation models (DSEMs) may be well-suited for mediation modeling with intensive longitudinal data and can overcome some of the limitations associated with traditional approaches. We describe four increasingly complex intensive longitudinal mediation models: (a) stationary models where the indirect effect is constant over time and people, (b) person-specific models where the indirect effect varies across people, (c) dynamic models where the indirect effect varies across time, and (d) cross-classified models where the indirect effect varies across both time and people. We apply each model to a running example featuring a mobile health intervention designed to improve health behavior of individuals with binge eating disorder. In each example, we provide annotated Mplus code and interpretation of the output

174

to guide empirical researchers through mediation modeling with this increasingly popular type of longitudinal data.

Mehl et al.: Handbook of research methods for studying daily life

Matthias R. Mehl, Tamlin S. Conner, and Mihaly. Csikszentmihalyi. *Handbook of research methods for studying daily life*. Guilford Publications, 2011. ISBN: 9781609187491.

Abstract: Laboratory-based experimental methods historically have been the strength and pride of psychology and related disciplines. Yet a comprehensive science of behavior also requires the study of humans in real life. Bringing together leading investigators, this book reviews the breadth of current approaches for studying how people think, feel, and behave in everyday environments. The Handbook is organized in four parts. Part I covers the theoretical and methodological foundations of conducting daily life research. Part II provides guidance for designing a high-quality study and selecting and implementing appropriate methods. The chapters describe experience sampling methods, diary methods, ambulatory physiological measures, and other tools—including recording technologies and computerized approaches—that allow repeated, real-time measurement in natural settings. Part III focuses on techniques for analyzing intensive data from daily life, featuring practical discussions of power analysis, psychometrics, data cleaning, multilevel modeling, time series analysis, and other topics. Part IV reviews how methods for studying daily life have been employed in different subfields and research areas, such as the study of emotion, close relationships, personality, health, development, psychopathology, and mental health treatment. Specific advantages and challenges inherent to using the methods in each area are discussed. Timely and authoritative, this handbook meets a key need for research psychologists and for graduate students in social/personality, health, developmental, industrial/organizational, and clinical psychology.

Mehta et al.: People are variables too: Multilevel structural equations modeling.

Mehta-Neale-2005

Paras D. Mehta and Michael C. Neale. "People are variables too: Multilevel structural equations modeling." In: *Psychological Methods* 10.3 (Sept. 2005), pp. 259–284. ISSN: 1082-989X. DOI: 10.1037/1082-989x.10.3.259.

Abstract: The article uses confirmatory factor analysis (CFA) as a template to explain didactically multilevel structural equation models (ML-SEM) and to demonstrate the equivalence of general mixed-effects models and ML-SEM. An intuitively appealing graphical representation of complex ML-SEMs is introduced that succinctly describes the underlying model and its assumptions. The use of definition variables (i.e., observed variables used to fix model parameters to individual specific data values) is extended to the case of ML-SEMs for clustered data with random slopes. Empirical examples of multilevel CFA and ML-SEM with random slopes are provided along with scripts for fitting such models in SAS Proc Mixed, Mplus, and Mx. Methodological issues regarding estimation of complex ML-SEMs and the evaluation of model fit are discussed. Further potential applications of ML-SEMs are explored.

Merkel: Docker: Lightweight Linux containers for consistent development and deployment

Merkel-2014

Dirk Merkel. "Docker: Lightweight Linux containers for consistent development and deployment". In: *Linux Journal* 2014.239 (2014), p. 2. URL: https://www.linuxjournal.com/content/docker-lightweight-linux-containers-consistent-development-and-deployment.

Micceri: The unicorn, the normal curve, and other improbable creatures

Micceri-1989

Theodore Micceri. "The unicorn, the normal curve, and other improbable creatures". In: *Psychological Bulletin* 105.1 (1989), pp. 156–166. DOI: 10.1037/0033-2909.105.1.156.

Millsap: Statistical approaches to measurement invariance   Millsap-2011

Roger E. Millsap. *Statistical approaches to measurement invariance.* Routledge, 2011. ISBN: 9780203821961. DOI: 10.4324/9780203821961.

Miocevic et al.: A tutorial in Bayesian potential outcomes mediation analysis

Miocevic-Gonzalez-Valente-etal-2017

Milica Miocevic et al. "A tutorial in Bayesian potential outcomes mediation analysis". In: *Structural Equation Modeling: A Multidisciplinary Journal* 25.1 (July 2017), pp. 121–136. ISSN: 1532-8007. DOI: 10.1080/10705511.2017.1342541.

Abstract: Statistical mediation analysis is used to investigate intermediate variables in the relation between independent and dependent variables. Causal interpretation of mediation analyses is challenging because randomization of subjects to levels of the independent variable does not rule out the possibility of unmeasured confounders of the mediator to outcome relation. Furthermore, commonly used frequentist methods for mediation analysis compute the probability of the data given the null hypothesis, which is not the probability of a hypothesis given the data as in Bayesian analysis. Under certain assumptions, applying the potential outcomes framework to mediation analysis allows for the computation of causal effects, and statistical mediation in the Bayesian framework gives indirect effects probabilistic interpretations. This tutorial combines causal inference and Bayesian methods for mediation analysis so the indirect and direct effects have both causal and probabilistic

interpretations. Steps in Bayesian causal mediation analysis are shown in the application to an empirical example.

P. C. M. Molenaar: A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever   Molenaar-2004

Abstract: Psychology is focused on variation between cases (interindividual variation). Results thus obtained are considered to be generalizable to the understanding and explanation of variation within single cases (intraindividual variation). It is indicated, however, that the direct consequences of the classical ergodic theorems for psychology and psychometrics invalidate this conjectured generalizability: only under very strict conditions-which are hardly obtained in real psychological processes-can a generalization be made from a structure of interindividual variation to the analogous structure of intraindividual variation. Illustrations of the lack of this generalizability are given in the contexts of psychometrics, developmental psychology, and personality theory.

P. C. M. Molenaar: Equivalent Dynamic Models   Molenaar-2017

Abstract: Equivalences of two classes of dynamic models for weakly stationary multivariate time series are discussed: dynamic factor models and autoregressive models. It is shown that exploratory dynamic factor models can be rotated, yielding an infinite set of equivalent solutions for any observed series. It also is shown that dynamic factor models with lagged factor loadings are not equivalent to the currently popular state-space models, and that restriction of attention to the latter type of models may yield invalid results. The known equivalent vector autoregressive model types, standard

and structural, are given a new interpretation in which they are conceived of as the extremes of an innovating type of hybrid vector autoregressive models. It is shown that consideration of hybrid models solves many problems, in particular with Granger causality testing.

P. C. M. Molenaar et al.: The new person-specific paradigm in psychology

Molenaar-Campbell-2009

Abstract: Most research methodology in the behavioral sciences employs interindividual analyses, which provide information about the state of affairs of the population. However, as shown by classical mathematical-statistical theorems (the ergodic theorems), such analyses do not provide information for, and cannot be applied at, the level of the individual, except on rare occasions when the processes of interest meet certain stringent conditions. When psychological processes violate these conditions, the interindividual analyses that are now standardly applied have to be replaced by analysis of intraindividual variation in order to obtain valid results. Two illustrations involving analysis of intraindividual variation of personality and emotional processes are given.

Moneta et al.: Causal search in structural vector autoregressive models

Moneta-Chlas-Entner-etal-2011

Abstract: This paper reviews a class of methods to perform causal inference in the framework of a structural vector autoregressive model. We consider three different settings. In the first setting the underlying system is linear with normal disturbances and the structural model is identified by exploiting the information incorporated in the partial correlations of the estimated residuals.

Zero partial correlations are used as input of a search algorithm formalized via graphical causal models. In the second, semi-parametric, setting the underlying system is linear with non-Gaussian disturbances. In this case the structural vector autoregressive model is identified through a search procedure based on independent component analysis. Finally, we explore the possibility of causal search in a nonparametric setting by studying the performance of conditional independence tests based on kernel density estimations.

Morris et al.: The role of the family context in the development of emotion regulation
Morris-Silk-Steinberg-etal-2007

Abstract: This article reviews current literature examining associations between components of the family context and children and adolescents' emotion regulation (ER). The review is organized around a tripartite model of familial influence. Firstly, it is posited that children learn about ER through observational learning, modeling and social referencing. Secondly, parenting practices specifically related to emotion and emotion management affect ER. Thirdly, ER is affected by the emotional climate of the family via parenting style, the attachment relationship, family expressiveness and the marital relationship. The review ends with discussions regarding the ways in which child characteristics such as negative emotionality and gender affect ER, how socialization practices change as children develop into adolescents, and how parent characteristics such as mental health affect the socialization of ER.

Mulder: Power analysis for the random intercept cross-lagged panel model using the powRICLPM R-package
Mulder-2022

Abstract: The random intercept cross-lagged panel model (RI-CLPM) is a popular model among psychologists for studying reciprocal effects in longitudinal panel data. Although various texts and software packages have been published concerning power analyses for structural equation models (SEM) generally, none have proposed a power analysis strategy that is tailored to the particularities of the RI-CLPM. This can be problematic because mismatches between the power analysis design, the model, and reality, can negatively impact the validity of the recommended sample size and number of repeated measures. As power analyses play an increasingly important role in the preparation phase of research projects, an RI-CLPM-specific strategy for the design of a power analysis is detailed, and implemented in the R-package powRICLPM. This paper focuses on the (basic) bivariate RI-CLPM, and extensions to include constraints over time, measurement error (leading to the stable trait autoregressive trait state model), non-normal data, and bounded estimation.

Mulder et al.: Three extensions of the random intercept cross-lagged panel model
Mulder-Hamaker-2020

Abstract: The random intercept cross-lagged panel model (RI-CLPM) is rapidly gaining popularity in psychology and related fields as a structural equation modeling (SEM) approach to longitudinal data. It decomposes observed scores into within-unit dynamics and stable, between-unit differences. This paper discusses three extensions of the RI-CLPM that researchers may be interested in, but are unsure of how to accomplish: (a) including stable, person-level characteristics as predictors and/or outcomes; (b) specifying a multiple-group version; and (c) including multiple indicators. For each extension, we discuss which models need to be run in order to investigate underlying assumptions,

and we demonstrate the various modeling options using a motivating example. We provide fully annotated code for lavaan (R-package) and Mplus on an accompanying website.

B. O. Muthén et al.: Latent transition analysis with random intercepts (RI-LTA)

Muthen-Asparouhov-2022

Abstract: This article demonstrates that the regular LTA model is unnecessarily restrictive and that an alternative model is readily available that typically fits the data much better, leads to better estimates of the transition probabilities, and extracts new information from the data. By allowing random intercept variation in the model, between-subject variation is separated from the within-subject latent class transitions over time allowing a clearer interpretation of the data. Analysis of two examples from the literature demonstrates the advantages of random intercept LTA. Model variations include Mover-Stayer analysis, measurement invariance analysis, and analysis with covariates.

B. O. Muthén et al.: General longitudinal modeling of individual differences in experimental designs: A latent variable framework for analysis and power estimation

Muthen-Curran-1997

Abstract: The generality of latent variable modeling of individual differences in development over time is demonstrated with a particular emphasis on randomized intervention studies. First, a brief overview is given of biostatistical and psychometric approaches to repeated measures analysis. Sec-

ond, the generality of the psychometric approach is indicated by some nonstandard models. Third, a multiple-population analysis approach is proposed for the estimation of treatment effects. The approach clearly describes the treatment effect as development that differs from normative, control-group development. This framework allows for interactions between treatment and initial status in their effects on development. Finally, an approach for the estimation of power to detect treatment effects in this framework is demonstrated. Illustrations of power calculations are carried out with artificial data, varying the sample sizes, number of timepoints, and treatment effect sizes. Real data are used to illustrate analysis strategies and power calculations. Further modeling extensions are discussed.

L. K. Muthén et al.: Mplus user's guide. Eighth edition   **Muthen-Muthen-2017**

Linda K. Muthén and Bengt O. Muthén. *Mplus user's guide. Eighth edition.* Los Angeles, CA: Muthén & Muthén, 2017.

National Research Council: An assessment of research-doctorate programs in the United States: Social and behavioral sciences   **NationalResearchCouncil-1982**

National Research Council. *An assessment of research-doctorate programs in the United States: Social and behavioral sciences.* Washington, D.C.: National Academies Press, Jan. 1982. DOI: 10.17226/9781.

Neale et al.: OpenMx 2.0: Extended structural equation and statistical modeling
                                                         Neale-Hunter-Pritikin-etal-2015

Michael C. Neale et al. "OpenMx 2.0: Extended structural equation and statistical modeling". In: *Psychometrika* 81.2 (Jan. 2015), pp. 535–549. DOI: 10.1007/s11336-014-9435-8.

Abstract: The new software package OpenMx 2.0 for structural equation and other statistical modeling is introduced and its features are described. OpenMx is evolving in a modular direction and

now allows a mix-and-match computational approach that separates model expectations from fit functions and optimizers. Major backend architectural improvements include a move to swappable open-source optimizers such as the newly written CSOLNP. Entire new methodologies such as item factor analysis and state space modeling have been implemented. New model expectation functions including support for the expression of models in LISREL syntax and a simplified multigroup expectation function are available. Ease-of-use improvements include helper functions to standardize model parameters and compute their Jacobian-based standard errors, access to model components through standard R $ mechanisms, and improved tab completion from within the R Graphical User Interface.

Nel: A matrix derivation of the asymptotic covariance matrix of sample correlation coefficients   Nel-1985

Abstract: The asymptotic covariance matrix of the sample correlation matrix is derived in matrix form as an application of some new matrix theory in multivariate statistics.

Nesselroade et al.: Sometimes, it's okay to factor difference scores" - The separation of state and trait anxiety   Nesselroade-Cable-1974

Abstract: Contemporary psychometric policy and practice have tended to make the use of algebraic difference scores in psychological research taboo. Within the more limited domain of factor analytic research on personality, difference scores have been the subject of sporadic debate for more than 30

years. Using the personality trait versus state distinction as a substantive context, the fit of the factor analytic model to difference score data is investigated and found to be quite good. Methodological issues related to properties of difference scores and their implications for personality research are briefly discussed.

Nesselroade et al.: Dynamic factor analysis models for representing process in multivariate time-series

<div align="right">Nesselroade-McArdle-Aggen-etal-2002</div>

John R. Nesselroade, John J. McArdle, et al. "Dynamic factor analysis models for representing process in multivariate time-series". In: *Modeling intraindividual variability with repeated measures data: Methods and applications.* Ed. by Scott L. Hershberger and Debbie S. Moskowitz. Psychology Press, 2002, pp. 235–265. ISBN: 9781410604477. DOI: 10.4324/9781410604477-9.

Abstract: The collection of multivariate time-series and their analysis with mathematical models is necessary if we are effectively and fully to represent process and change. Among the promising applications currently available are variations of the common factor model that integrate factor and time series modeling features in a common analytic framework. We highlight some differences and similarities between two kinds of time series models for common factors: a direct autoregressive factor score (DAFS) model and a white-noise factor score (WNFS) model. Particular specifications of these models are fitted to data reflecting short-term changes in an intensively measured individual's self-reported affect. Results of the model fitting underscore the importance of explicit differences in model specifications that define one's view of the nature of process and change.

Newey et al.: A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix

<div align="right">Newey-West-1987</div>

Whitney K. Newey and Kenneth D. West. "A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix". In: *Econometrica* 55.3 (May 1987), p. 703. DOI: 10.2307/1913610.

Abstract: Introduction and Aims: Comprehensively investigating alcohol-related behaviours in the context of a dynamic multi-day alcohol-licensed event is important for understanding and minimising patron risk. We aimed to assess the measurement utility of implementing a multi-dimensional alcohol assessment battery using biometric data collection, real-time drink logs and retrospective self-report measures over the course of a 4-day music festival. Methods: Fourteen adults participated ($n = 7$ male, mean age 21.9 years). Breath and transdermal alcohol concentration (BrAC and TAC, respectively) were measured using breathalysers and transdermal alcohol bracelets. A real-time drink log was completed via smartphones on initiating each drink, and a retrospective questionnaire was administered up to twice daily throughout the event (6 timepoints total). Results: While almost all participants (92.9%) logged significantly fewer drinks in real-time than they retrospectively reported via the twice-daily questionnaires, logs provided important contextual information including the types of drinks consumed and drinking intensity. Compared to BrAC, TAC provided a better understanding of the time course of intoxication, indicating highest alcohol consumption outside of static BrAC assessment windows. However, BrAC provided a better assessment of present state: all participants were 0.00% BrAC at departure despite over two-fifths (42.9%) of the sample's last TAC reading exceeding 0.00%. Conclusions: As standalone assessments, each method possessed limitations. As a combined battery, they were successfully administered simultaneously, resulting in a more comprehensive overview of alcohol consumption/intoxication over the prolonged drinking session. However, the marked burden of simultaneous administration should be considered, and measures should be chosen judiciously based on research needs.

Abstract: Aims: As a formative step towards determining the accuracy of self-reported drinking levels commonly used for estimating population alcohol use, the validity of a 'last occasion' self-reporting approach is tested with corresponding field observations of participants' drinking quantity. This study is the first known attempt to validate the accuracy of self-reported alcohol consumption using data from a natural setting. Methods: A total of 81 young adults (aged 18-25 years) were purposively selected in Perth, Western Australia. Participants were asked to report the number of alcoholic drinks consumed at nightlife venues 1-2 days after being observed by peer-based researchers on 239 occasions. Complete observation data and self-report estimates were available for 129 sessions, which were fitted with multi-level models assessing the relationship between observed and reported consumption. Results: Participants accurately estimated their consumption when engaging in light to moderate drinking (eight or fewer drinks in a single session), with no significant difference between the mean reported consumption and the mean observed consumption. In contrast, participants underestimated their own consumption by increasing amounts when engaging in heavy drinking of more than eight drinks. Conclusion: It is suggested that recent recall methods in self-report surveys are potentially reasonably accurate measures of actual drinking levels for light to moderate drinkers, but that underestimating of alcohol consumption increases with heavy consumption. Some of the possible reasons for underestimation of heavy drinking are discussed, with both cognitive and socio-cultural factors considered.

Abstract: The Rocker Project provides widely used Docker images for R across different application scenarios. This article surveys downstream projects that build upon the Rocker Project images and presents the current state of R packages for managing Docker images and controlling containers. These use cases cover diverse topics such as package development, reproducible research, collaborative work, cloud-based data processing, and production deployment of services. The variety of applications demonstrates the power of the Rocker Project specifically and containerisation in general. Across the diverse ways to use containers, we identified common themes: reproducible environments, scalability and efficiency, and portability across clouds. We conclude that the current growth and diversification of use cases is likely to continue its positive impact, but see the need for consolidating the Rockerverse ecosystem of packages, developing common practices for applications, and exploring alternative containerisation software.

Nylund et al.: Deciding on the number of classes in latent class analysis and growth mixture modeling: A monte carlo simulation study

Nylund-Asparouhov-Muthen-2007

Abstract: Mixture modeling is a widely applied data analysis technique used to identify unobserved heterogeneity in a population. Despite mixture models' usefulness in practice, one unresolved issue

in the application of mixture models is that there is not one commonly accepted statistical indicator for deciding on the number of classes in a study population. This article presents the results of a simulation study that examines the performance of likelihood-based tests and the traditionally used Information Criterion (ICs) used for determining the number of classes in mixture modeling. We look at the performance of these tests and indexes for 3 types of mixture models: latent class analysis (LCA), a factor mixture model (FMA), and a growth mixture models (GMM). We evaluate the ability of the tests and indexes to correctly identify the number of classes at three different sample sizes ($n = 200, 500, 1,000$). Whereas the Bayesian Information Criterion performed the best of the ICs, the bootstrap likelihood ratio test proved to be a very consistent indicator of classes across all of the models considered.

O'Laughlin et al.: Cross-sectional analysis of longitudinal mediation processes

<div align="right">

OLaughlin-Martin-Ferrer-2018

</div>

Abstract: Statistical mediation analysis can help to identify and explain the mechanisms behind psychological processes. Examining a set of variables for mediation effects is a ubiquitous process in the social sciences literature; however, despite evidence suggesting that cross-sectional data can misrepresent the mediation of longitudinal processes, cross-sectional analyses continue to be used in this manner. Alternative longitudinal mediation models, including those rooted in a structural equation modeling framework (cross-lagged panel, latent growth curve, and latent difference score models) are currently available and may provide a better representation of mediation processes for longitudinal data. The purpose of this paper is twofold: first, we provide a comparison of cross-sectional and longitudinal mediation models; second, we advocate using models to evaluate mediation effects that capture the temporal sequence of the process under study. Two separate empirical examples are presented to illustrate differences in the conclusions drawn from cross-sectional and longitudinal

mediation analyses. Findings from these examples yielded substantial differences in interpretations between the cross-sectional and longitudinal mediation models considered here. Based on these observations, researchers should use caution when attempting to use cross-sectional data in place of longitudinal data for mediation analyses.

O'Rourke et al.: Reasons for testing mediation in the absence of an intervention effect: A research imperative in prevention and intervention research

Abstract: Objective: Mediation models are used in prevention and intervention research to assess the mechanisms by which interventions influence outcomes. However, researchers may not investigate mediators in the absence of intervention effects on the primary outcome variable. There is emerging evidence that in some situations, tests of mediated effects can be statistically significant when the total intervention effect is not statistically significant. In addition, there are important conceptual and practical reasons for investigating mediation when the intervention effect is non-significant. Method: This article discusses the conditions under which mediation may be present when an intervention effect does not have a statistically significant effect and why mediation should always be considered important. Results: Mediation may be present in the following conditions: when the total and mediated effects are equal in value, when the mediated and direct effects have opposing signs, when mediated effects are equal across single and multiple-mediator models, and when specific mediated effects have opposing signs. Mediation should be conducted in every study because it provides the opportunity to test known and replicable mediators, to use mediators as an intervention manipulation check, and to address action and conceptual theory in intervention models. Conclusions: Mediators are central to intervention programs, and mediators should be investigated for the valuable information they provide about the success or failure of interventions.

Abstract: This chapter describes the theoretical and practical importance of mediation analysis in substance-use prevention research. The most important reason for including mediators in a research study is to examine the mechanisms by which prevention programs influence substance-use outcomes. Understanding the mechanisms by which prevention programs achieve effects helps reduce the cost and increases effectiveness of prevention programs. This chapter first describes the theoretical foundations of the mediation model in prevention, and reasons for using mediation analysis in substance-use prevention. Next, we provide an overview of statistical mediation analysis for single and multiple mediator models. We summarize mediation analyses in substance-use prevention and discuss future directions for application of mediation analysis in substance-use research.

Abstract: Objective: Psychiatric symptoms play a crucial role in psychology and psychiatry. However, little is known about how dimensions of symptoms—other than symptom level—relate to psychiatric outcomes. Until recently, methods for measuring dynamic aspects of symptoms have not been available to clinicians or researchers. The authors sought to test whether systematic patterns of change in psychiatric symptoms can be recovered across weekly assessments of individ-

uals at high risk for violence. A secondary objective was to explore whether dynamic features of symptoms (specifically, oscillation speed and dysregulation) are concurrently associated with violence, an important indicator of functional impairment for these individuals. Method: Participants ($N = 132$) were drawn from a sample of patients evaluated at the emergency room of an urban psychiatric hospital. Patients actuarially classified as being at high risk for violence were eligible for participation in the study. Participants and collateral informants were interviewed weekly for 26 weeks following an acute psychiatric evaluation. Psychiatric symptoms were assessed using the Brief Symptom Inventory. Measures of symptom fluctuation and regulation were derived using dynamical systems models. Involvement in violence was assessed using self, informant, and official reports. Results: Individuals' symptom dynamics were recovered by a linear oscillator model that described how quickly symptoms oscillated and whether symptoms were amplifying or moving back toward equilibrium across time. Patterns of rapid symptom fluctuation and symptom amplification were concurrently associated with violence. Conclusions: Psychiatric researchers and clinicians have long been interested in adopting more dynamic approaches to understanding symptom change. This study is the first to demonstrate that systematic fluctuations in symptom patterns may be captured by dynamic models. Moreover, the concurrent association between symptom dynamics and violence suggests avenues for future research to test how features of symptom fluctuation could affect behavior.

Odgers et al.: Violence exposure is associated with adolescents' same- and next-day mental health symptoms Odgers-Russell-2017

Abstract: Background: Young people exposed to violence are at increased risk for mental health and behavioral problems. However, very little is known about the immediate, or same-day, associations between violence exposure and adolescents' mental health symptoms or whether daily symptom

or behavioral reactivity marks future problems. Methods: Young adolescents were assessed three times a day for 30 consecutive days using mobile-phone-based Ecological Momentary Assessment (EMA) ($N = 151$ adolescents). Over 12,500 assessments and 4,329 person days were obtained via the EMA. Adolescents were recruited from low-income neighborhoods based on parent-reported risk for externalizing symptoms. Mental health symptoms were assessed via parent and child report at baseline, multiple times per day via EMA assessments of the adolescents, and again 18 months later when 93% of the adolescents were reinterviewed. Results: Results from multilevel models illustrated that young adolescents were more likely to experience symptoms of anger (OR = 1.74, CI: 1.31-2.30), depression (OR = 1.66, CI: 1.26-2.19), and conduct problems (OR = 2.63, CI: 1.71-4.04) on days that they were exposed versus not exposed to violence. Increases in depressive symptoms were also observed on days following violence exposure (OR = 1.46, CI: 1.09-1.97). Adolescents with the highest levels of violence exposure across the 30-day EMA were less behaviorally reactive to violence exposures in daily life, and heightened behavioral reactivity predicted increased risk for substance use across early adolescence. Conclusions: Findings support the need to focus on both the immediate and long-term associations between violence exposure and adolescents' mental health and behavior. Results also suggest that heightened behavioral reactivity during early adolescence may signal emerging substance use problems.

Oehlert: A note on the delta method Oehlert-1992

Abstract: The delta method is an intuitive technique for approximating the moments of functions of random variables. This note reviews the delta method and conditions under which delta-method approximate moments are accurate.

Oh et al.: Measurement model misspecification in dynamic structural equation models: Power, reliability, and other considerations   **Oh-Hunter-Chow-2025**

Abstract: Dynamic Structural Equation Models (DSEMs) integrate multilevel modeling, time series analysis, and structural equation modeling within a Bayesian estimation framework, offering a versatile tool for analyzing intensive longitudinal data (ILD). However, the impact of measurement structure misspecification in DSEMs, especially under varying reliability conditions and model complexities, remains underexplored. Our Monte Carlo simulation revealed that omitting measurement errors when present led to severe biases in dynamic parameters regardless of reliability conditions, though power remained high. Increasing the number of participants and time points ameliorated but did not eliminate all biases. A single-indicator DSEMs with a measurement structure using composite scores showed similar performance to multiple indicators DSEMs. Empirical applications showed discrepancies in dynamic parameters based on the number of indicators and measurement structures used. Leveraging these findings, we provide design recommendations, functions for extending reliability indices from single-indicator to multiple-indicator models, and guidelines for power evaluations under different reliability conditions.

Ollendick et al.: Advances in clinical child psychology   **Ollendick-Prinz-1996**

Abstract: As in past volumes, the current volume of Advances in Clinical Child Psychology strives for a broad range of timely topics on the study and treatment of children, adolescents, and families. Volume 18 includes a new array of contributions covering issues pertaining to treatment, etiology,

and psychosocial context. The first two contributions address conduct problems. Using qualitative research methods, Webster-Stratton and Spitzer take a unique look at what it is like to be a parent of a young child with conduct problems as well as what it is like to be a participant in a parent training program. Chamberlain presents research on residential and foster-care treatment for adolescents with conduct disorder. As these chapters well reflect, Webster-Stratton, Spitzer, and Chamberlain are all veterans of programmatic research on treatment of child and adolescent conduct problems. Wills and Filer describe an emerging stress-coping model that has been applied to adolescent substance use and is empirically well justified. This model has implications for furthering intervention strategies as well as enhancing our scientific understanding of adolescents and the development of substance abuse. Foster, Martinez, and Kulberg confront the issue that researchers face pertaining to race and ethnicity as it relates to our understanding of peer relations. This chapter addresses some of the measurement and conceptual challenges relative to assessing ethnic variables and relating these to social cognitions of peers, friendship patterns, and peer acceptance.

Oravecz et al.: A hierarchical latent stochastic differential equation model for affective dynamics   Oravecz-Tuerlinckx-Vandekerckhove-2011

Abstract: In this article a continuous-time stochastic model (the Ornstein-Uhlenbeck process) is presented to model the perpetually altering states of the core affect, which is a 2-dimensional concept underlying all our affective experiences. The process model that we propose can account for the temporal changes in core affect on the latent level. The key parameters of the model are the average position (also called home base), the variances and covariances of the process, and the regulatory mechanisms that keep the process in the vicinity of the average position. To account for individual differences, the model is extended hierarchically. A particularly novel contribution is that in principle all parameters of the stochastic process (not only the mean but also its variance

195

and the regulatory parameters) are allowed to differ between individuals. In this way, the aim is to understand the affective dynamics of single individuals and at the same time investigate how these individuals differ from one another. The final model is a continuous-time state-space model for repeated measurement data taken at possibly irregular time points. Both time-invariant and time-varying covariates can be included to investigate sources of individual differences. As an illustration, the model is applied to a diary study measuring core affect repeatedly for several individuals (thereby generating intensive longitudinal data).

Oravecz et al.: On fitting a continuous-time stochastic process model in the Bayesian framework   Oravecz-Wood-Ram-2018

Abstract: Process models can be viewed as mathematical tools that allow researchers to formulate and test theories on the data-generating mechanism underlying observed data. In this chapter we highlight the advantages of this approach by proposing a multilevel, continuous-time stochastic process model to capture the dynamical homeostatic process that underlies observed intensive longitudinal data. Within the multilevel framework, we also link the dynamical processes parameters to time-varying and time-invariant covariates. However, estimating all model parameters (e.g., process model parameters and regression coefficients) simultaneously requires custom-made implementation of the parameter estimation; therefore we advocate the use of a Bayesian statistical framework for fitting these complex process models. We illustrate application to data on self-reported affective states collected in an ecological momentary assessment setting.

Abstract: In virtually all areas of psychology, the question of whether a particular construct has a prospective effect on another is of fundamental importance. For decades, the cross-lagged panel model (CLPM) has been the model of choice for addressing this question. However, CLPMs have recently been critiqued, and numerous alternative models have been proposed. Using the association between low self-esteem and depression as a case study, we examined the behavior of seven competing longitudinal models in 10 samples, each with at least four waves of data and sample sizes ranging from 326 to 8,259. The models were compared in terms of convergence, fit statistics, and consistency of parameter estimates. The traditional CLPM and the random intercepts cross-lagged panel model (RI-CLPM) converged in every sample, whereas the other models frequently failed to converge or did not converge properly. The RI-CLPM exhibited better model fit than the CLPM, whereas the CLPM produced more consistent cross-lagged effects (both across and within samples) than the RI-CLPM. We discuss the models from a conceptual perspective, emphasizing that the models test conceptually distinct psychological and developmental processes, and we address the implications of the empirical findings with regard to model selection. Moreover, we provide practical recommendations for researchers interested in testing prospective associations between constructs and suggest using the CLPM when focused on between-person effects and the RI-CLPM when focused on within-person effects.

Abstract: Regularized continuous time structural equation models are proposed to address two recent challenges in longitudinal research: Unequally spaced measurement occasions and high model complexity. Unequally spaced measurement occasions are part of most longitudinal studies, sometimes intentionally (e.g., in experience sampling methods) sometimes unintentionally (e.g., due to missing data). Yet, prominent dynamic models, such as the autoregressive cross-lagged model, assume equally spaced measurement occasions. If this assumption is violated parameter estimates can be biased, potentially leading to false conclusions. Continuous time structural equation models (CTSEM) resolve this problem by taking the exact time point of a measurement into account. This allows for any arbitrary measurement scheme. We combine CTSEM with LASSO and adaptive LASSO regularization. Such regularization techniques are especially promising for the increasingly complex models in psychological research, the most prominent example being network models with often dozens or hundreds of parameters. Here, LASSO regularization can reduce the risk of overfitting and simplify the model interpretation. In this article we highlight unique challenges in regularizing continuous time dynamic models, such as standardization or the optimization of the objective function, and offer different solutions. Our approach is implemented in the R (R Core Team, 2022) package regCtsem. We demonstrate the use of regCtsem in a simulation study, showing that the proposed regularization improves the parameter estimates, especially in small samples. The approach correctly eliminates true-zero parameters while retaining true-nonzero parameters. We present two empirical examples and end with a discussion on current limitations and future research directions.

Osborne et al.: A Longitudinal Investigation of the Intellectual Differentiation Hypothesis   **Osborne-Suddick-1972**

R. T. Osborne and D. E. Suddick. "A Longitudinal Investigation of the Intellectual Differentiation Hypothesis". In: *The Journal of Genetic Psychology* 121.1 (Sept. 1972), pp. 83–89. ISSN: 1940-0896. DOI: 10.1080/00221325.1972.10533131.

Ou et al.: What's for dynr: A package for linear and nonlinear dynamic modeling in R   **Ou-Hunter-Chow-2019**

Lu Ou, Michael D. Hunter, and Sy-Miin Chow. "What's for dynr: A package for linear and nonlinear dynamic modeling in R". In: *The R Journal* 11.1 (2019), p. 91. DOI: 10.32614/rj-2019-012.

Abstract: Intensive longitudinal data in the behavioral sciences are often noisy, multivariate in nature, and may involve multiple units undergoing regime switches by showing discontinuities interspersed with continuous dynamics. Despite increasing interest in using linear and nonlinear differential/difference equation models with regime switches, there has been a scarcity of software packages that are fast and freely accessible. We have created an R package called dynr that can handle a broad class of linear and nonlinear discreteand continuous-time models, with regime-switching properties and linear Gaussian measurement functions, in C, while maintaining simple and easy-to learn model specification functions in R. We present the mathematical and computational bases used by the dynr R package, and present two illustrative examples to demonstrate the unique features of dynr.

J. H. Oud et al.: Longitudinal factor score estimation using the Kalman filter

Oud-vandenBercken-Essers-1990

Johan H. Oud, John H. van den Bercken, and Raymond J. Essers. "Longitudinal factor score estimation using the Kalman filter". In: *Applied Psychological Measurement* 14.4 (Dec. 1990), pp. 395–418. DOI: 10.1177/014662169001400406.

Abstract: The advantages of the Kalman filter as a factor score estimator in the presence of longitudinal data are described. Because the Kalman filter presupposes the availability of a dynamic state space model, the state space model is reviewed first, and it is shown to be translatable into the LISREL model. Several extensions of the LISREL model specification are discussed in order to enhance the applicability of the Kalman filter for behavioral research data. The Kalman filter and its main properties are summarized. Relationships are shown between the Kalman filter and two well-known cross-sectional factor score estimators: the regression estimator, and the Bartlett estimator. The indeterminacy problem of factor scores is also discussed in the context of Kalman filtering, and the differences are described between Kalman filtering on the basis of a zero-means and a structured-means LISREL model. By using a structured-means LISREL model, the Kalman filter is capable of estimating absolute latent developmental curves. An educational research example is presented. Index terms: factor score estimation, indeterminacy of factor scores, Kalman filter, L,ISREL longitudinal LISREL modeling, longitudinal factor analysis, state space modeling.

J. H. L. Oud et al.: Continuous time modeling of panel data by means of SEM

Oud-Delsing-2010

Abstract: After a brief history of continuous time modeling and its implementation in panel analysis by means of structural equation modeling (SEM), the problems of discrete time modeling are discussed in detail. This is done by means of the popular cross-lagged panel design. Next, the exact discrete model (EDM) is introduced, which accounts for the exact nonlinear relationship between the underlying continuous time model and the resulting discrete time model for data analysis. In addition, a linear approximation of the EDM is discussed: the approximate discrete model (ADM). It is recommended to apply the ADM-SEM procedure by means of a SEM program such as LISREL

in the model building phase and the EDM-SEM procedure by means of Mx in the final model estimation phase. Both procedures are illustrated in detail by two empirical examples: Externalizing and Internalizing Problem Behavior in children; Individualism, Nationalism and Ethnocentrism in the Flemish electorate.

J. H. L. Oud et al.: Continuous time state space modeling of panel data by means of SEM   Oud-Jansen-2000

Abstract: Maximum likelihood parameter estimation of the continuous time linear stochastic state space model is considered on the basis of largeN discrete time data using a structural equation modeling (SEM) program. Random subject effects are allowed to be part of the model. The exact discrete model (EDM) is employed which links the discrete time model parameters to the underlying continuous time model parameters by means of nonlinear restrictions. The EDM is generalized to cover not only time-invariant parameters but also the cases of stepwise time-varying (piecewise time-invariant) parameters and parameters varying continuously over time according to a general polynomial scheme. The identification of the continuous time parameters is discussed and an educational example is presented.

Park et al.: Subgrouping with chain graphical VAR models

Park-Chow-Epskamp-etal-2024

Abstract: Recent years have seen the emergence of an "idio-thetic" class of methods to bridge the gap between nomothetic and idiographic inference. These methods describe nomothetic trends in

idiographic processes by pooling intraindividual information across individuals to inform group-level inference or vice versa. The current work introduces a novel "idio-thetic" model: the subgrouped chain graphical vector autoregression (scGVAR). The scGVAR is unique in its ability to identify subgroups of individuals who share common dynamic network structures in both lag(1) and contemporaneous effects. Results from Monte Carlo simulations indicate that the scGVAR shows promise over similar approaches when clusters of individuals differ in their contemporaneous dynamics and in showing increased sensitivity in detecting nuanced group differences while keeping Type-I error rates low. In contrast, a competing approach-the Alternating Least Squares VAR (ALS VAR) performs well when groups were separated by larger distances. Further considerations are provided regarding applications of the ALS VAR and scGVAR on real data and the strengths and limitations of both methods.

Park et al.: On subgrouping continuous processes in discrete time

Jonathan J. Park, Zachary Fisher, et al. "On subgrouping continuous processes in discrete time". In: *Multivariate Behavioral Research* 58.1 (Jan. 2023), pp. 154–155. ISSN: 1532-7906. DOI: 10.1080/00273171.2022.2160957.

Park et al.: Evaluating discrete time methods for subgrouping continuous processes

Jonathan J. Park, Zachary F. Fisher, Sy-Miin Chow, et al. "Evaluating discrete time methods for subgrouping continuous processes". In: *Multivariate Behavioral Research* 59.6 (Aug. 2023), pp. 1240–1252. ISSN: 1532-7906. DOI: 10.1080/00273171.2023.2235685.

Abstract: Rapid developments over the last several decades have brought increased focus and attention to the role of time scales and heterogeneity in the modeling of human processes. To address these emerging questions, subgrouping methods developed in the discrete-time framework—such

as the vector autoregression (VAR)—have undergone widespread development to identify shared nomothetic trends from idiographic modeling results. Given the dependence of VAR-based parameters on the measurement intervals of the data, we sought to clarify the strengths and limitations of these methods in recovering subgroup dynamics under different measurement intervals. Building on the work of Molenaar and collaborators for subgrouping individual time-series by means of the subgrouped chain graphical VAR (scgVAR) and the subgrouping option in the group iterative multiple model estimation (S-GIMME), we present results from a Monte Carlo study aimed at addressing the implications of identifying subgroups using these discrete-time methods when applied to continuous-time data. Results indicate that discrete-time subgrouping methods perform well at recovering true subgroups when the measurement intervals are large enough to capture the full range of a system's dynamics, either via lagged or contemporaneous effects. Further implications and limitations are discussed therein.

Park et al.: Unsupervised model construction in continuous-time

Park-Fisher-Hunter-etal-2025

Jonathan J. Park, Zachary F. Fisher, Michael D. Hunter, et al. "Unsupervised model construction in continuous-time". In: *Structural Equation Modeling: A Multidisciplinary Journal* 32.3 (2025), pp. 377–399. ISSN: 1532-8007. DOI: 10.1080/10705511.2024.2429544.

abstract
Abstract: Many of the advancements reconciling individual- and group-level results have occurred in the context of a discrete-time modeling framework. Discrete-time models are intuitive and offer relatively simple interpretations for the resulting dynamic structures; however, they do not possess the flexibility of models fitted in the continuous-time framework. We introduce ct-gimme, a continuous-time extension of the group iterative multiple model estimation (GIMME) procedure which enables researchers to fit complex, high-dimensional dynamic networks in continuous time. Our results indicate that ct-gimme outperforms $N = 1$ model fitting in continuous time by pooling information across multiple subjects. Likewise, ct-gimme outperforms group-level model fitting

203

in the presence of within-sample heterogeneity. We conclude with an empirical illustration and highlight the limitations of the approach relating to the identification of meaningful starting values.

Pastor et al.: On the multilevel nature of meta-analysis: A tutorial, comparison of software programs, and discussion of analytic choices   Pastor-Lazowski-2017

Abstract: The term "multilevel meta-analysis" is encountered not only in applied research studies, but in multilevel resources comparing traditional meta-analysis to multilevel meta-analysis. In this tutorial, we argue that the term "multilevel meta-analysis" is redundant since all meta-analysis can be formulated as a special kind of multilevel model. To clarify the multilevel nature of meta-analysis the four standard meta-analytic models are presented using multilevel equations and fit to an example data set using four software programs: two specific to meta-analysis (metafor in R and SPSS macros) and two specific to multilevel modeling (PROC MIXED in SAS and HLM). The same parameter estimates are obtained across programs underscoring that all meta-analyses are multilevel in nature. Despite the equivalent results, not all software programs are alike and differences are noted in the output provided and estimators available. This tutorial also recasts distinctions made in the literature between traditional and multilevel meta-analysis as differences between meta-analytic choices, not between meta-analytic models, and provides guidance to inform choices in estimators, significance tests, moderator analyses, and modeling sequence. The extent to which the software programs allow flexibility with respect to these decisions is noted, with metafor emerging as the most favorable program reviewed.

M. Patrick et al.: Monitoring the Future Panel Study annual report: National data on substance use among adults ages 19 to 60, 1976-2022

Patrick-Miech-Johnston-etal-2023

Megan Patrick et al. *Monitoring the Future Panel Study annual report: National data on substance use among adults ages 19 to 60, 1976-2022*. Ann Arbor, MI: Institute for Social Research, The University of Michigan, 2023. DOI: 10.7826/isr-um.06.585140.002.07.0002.2023.

M. E. Patrick et al.: Short-term changes in plans to drink and importance of positive and negative alcohol consequences Patrick-Maggs-2007

Megan E. Patrick and Jennifer L. Maggs. "Short-term changes in plans to drink and importance of positive and negative alcohol consequences". In: *Journal of Adolescence* 31.3 (July 2007), pp. 307–321. ISSN: 1095-9254. DOI: 10.1016/j.adolescence.2007.06.002.

Abstract: Experienced consequences predicted short-term changes in alcohol use plans and perceptions of the importance of alcohol-related consequences. Participants were 176 traditionally aged first-year university students who completed a 10-week telephone diary study (total weeks = 1735). In multi-level models, men and students who experienced more positive and negative consequences on average planned to drink more and rated avoiding negative consequences as less important. Students who experienced more positive consequences rated them as more important (between-person analyses). Following weeks of experiencing relatively more positive drinking consequences, students planned to drink more and rated experiencing positive consequences as more important for the subsequent week (within-person analyses). Challenges for intervening in the ongoing formation of anticipatory cognitions regarding alcohol use are discussed.

M. E. Patrick et al.: Profiles of motivations for alcohol use and sexual behavior among first-year university students Patrick-Maggs-2009

Megan E. Patrick and Jennifer L. Maggs. "Profiles of motivations for alcohol use and sexual behavior

among first-year university students". In: *Journal of Adolescence* 33.5 (Nov. 2009), pp. 755–765. ISSN: 1095-9254. DOI: 10.1016/j.adolescence.2009.10.003.

Abstract: The links between motivations for alcohol use and for sexual behaviors are not well understood. Latent profile analysis was used to identify drinking motivational profiles (based on motivations for: fun/social, relaxation/coping, image, sex; motivations against: physical, behavioral) and sex motivational profiles (motivations for: enhancement, intimacy, coping; motivations against: not ready, health, values) among college students (N = 227, 51% male). Latent profiles for drinking were: low for/high against drinking (34%), average drinking motives (53%), and high for/low against drinking (13%). Profiles for sex were: low for/high against sex (35%), high for/low against sex (42%), and high for with coping/moderate against sex (23%). Motivational profiles were related across behaviors. Drinking motivational profiles were associated with alcohol use and psychosocial adjustment; sex motivational profiles were associated with sexual experiences. Distinct profiles of motivations support the need for differentiated intervention programs targeting individuals with different patterns of reasons for engaging in risk behaviors during late adolescence.

M. E. Patrick et al.: Monitoring the Future Panel Study annual report: National data on substance use among adults ages 19 to 65, 1976-2024

Patrick-Miech-Johnston-etal-2025

Megan E. Patrick, Richard A. Miech, et al. *Monitoring the Future Panel Study annual report: National data on substance use among adults ages 19 to 65, 1976-2024.* Monitoring the Future Monograph Series. Ann Arbor, Michigan: Institute for Social Research, University of Michigan, 2025. DOI: [10.7302/26783](10.7302/26783).

M. E. Patrick et al.: High-intensity drinking by underage young adults in the United States: Underage high-intensity drinking **Patrick-TerryMcElrath-2016**

Megan E. Patrick and Yvonne M. Terry-McElrath. "High-intensity drinking by underage young

adults in the United States: Underage high-intensity drinking". In: *Addiction* 112.1 (Sept. 2016), pp. 82–93. ISSN: 0965-2140. DOI: 10.1111/add.13556.

Abstract: Aims. To estimate (1) the prevalence of underage binge drinking, high-intensity drinking and intoxication among young adults aged 19/20 years; (2) change in these behaviors across the transition out of high school and across historical time; and (3) associations between these behaviors and key covariates, including college status. Design, Setting, Participants. Longitudinal data from the US nationally representative Monitoring the Future study included 1657 respondents first surveyed as 12th graders (modal age 18 years) in 2005-13 and again at modal age 19/20 years in 2006-14. Measurements. Self-reported measures of alcohol use, demographics, college attendance and living situation. Findings. Binge drinking (5+ drinks on one occasion) was reported by 24.2% [95% confidence interval (CI) = 22.0, 26.5] of young adults aged 19/20; 10.3% (CI = 8.7, 11.9) reported high-intensity drinking of 10+ drinks; 4.2% (CI = 3.1, 5.2) reported 15+ drinks. Usual moderate/high intoxication when drinking was reported by 33.1% (CI = 30.6, 35.6); 29.6% (CI = 27.2, 32.0) reported usual sustained intoxication of 3+ hours. Significant variability (P ¡ 0.001) in these behaviors from ages 18 to 19/20 was observed. Significant decreases (P ¡ 0.05) across historical time in 5+ and 10+ drinking were found. Four-year college students not residing with parents had significantly higher odds of moderate/high intoxication, binge drinking and high-intensity drinking compared with other groups (P ¡ 0.001). Conclusions. Young adult underage binge drinking (5+ drinks on one occasion), high-intensity drinking (10+ and 15+ drinks) and intoxication are relatively common in the United States, and show meaningful variability across the transition out of high school. Four-year college students and those who do not live with their parents are more likely to engage in high-intensity drinking than their peers.

Pawitan: In all likelihood: Statistical modelling and inference using likelihood

Pawitan-2013

Yudi Pawitan. *In all likelihood: Statistical modelling and inference using likelihood*. Oxford University Press, Jan. 17, 2013. 544 pp. ISBN: 9780199671229.

207

Abstract: Based on a course in the theory of statistics this text concentrates on what can be achieved using the likelihood/Fisherian method of taking account of uncertainty when studying a statistical problem. It takes the concept ot the likelihood as providing the best methods for unifying the demands of statistical modelling and the theory of inference. Every likelihood concept is illustrated by realistic examples, which are not compromised by computational problems. Examples range from a simile comparison of two accident rates, to complex studies that require generalised linear or semiparametric modelling. The emphasis is that the likelihood is not simply a device to produce an estimate, but an important tool for modelling. The book generally takes an informal approach, where most important results are established using heuristic arguments and motivated with realistic examples. With the currently available computing power, examples are not contrived to allow a closed analytical solution, and the book can concentrate on the statistical aspects of the data modelling. In addition to classical likelihood theory, the book covers many modern topics such as generalized linear models and mixed models, non parametric smoothing, robustness, the EM algorithm and empirical likelihood.

Pe et al.: Updating in working memory predicts greater emotion reactivity to and facilitated recovery from negative emotion-eliciting stimuli

Pe-Koval-Houben-etal-2015

Abstract: That emotions change in response to emotion-eliciting events is a natural part of human life. However, it is equally important for emotions to return to baseline once the emotion-eliciting events have passed. This suggests that the ability to emotionally react to and recover from emotion-eliciting events is critical for healthy psychological functioning. But why do individuals differ in their emotion reactivity and recovery? The present work postulates that the ability to update emotional information in working memory (WM) may explain individual differences in emotion reactivity

and recovery. Two studies are presented, which examined whether updating ability was related to emotion reactivity and recovery. In Study 1, we assessed participants' self-reported affect as they viewed negative and positive films. Our results revealed that better updating ability was related to greater emotion reactivity and facilitated (i.e., quicker) recovery from watching negative films. In Study 2, participants recalled a recent angering event, and were then instructed to either ruminate about or reappraise the event. Results revealed that updating ability was again related to greater emotion reactivity and facilitated (i.e., successful) emotion recovery in response to the angering event, and that this was unrelated to the emotion regulation strategy used. These findings identify the ability to update emotional information in WM as a possible mechanism in emotion responding.

Pesigan: Confidence intervals for standardized coefficients: Applied to regression coefficients in primary studies and indirect effects in meta-analytic structural equation modeling   Pesigan-2022

Ivan Jacob Agaloos Pesigan. "Confidence intervals for standardized coefficients: Applied to regression coefficients in primary studies and indirect effects in meta-analytic structural equation modeling". PhD thesis. University of Macau, 2022.

Pesigan et al.: SEM-based methods to form confidence intervals for indirect effect: Still applicable given nonnormality, under certain conditions   Pesigan-Cheung-2020

Ivan Jacob Agaloos Pesigan and Shu Fai Cheung. "SEM-based methods to form confidence intervals for indirect effect: Still applicable given nonnormality, under certain conditions". In: *Frontiers in Psychology* 11 (Dec. 2020). DOI: 10.3389/fpsyg.2020.571928.

Abstract: A SEM-based approach using likelihood-based confidence interval (LBCI) has been proposed to form confidence intervals for unstandardized and standardized indirect effect in mediation models. However, when used with the maximum likelihood estimation, this approach requires that the variables are multivariate normally distributed. This can affect the LBCIs of unstandardized

and standardized effect differently. In the present study, the robustness of this approach when the predictor is not normally distributed but the error terms are conditionally normal, which does not violate the distributional assumption of ordinary least squares (OLS) estimation, is compared to four other approaches: nonparametric bootstrapping, two variants of LBCI, LBCI assuming the predictor is fixed (LBCI-Fixed-X) and LBCI based on ADF estimation (LBCI-ADF), and Monte Carlo. A simulation study was conducted using a simple mediation model and a serial mediation model, manipulating the distribution of the predictor. The Monte Carlo method performed worst among the methods. LBCI and LBCI-Fixed-X had suboptimal performance when the distributions had high kurtosis and the population indirect effects were medium to large. In some conditions, the problem was severe even when the sample size was large. LBCI-ADF and nonparametric bootstrapping had coverage probabilities close to the nominal value in nearly all conditions, although the coverage probabilities were still suboptimal for the serial mediation model when the sample size was small with respect to the model. Implications of these findings in the context of this special case of nonnormal data were discussed.

Pesigan et al.: Monte Carlo confidence intervals for the indirect effect with missing data Pesigan-Cheung-2024

Abstract: Missing data is a common occurrence in mediation analysis. As a result, the methods used to construct confidence intervals around the indirect effect should consider missing data. Previous research has demonstrated that, for the indirect effect in data with complete cases, the Monte Carlo method performs as well as nonparametric bootstrap confidence intervals (see MacKinnon et al., Multivariate Behavioral Research, 39(1), 99-128, 2004; Preacher & Selig, Communication Methods and Measures, 6(2), 77-98, 2012; Tofighi & MacKinnon, Structural Equation Modeling: A Multidisciplinary Journal, 23(2), 194-205, 2015). In this manuscript, we propose a simple, fast,

and accurate two-step approach for generating confidence intervals for the indirect effect, in the presence of missing data, based on the Monte Carlo method. In the first step, an appropriate method, for example, full-information maximum likelihood or multiple imputation, is used to estimate the parameters and their corresponding sampling variance-covariance matrix in a mediation model. In the second step, the sampling distribution of the indirect effect is simulated using estimates from the first step. A confidence interval is constructed from the resulting sampling distribution. A simulation study with various conditions is presented. Implications of the results for applied research are discussed.

Pesigan et al.: How plausible is my model? Assessing model plausibility of structural equation models using Bayesian posterior probabilities (BPP)

Ivan Jacob Agaloos Pesigan, Shu Fai Cheung, et al. "How plausible is my model? Assessing model plausibility of structural equation models using Bayesian posterior probabilities (BPP)". In: *Behavior Research Methods* (Dec. 2025).

Abstract: In structural equation modeling (SEM), one method to select the most plausible model from several candidates, or to compare one or more hypothesized models with similar alternatives on plausibility, is to compare the models using Bayesian posterior probability (BPP). BPP can be computed from the Bayesian Information Criterion (BIC) scores (Wu et al., 2020). This approach complements conventional goodness-of-fit indices such as the Comparative Fit Index (CFI), the Root Mean Square Error of Approximation (RMSEA), and the standardized root mean square residual (SRMR) in giving concise BPP for assessing uncertainties among all models considered. It can also reveal evidence against a model otherwise hidden by these indices. However, Wu et al. (2020) did not provide guidelines on deciding the models that should be considered. To facilitate the use of BPP, we proposed a novel method for selecting this set of models, called *neighboring models*, to make it easier for researchers to decide on the initial set. This novel method integrates seamlessly into the typical workflow for SEM analysis. Researchers can fit a model as usual and then use this

method to assess whether it is the most plausible model compared with the neighboring models. We believe the proposed method will make it easier for researchers to make better-informed decisions when evaluating their models. We developed a user-friendly R package, modelbpp, to automate all the steps: generating the set of neighboring models, fitting them, and computing the BPPs, all in a single function.

Pesigan et al.: Brief report: Identity processes in Filipino late adolescents and young adults: Parental influences and mental health outcomes

Pesigan-Luyckx-Alampay-2014

Abstract: This study focused on a process-oriented approach to identity formation using a sample of Filipino late adolescents and young adults (17-30 years; $N = 779$). Indirect relations between parenting and mental health via identity formation processes were examined. Two parenting dimensions (psychological control and support), two types of mental health outcomes (depression and psychological well-being), and five identity dimensions (commitment making (CM), identification with commitment (IC), exploration in breadth (EB), exploration in depth (ED), and ruminative exploration (RE)) were assessed. Recursive path analysis showed indirect relations between parenting and mental health via EB, ED, RE, and IC. Model differences between late adolescents (17-21 year olds) and young adults (22-30 year olds) were examined using multigroup path analysis. Results showed that the direct effect of psychological control on RE, and its indirect effect on depression through RE differed between the age groups. Implications and suggestions for future research are provided.

Abstract: Objective: This article introduces the common and unique latent transition analysis (CULTA), a novel approach to studying alcohol intoxication dynamics in young adults engaged in heavy episodic drinking. CULTA merges the common and unique trait-state model with latent transition analysis to separate stable, traitlike intoxication components from transient fluctuations while modeling transitions between distinct drinking profiles. Method: A sample of 222 young adults wore transdermal alcohol concentration sensors for 6 days, capturing real-time alcohol levels. The CULTA model decomposed intoxication variability into common and unique influences across four transdermal alcohol concentration features—peak, rise rate, fall rate, and duration. Latent intoxication profiles were identified, and transition probabilities between profiles were estimated with a focus on the influence of alcohol use disorder risk measured by the Alcohol Use Disorders Identification Test. Results: Two latent intoxication profiles emerged. The first, chronic heavy episodic drinking, was characterized by persistently high intoxication without significant inertia, while the second, inertia-driven drinking, featured moderate episodic intoxication with a strong autoregressive effect, reflecting lingering intoxication that dissipates over time. Individuals with higher Alcohol Use Disorders Identification Test scores were more likely to remain in or transition to the chronic heavy episodic drinking profile. Although peak intoxication and rise rate showed limited individual variability, fall rate and duration varied substantially, marking them as potential targets for intervention. Conclusions: CULTA advances our understanding of alcohol intoxication by distinguishing stable from transient influences and modeling transitions between drinking states. These findings suggest that interventions should address both persistent and situational aspects of intoxication—

213

especially by reducing duration and fall rate—and encourage research across longer periods and populations.

Pesigan et al.: Inferences and effect sizes for direct, indirect, and total effects in continuous-time mediation models   Pesigan-Russell-Chow-2025o

Abstract: Mediation modeling using longitudinal data is an exciting field that captures the interrelations in dynamic changes, such as mediated changes, over time. Even though discrete-time vector autoregressive approaches are commonly used to estimate indirect effects in longitudinal data, they have known limitations due to the dependency of inferential results on the time intervals between successive occasions and the assumption of regular spacing between measurements. Continuous-time vector autoregressive models have been proposed as an alternative to address these issues. Previous work in the area (e.g., Deboeck & Preacher, 2015; Ryan & Hamaker, 2021) has shown how the direct, indirect, and total effects, for a range of time-interval values, can be calculated using parameters estimated from continuous-time vector autoregressive models for causal inferential purposes. However, both standardized effects size measures and methods for calculating the uncertainty around the direct, indirect, and total effects in continuous-time mediation have yet to be explored. Drawing from the mediation model literature, we present and compare results using the delta, Monte Carlo, and parametric bootstrap methods to calculate SEs and confidence intervals for the direct, indirect, and total effects in continuous-time mediation for inferential purposes. Options to automate these inferential procedures and facilitate interpretations are available in the `cTMed` R package.

Abstract: The multivariate delta method was used by Yuan and Chan to estimate standard errors and confidence intervals for standardized regression coefficients. Jones and Waller extended the earlier work to situations where data are nonnormal by utilizing Browne's asymptotic distribution-free (ADF) theory. Furthermore, Dudgeon developed standard errors and confidence intervals, employing heteroskedasticity-consistent (HC) estimators, that are robust to nonnormality with better performance in smaller sample sizes compared to Jones and Waller's ADF technique. Despite these advancements, empirical research has been slow to adopt these methodologies. This can be a result of the dearth of user-friendly software programs to put these techniques to use. We present the betaDelta and the betaSandwich packages in the R statistical software environment in this manuscript. Both the normal-theory approach and the ADF approach put forth by Yuan and Chan and Jones and Waller are implemented by the betaDelta package. The HC approach proposed by Dudgeon is implemented by the betaSandwich package. The use of the packages is demonstrated with an empirical example. We think the packages will enable applied researchers to accurately assess the sampling variability of standardized regression coefficients.

Abstract: Missing data analyses have received considerable recent attention in the methodological literature, and two "modern" methods, multiple imputation and maximum likelihood estimation, are recommended. The goals of this article are to (a) provide an overview of missing-data theory, maximum likelihood estimation, and multiple imputation; (b) conduct a methodological review of missing-data reporting practices in 23 applied research journals; and (c) provide a demonstration of multiple imputation and maximum likelihood estimation using the Longitudinal Study of American Youth data. The results indicated that explicit discussions of missing data increased substantially between 1999 and 2003, but the use of maximum likelihood estimation or multiple imputation was rare; the studies relied almost exclusively on listwise and pairwise deletion.

Piasecki: Assessment of alcohol use in the natural environment — Piasecki-2019

Abstract: The current article critically reviews 3 methodological options for assessing drinking episodes in the natural environment. Ecological momentary assessment (EMA) typically involves using mobile devices to collect self-report data from participants in daily life. This technique is now widely used in alcohol research, but investigators have implemented diverse assessment strategies. This article focuses on "high-resolution" EMA protocols that oversample experiences and behaviors within individual drinking episodes. A number of approaches have been used to accomplish this, including using signaled follow-ups tied to drinking initiation, asking participants to log entries before and after individual drinks or drinking episodes, and delivering frequent signaled assessments during periods of the day when alcohol use is most common. Transdermal alcohol sensors (TAS) are devices that are worn continuously and are capable of detecting alcohol eliminated through the skin. These methods are appealing because they do not rely upon drinkers' self-report. Studies using TAS have been appearing with greater frequency over the past several years. New methods are making the use of TAS more tractable by permitting back-translation of transdermal alcohol

concentration data to more familiar estimates of blood alcohol concentration or breath alcohol concentration. However, the current generation of devices can have problems with missing data and tend to be relatively insensitive to low-level drinking. An emerging area of research investigates the possibility of using mobile device data and machine learning to passively detect the user's drinking, with promising early findings. EMA, TAS, and sensor-based approaches are all valid, and tend to produce convergent information when used in conjunction with one another. Each has a unique profile of advantages, disadvantages, and threats to validity. Therefore, the nature of the underlying research question must dictate the method(s) investigators select.

Politis: The impact of bootstrap methods on time series analysis   Politis-2003

Abstract: Sparked by Efron's seminal paper, the decade of the 1980s was a period of active research on bootstrap methods for independent data–mainly i.i.d. or regression set-ups. By contrast, in the 1990s much research was directed towards resampling dependent data, for example, time series and random fields. Consequently, the availability of valid nonparametric inference procedures based on resampling and/or subsampling has freed practitioners from the necessity of resorting to simplifying assumptions such as normality or linearity that may be misleading.

Politis et al.: The stationary bootstrap   Politis-Romano-1994

Abstract: This article introduces a resampling procedure called the stationary bootstrap as a means of calculating standard errors of estimators and constructing confidence regions for parameters based on weakly dependent stationary observations. Previously, a technique based on resampling blocks

of consecutive observations was introduced to construct confidence intervals for a parameter of the m-dimensional joint distribution of m consecutive observations, where m is fixed. This procedure has been generalized by constructing a "blocks of blocks" resampling scheme that yields asymptotically valid procedures even for a multivariate parameter of the whole (i.e., infinite-dimensional) joint distribution of the stationary sequence of observations. These methods share the construction of resampling blocks of observations to form a pseudo-time series, so that the statistic of interest may be recalculated based on the resampled data set. But in the context of applying this method to stationary data, it is natural to require the resampled pseudo-time series to be stationary (conditional on the original data) as well. Although the aforementioned procedures lack this property, the stationary procedure developed here is indeed stationary and possesses other desirable properties. The stationary procedure is based on resampling blocks of random length, where the length of each block has a geometric distribution. In this article, fundamental consistency and weak convergence properties of the stationary resampling scheme are developed.

Preacher et al.: Computational tools for probing interactions in multiple linear regression, multilevel modeling, and latent curve analysis   Preacher-Curran-Bauer-2006

Abstract: Simple slopes, regions of significance, and confidence bands are commonly used to evaluate interactions in multiple linear regression (MLR) models, and the use of these techniques has recently been extended to multilevel or hierarchical linear modeling (HLM) and latent curve analysis (LCA). However, conducting these tests and plotting the conditional relations is often a tedious and error-prone task. This article provides an overview of methods used to probe interaction effects and describes a unified collection of freely available online resources that researchers can use to obtain significance tests for simple slopes, compute regions of significance, and obtain confidence bands

218

for simple slopes across the range of the moderator in the MLR, HLM, and LCA contexts. Plotting capabilities are also provided.

Preacher et al.: SPSS and SAS procedures for estimating indirect effects in simple mediation models   Preacher-Hayes-2004

Abstract: Researchers often conduct mediation analysis in order to indirectly assess the effect of a proposed cause on some outcome through a proposed mediator. The utility of mediation analysis stems from its ability to go beyond the merely descriptive to a more functional understanding of the relationships among variables. A necessary component of mediation is a statistically and practically significant indirect effect. Although mediation hypotheses are frequently explored in psychological research, formal significance tests of indirect effects are rarely conducted. After a brief overview of mediation, we argue the importance of directly testing the significance of indirect effects and provide SPSS and SAS macros that facilitate estimation of the indirect effect with a normal theory approach and a bootstrap approach to obtaining confidence intervals, as well as the traditional approach advocated by Baron and Kenny (1986). We hope that this discussion and the macros will enhance the frequency of formal mediation tests in the psychology literature. Electronic copies of these macros may be downloaded from the Psychonomic Society's Web archive at www.psychonomic.org/archive/.

Preacher et al.: Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models   Preacher-Hayes-2008

Kristopher J. Preacher and Andrew F. Hayes. "Asymptotic and resampling strategies for assessing

and comparing indirect effects in multiple mediator models". In: *Behavior Research Methods* 40.3 (Aug. 2008), pp. 879–891. DOI: 10.3758/brm.40.3.879.

Abstract: Hypotheses involving mediation are common in the behavioral sciences. Mediation exists when a predictor affects a dependent variable indirectly through at least one intervening variable, or mediator. Methods to assess mediation involving multiple simultaneous mediators have received little attention in the methodological literature despite a clear need. We provide an overview of simple and multiple mediation and explore three approaches that can be used to investigate indirect processes, as well as methods for contrasting two or more mediators within a single model. We present an illustrative example, assessing and contrasting potential mediators of the relationship between the helpfulness of socialization agents and job satisfaction. We also provide SAS and SPSS macros, as well as Mplus and LISREL syntax, to facilitate the use of these methods in applications.

Preacher et al.: Effect size measures for mediation models: Quantitative strategies for communicating indirect effects   Preacher-Kelley-2011

Abstract: The statistical analysis of mediation effects has become an indispensable tool for helping scientists investigate processes thought to be causal. Yet, in spite of many recent advances in the estimation and testing of mediation effects, little attention has been given to methods for communicating effect size and the practical importance of those effect sizes. Our goals in this article are to (a) outline some general desiderata for effect size measures, (b) describe current methods of expressing effect size and practical importance for mediation, (c) use the desiderata to evaluate these methods, and (d) develop new methods to communicate effect size in the context of mediation analysis. The first new effect size index we describe is a residual-based index that quantifies the amount of variance explained in both the mediator and the outcome. The second new

effect size index quantifies the indirect effect as the proportion of the maximum possible indirect effect that could have been obtained, given the scales of the variables involved. We supplement our discussion by offering easy-to-use R tools for the numerical and visual communication of effect size for mediation effects.

Preacher et al.: Advantages of Monte Carlo confidence intervals for indirect effects Preacher-Selig-2012

Kristopher J. Preacher and James P. Selig. "Advantages of Monte Carlo confidence intervals for indirect effects". In: *Communication Methods and Measures* 6.2 (Apr. 2012), pp. 77–98. DOI: 10.1080/19312458.2012.679848.

Abstract: Monte Carlo simulation is a useful but underutilized method of constructing confidence intervals for indirect effects in mediation analysis. The Monte Carlo confidence interval method has several distinct advantages over rival methods. Its performance is comparable to other widely accepted methods of interval construction, it can be used when only summary data are available, it can be used in situations where rival methods (e.g., bootstrapping and distribution of the product methods) are difficult or impossible, and it is not as computer-intensive as some other methods. In this study we discuss Monte Carlo confidence intervals for indirect effects, report the results of a simulation study comparing their performance to that of competing methods, demonstrate the method in applied examples, and discuss several software options for implementation in applied settings.

Prince et al.: Trajectories of college alcohol involvement and their associations with later alcohol use disorder symptoms Prince-Read-Colder-2019

Mark A. Prince, Jennifer P. Read, and Craig R. Colder. "Trajectories of college alcohol involvement and their associations with later alcohol use disorder symptoms". In: *Prevention Science* 20.5 (Jan. 2019), pp. 741–752. ISSN: 1573-6695. DOI: 10.1007/s11121-018-0974-6.

Abstract: Little is known about what differentiates individuals whose drinking patterns escalate into problematic use following the transition out of college compared to those who learn to drink in a way that is consistent with independent adult roles. Patterns of alcohol use and problems during college may pre-sage progression toward problem drinking in adulthood. The present study sought to examine such patterns in an effort to delineate those at greatest risk. Research has not yet elucidated whether, when, and how these groups diverge. Our results indicate that students who report AUD symptoms one year following graduation reported greater alcohol involvement from the first semester and escalated their involvement with alcohol at a more rapid pace. We observed marked and measurable differences in drinking patterns between those who go on to exhibit AUD symptoms following college and those who do not. A close inspection of these differences reveals that relatively small absolute differences in alcohol consumption add up to large differences in alcohol-related consequences. Thus, markers of longer-term risk are present early in college, and greater escalation of drinking across college is an indicator that intervention is needed. Brief Motivational Interventions could help students to anticipate some of the challenges ahead as they transition from the college environment, as well as the potential deleterious effects of immoderate alcohol use on making a successful transition into adult roles. In addition to the beginning of college, our findings also point to critical periods during which screening and brief intervention may be optimally timed.

R Core Team: R: A language and environment for statistical computing

RCoreTeam-2021

R Core Team. *R: A language and environment for statistical computing.* R Foundation for Statistical Computing. Vienna, Austria, 2021. URL: https://www.R-project.org/.

R Core Team: R: A language and environment for statistical computing

RCoreTeam-2022

R Core Team. *R: A language and environment for statistical computing.* R Foundation for Statistical Computing. Vienna, Austria, 2022. URL: https://www.R-project.org/.

R Core Team: R: A language and environment for statistical computing

<div align="right">**RCoreTeam-2023**</div>

R Core Team. *R: A language and environment for statistical computing.* R Foundation for Statistical Computing. Vienna, Austria, 2023. URL: https://www.R-project.org/.

R Core Team: R: A language and environment for statistical computing

<div align="right">**RCoreTeam-2024**</div>

R Core Team. *R: A language and environment for statistical computing.* R Foundation for Statistical Computing. Vienna, Austria, 2024. URL: https://www.R-project.org/.

R Core Team: R: A language and environment for statistical computing

<div align="right">**RCoreTeam-2025**</div>

R Core Team. *R: A language and environment for statistical computing.* R Foundation for Statistical Computing. Vienna, Austria, 2025. URL: https://www.R-project.org/.

Trivellore E. Raghunathan et al.: A multivariate technique for multiply imputing missing values using a sequence of regression models

<div align="right">**Raghunathan-Lepkowski-vanHoewyk-etal-2001**</div>

Trivellore E. Raghunathan et al. "A multivariate technique for multiply imputing missing values using a sequence of regression models". In: *Survey Methodology* 27.1 (2001), pp. 85–95.

Abstract: This article describes and evaluates a procedure for imputing missing values for a relatively complex data structure when the data are missing at random. The imputations are obtained by fitting a sequence of regression models and drawing values from the corresponding predictive distributions. The types of regression models used are linear, logistic, Poisson, generalized logit or a mixture of these depending on the type of variable being imputed. Two additional common

features in the imputation process are incorporated: restriction to a relevant subpopulation for some variables and logical bounds or constraints for the imputed values. The restrictions involve subsetting the sample individuals that satisfy certain criteria while fitting the regression models. The bounds involve drawing values from a truncated predictive distribution. The development of this method was partly motivated by the analysis of two data sets which are used as illustrations. The sequential regression procedure is applied to perform multiple imputation analysis for the two applied problems. The sampling properties of inferences from multiply imputed data sets created using the sequential regression method are evaluated through simulated data sets.

Rasmussen: Estimating correlation coefficients: Bootstrap and parametric approaches
Rasmussen-1987

Jeffrey L. Rasmussen. "Estimating correlation coefficients: Bootstrap and parametric approaches". In: *Psychological Bulletin* 101.1 (1987), pp. 136–139. DOI: 10.1037/0033-2909.101.1.136.

Abstract: The bootstrap, a computer-intensive approach to statistical data analysis, has been recommended as an alternative to parametric approaches. Advocates claim it is superior because it is not burdened by potentially unwarranted normal theory assumptions and because it retains information about the form of the original sample. Empirical support for its superiority, however, is quite limited. The present article compares the bootstrap and parametric approaches to estimating confidence intervals and Type I error rates of the correlation coefficient. The parametric approach is superior to the bootstrap under both assumption violation and nonviolation. The bootstrap results in overly restricted confidence intervals and overly liberal Type I error rates.

Ray et al.: Capturing habitualness of drinking and smoking behavior in humans
Ray-Du-Grodin-etal-2020

Lara A. Ray et al. "Capturing habitualness of drinking and smoking behavior in humans". In: *Drug*

Abstract: Background: Recent findings suggest that overreliance on habit may be common in individuals diagnosed with addiction. To advance our understanding of habit in clinical samples and from behavioral measures, this study examines the interrelations between self-reported habit index for smoking and drinking as well as behavioral measures of intraindividual variability in smoking and drinking. Methods: Treatment-seeking heavy drinking smokers ($N = 416$) completed the Self-Report Habit Index (SRHI) adapted for both smoking and drinking. "Behavioral habitualness" was computed from the degree of intraindividual variability in patterns of smoking and drinking over the past month. Using the 28-day Timeline-Follow Back (TLFB) interview, we derived two measures of intraindividual variability: interclass correlation (ICC) and autocorrelation [AR(7) coefficients]. Results: Self-report measures of habit were robustly associated with clinical severity of drinking and smoking with higher habit scores indicating greater severity of drinking and smoking, respectively. ICC and AR(7) coefficients, the behavioral measure of "patterness" and putative habit, were not associated with SRHI scores. While ICC for smoking was associated with higher nicotine dependence scores, this pattern was not found for drinking ICC and alcohol problem severity. Conclusions: These results support the construct validity of the self-report measures of habit for smoking and drinking, as well an initial evaluation of behavioral measure of smoking "patterness" as a potential proxy for habit smoking. Because habit represents a complex phenotype with limited clinical translation, additional studies capturing a wider range of substance use severity and coupled with brain-based validation methods are warranted.

Raykov et al.: Using the delta method for approximate interval estimation of parameter functions in SEM   Raykov-Marcoulides-2004

Abstract: In applications of structural equation modeling, it is often desirable to obtain measures of uncertainty for special functions of model parameters. This article provides a didactic discussion of how a method widely used in applied statistics can be employed for approximate standard error and confidence interval evaluation of such functions. The described approach is illustrated with data from a cognitive intervention study, in which it is used to estimate time-invariant reliability in multiwave, multiple indicator models.

Reichardt: Commentary: Are three waves of data sufficient for assessing mediation? Reichardt-2011

Abstract: Maxwell, Cole, and Mitchell (2011) demonstrated that simple structural equation models, when used with cross-sectional data, generally produce biased estimates of meditated effects. I extend those results by showing how simple structural equation models can produce biased estimates of meditated effects when used even with longitudinal data. Even with longitudinal data, simple autoregressive structural equation models can imply the existence of indirect effects when only direct effects exist and the existence of direct effects when only indirect effects exist.

Reinert et al.: The alcohol use disorders identification test: An update of research findings Reinert-Allen-2007

Abstract: Background: The Alcohol Use Disorders Identification Test (AUDIT) has been extensively researched to determine its capability to accurately and practically screen for alcohol prob-

lems. Methods: During the 5 years since our previous review of the literature, a large number of additional studies have been published on the AUDIT, abbreviated versions of it, its psychometric properties, and the applicability of the AUDIT for a diverse array of populations. The current article summarizes new findings and integrates them with results of previous research. It also suggests some issues that we believe are particularly in need of further study. Results: A growing body of research evidence supports the criterion validity of English version of the AUDIT as a screen for alcohol dependence as well as for less severe alcohol problems. Nevertheless, the cut-points for effective detection of hazardous drinking as well as identification of alcohol dependence or harmful use in women need to be lowered from the originally recommended value of 8 points. The AUDIT-C, the most popular short version of the AUDIT consisting solely of its 3 consumption items, is approximately equal in accuracy to the full AUDIT. Psychometric properties of the AUDIT, such as test-retest reliability and internal consistency, are quite favorable. Continued research is urged to establish the psychometric properties of non-English versions of the AUDIT, use of the AUDIT with adolescents and with older adults, and selective inclusion of alcohol biomarkers with the AUDIT in some instances. Conclusions: Research continues to support use of the AUDIT as a means of screening for the spectrum of alcohol use disorders in various settings and with diverse populations.

Rhemtulla et al.: Worse than measurement error: Consequences of inappropriate latent variable measurement models   Rhemtulla-vanBork-Borsboom-2020

Abstract: Previous research and methodological advice has focused on the importance of accounting for measurement error in psychological data. That perspective assumes that psychological variables conform to a common factor model. We explore what happens when data that are not generated from a common factor model are nonetheless modeled as reflecting a common factor. Through a series of hypothetical examples and an empirical reanalysis, we show that when a common factor model is

misused, structural parameter estimates that indicate the relations among psychological constructs can be severely biased. Moreover, this bias can arise even when model fit is perfect. In some situations, composite models perform better than common factor models. These demonstrations point to a need for models to be justified on substantive, theoretical bases in addition to statistical ones.

Richards et al.: Correspondence between alcohol use measured by a wrist-worn alcohol biosensor and self-report via ecological momentary assessment over a 2-week period
Richards-Barnett-Cook-etal-2022

Abstract: Background: Transdermal alcohol biosensors measure alcohol use continuously, passively, and non-invasively. There is little field research on the Skyn biosensor, a new-generation, wrist-worn transdermal alcohol biosensor, and little evaluation of its sensitivity and specificity and the day-level correspondence between transdermal alcohol concentration (TAC) and number of self-reported drinks. Methods: Participants ($N = 36$; 61% male, $M_{age} = 34.3$) wore the Skyn biosensor and completed ecological momentary assessment (EMA) surveys about their alcohol use over 2 weeks. A total of 497 days of biosensor and EMA data were collected. Skyn-measured drinking episodes were defined by TAC ¿ $5\mu g/L$. Skyn data were compared to self-reported drinking to calculate sensitivity and specificity (for drinking day vs. nondrinking day). Generalized estimating equations models were used to evaluate the correspondence between TAC features (peak TAC and TAC-area under the curve (AUC)) and number of drinks. Individual-level factors (sex, age, race/ethnicity, body mass index, human immunodeficiency virus status, and hazardous drinking) were examined to explore associations with TAC controlling for number of drinks. Results: Using a minimum TAC threshold of $5\mu g/L$ plus coder review, the biosensor had sensitivity of 54.7%

and specificity of 94.6% for distinguishing drinking from nondrinking days. Without coder review, the sensitivity was 78.1% and the specificity was 55.2%. Peak TAC ($\beta = 0.92, p < 0.0001$) and TAC-AUC ($\beta = 1.60, p < 0.0001$) were significantly associated with number of drinks. Females had significantly higher TAC levels than males for the same number of drinks. Conclusions: Skyn-derived TAC can be used to measure alcohol use under naturalistic drinking conditions, additional research is needed to accurately identify drinking episodes based on Skyn TAC readings.

Richards et al.: Transdermal alcohol concentration features predict alcohol-induced blackouts in college students   Richards-Glenn-Turrisi-etal-2024

Abstract: Background: Alcohol-induced blackouts (AIBs) are common in college students. Individuals with AIBs also experience acute and chronic alcohol-related consequences. Research suggests that how students drink is an important predictor of AIBs. We used transdermal alcohol concentration (TAC) sensors to measure biomarkers of increasing alcohol intoxication (rise rate, peak, and rise duration) in a sample of college students. We hypothesized that the TAC biomarkers would be positively associated with AIBs. Methods: Students were eligible to participate if they were aged 18-22 years, in their second or third year of college, reported drinking 4+ drinks on a typical Friday or Saturday, experienced ≥ 1 AIB in the past semester, owned an iPhone, and were willing to wear a sensor for 3 days each weekend. Students ($N = 79$, 55.7% female, 86.1% White, $M_{age} = 20.1$) wore TAC sensors and completed daily diaries over four consecutive weekends (89.9% completion rate). AIBs were assessed using the Alcohol-Induced Blackout Measure-2. Logistic multilevel models were conducted to test for main effects. Results: Days with faster TAC rise rates (OR = 2.69, 95% CI: 1.56, 5.90), higher peak TACs (OR = 2.93, 95% CI: 1.64, 7.11), and longer rise TAC durations (OR = 4.16, 95% CI: 2.08, 10.62) were associated with greater odds of experiencing an AIB. Conclusions: In a sample of "risky" drinking college students, three TAC drinking features identified as

being related to rising intoxication independently predicted the risk for daily AIBs. Our findings suggest that considering how an individual drinks (assessed using TAC biomarkers), rather than quantity alone, is important for assessing risk and has implications for efforts to reduce risk. Not only is speed of intoxication important for predicting AIBs, but the height of the peak intoxication and the time spent reaching the peak are important predictors, each with different implications for prevention.

Richards et al.: Profiles of transdermal alcohol concentration and their prediction of negative and positive alcohol-related consequences in young adults' natural settings
Richards-Mallett-Turrisi-etal-2025

Abstract: Objective: Transdermal alcohol concentration (TAC) sensors provide a multidimensional characterization of drinking events that self-reports cannot. These profiles may differ in their associated day-level alcohol-related consequences, but no research has tested this. We address this using multilevel latent profile analysis. Method: Two hundred twenty-two young adults who regularly engage in heavy drinking (M_{age} = 22.3, 64% female, 79% non-Hispanic White) responded to surveys and wore TAC sensors for 6 consecutive days. We tested whether four previously identified TAC profiles: (1) high-fast (8.5% of days), (2) moderate-fast (12.8%), (3) low-slow (20.4%), and (4) little-to-no-drinking days (58.2%) differed in numbers of negative and positive consequences and in the odds that both consequence types occurred on the same day. Results: High-fast (incident rate ratio [$IRR_{low-slow}$] = 6.18; $IRR_{little-to-no-drinking}$ = 9.47) and moderate-fast ($IRR_{low-slow}$ = 3.71; $IRR_{little-to-no-drinking}$ = 5.68) days contained more negative consequences compared to low-slow and little-to-no-drinking days. High-fast (IRR = 2.05), moderate-fast (IRR = 1.88), and low-slow (IRR = 1.43) days contained more positive consequences than little-to-no-drinking days. The odds

of having only positive consequences were highest on low-slow, $\chi^2(3) = 9.10$, $p < .05$, days but the odds of experiencing both consequence types increased on moderate-fast and high-fast days, $\chi^2(3) = 39.63$, $p < .001$. Conclusions: Compared to little-to-no-drinking days, TAC profiles indicative of drinking (high-fast, moderate-fast, and low-slow) contained more negative and positive consequences. However, the odds of experiencing only positive consequences were highest among low-slow days and decreased on moderate-fast and high-fast days as the odds of negative consequences rose. These findings provide novel evidence reinforcing harm reduction approaches that seek to maximize positives and minimize negatives of alcohol consumption through emphasis on slow-paced, low-volume drinking.

Richards et al.: Subjective intoxication predicts alcohol-related consequences at equivalent alcohol concentrations in young adults using ecological momentary assessment and alcohol sensors — Richards-Turrisi-Russell-2024

Abstract: Objective: Subjective intoxication (SI) when drinking may serve as an internal barometer of whether to continue drinking or engage in potentially unsafe behavior. Mobile assessments offer the potential to use SI as a prospective risk indicator during drinking episodes; little evidence exists for the validity of real-time SI measures. We test the correspondence of SI with estimated blood alcohol concentration and transdermal alcohol concentration (TAC) in young adults' natural settings. We provide a novel test of whether SI features (peak and mean SI) uniquely predict consequences adjusting for alcohol concentration. Method: Two hundred twenty-two heavy-drinking young adults ($M_{age} = 22.3$, 64% female, 79% non-Hispanic White, 84% undergraduates) participated in a 6-day study that used ecological momentary assessment of drinking and TAC sensors. SI was assessed every 30 min during drinking episodes. Multilevel modeling was used to test hypotheses. Results:

Momentary SI and estimated blood alcohol concentration had moderate associations at the moment and day levels (standardized βs = 0.5-0.6); SI was moderately associated with TAC at the day level (βs = 0.5). Associations between SI and alcohol concentration varied widely between persons and across days. Day-level SI features predicted consequences when adjusting for alcohol concentration (incidence rate ratios, IRRs = 1.29-1.70). Conclusions: Our two-item SI measure shows evidence of validity in real-world settings with heavy-drinking young adults. SI was significantly correlated with alcohol concentration and was a unique predictor of consequences. The strength of these associations varied greatly across persons and days. Real-time SI measurement may be useful in preventive interventions, but continued research is needed into when and for whom momentary SI is most predictive of risk.

Roache et al.: Processing transdermal alcohol concentration (TAC) data to detect low-level drinking Roache-KarnsWright-Goros-etal-2019

Abstract: Background: Several studies have objectively quantified drinking through the use of Alcohol Monitoring System's (AMS) transdermal alcohol concentration (TAC) device known as SCRAM CAM. Criteria that AMS uses to detect drinking are known to be conservative and only reliably detect heavy drinking equivalent to 5 or more standard drinks. Our group has developed Research Rules used to process TAC data in a manner that will detect low-level and moderate drinking even though it is below the AMS criteria for detection. Methods: Sixteen male and 14 female paid research volunteers wore TAC monitors for 28 days in their natural environments and responded daily to text message prompts to self-report the previous day's drinking. Current analyses describe the Research Rules that we developed and how use of those rules impacts the detection of self-reported drinking treated as the standard in sensitivity/specificity analysis. Results: We observed 606 occurrences of positive TAC events over a total of 867 days and processed the TAC data to retain

345 as possible drinking events, even though AMS criteria confirmed drinking for only 163 of these events. The kinds of TAC events removed or retained by our rules are illustrated as cases of low and moderate drinking days that were detected by our rules but not by the conservative AMS criteria. AMS-confirmed TAC events have a high specificity (99.8%) to detect primarily heavy drinking, but have a poor sensitivity to detect lower-level drinking and a poor specificity as an indicator of alcohol abstinence. In contrast, our Research Rules detected 100% of TAC events detected by AMS but also detected 31% of the lower-level drinking events not detected by AMS, with 91% specificity. Conclusions: Reliance upon the AMS criteria for alcohol detection affords a high specificity for detection of heavy drinking but is a poor indicator of abstinence rates. In contrast, use of our Research Rules provides more sensitive means to quantify either any drinking or low-moderate levels of drinking while still maintaining good specificity.

Robey et al.: Type I error and the number of iterations in Monte Carlo studies of robustness   Robey-Barcikowski-1992

Abstract: A recent survey of simulation studies concluded that an overwhelming majority of papers do not report a rationale for the decision regarding the number of Monte Carlo iterations. A surprisingly large number of reports do not contain a justifiable definition of robustness and many studies are conducted with an insufficient number of iterations to achieve satisfactory statistical conclusion validity. The implication is that we do not follow our own advice regarding the management of Type I and Type II errors when conducting Monte Carlo experiments. This paper reports a straightforward application of a well-known procedure for the purpose of objectively determining the exact number of iterations necessary to confidently detect departures from robustness in Monte Carlo results. A table of the number of iterations necessary to detect departures from a series of nominal Type I error rates is included.

Michael E. Robinson and Joseph III L. Riley. "The role of emotion in pain". In: *Psychosocial factors in pain: Critical perspectives*. Ed. by Robert J. Gatchel and Dennis C. Turk. The Guilford Press, 1999, pp. 74–88.

Abstract: The purpose of this chapter is to review the role of negative emotion in the experience of pain. The authors focus their attention on the broad categories of depression, anxiety, and anger. They will also discuss several issues and controversies surrounding the role of negative emotion in pain. These include (1) the prevalence of negative emotion in patients with pain conditions, (2) the measurement of negative affect in pain conditions, (3) the role of negative emotion in disability and outcomes, (4) causal relationships between pain and negative affect, and (5) models incorporating negative emotion and pain.

Rogosa: Causal models in longitudinal research: Rationale, formulation, and interpretation   Rogosa-1979

David R. Rogosa. "Causal models in longitudinal research: Rationale, formulation, and interpretation". In: *Longitudinal methodology in the study of behavior and development*. Ed. by John R. Nesselroade and Paul B. Baltes. New York, NY: Academic Press, 1979. ISBN: 9780125156608.

Rogosa: A critique of cross-lagged correlation   Rogosa-1980

David R. Rogosa. "A critique of cross-lagged correlation". In: *Psychological Bulletin* 88.2 (Sept. 1980), pp. 245–258. ISSN: 0033-2909. DOI: 10.1037/0033-2909.88.2.245.

Abstract: Comments that cross-lagged correlation (CLC) is not a useful procedure for the analysis of longitudinal panel data. In particular, the difference between CLCs is not a sound basis for causal inference. Demonstrations of the failure of CLC are based mainly on results for the 2-wave, 2-variable longitudinal panel design. Extensions of these results to panels with multiple waves and multiple

measures reveal additional problems; each 2-wave snapshot did not yield dependable results. Taken together, the 2-wave analyses were often contradictory and misleading.

Rosseel: lavaan: An R package for structural equation modeling   Rosseel-2012

Yves Rosseel. "lavaan: An R package for structural equation modeling". In: *Journal of Statistical Software* 48.2 (2012). DOI: 10.18637/jss.v048.i02.

Abstract: Structural equation modeling (SEM) is a vast field and widely used by many applied researchers in the social and behavioral sciences. Over the years, many software packages for structural equation modeling have been developed, both free and commercial. However, perhaps the best state-of-the-art software packages in this field are still closed-source and/or commercial. The R package lavaan has been developed to provide applied researchers, teachers, and statisticians, a free, fully open-source, but commercial-quality package for latent variable modeling. This paper explains the aims behind the development of the package, gives an overview of its most important features, and provides some examples to illustrate how lavaan works in practice.

Rousselet et al.: The percentile bootstrap: A primer with step-by-step instructions in R   Rousselet-Pernet-Wilcox-2021

Guillaume A. Rousselet, Cyril R. Pernet, and Rand R. Wilcox. "The percentile bootstrap: A primer with step-by-step instructions in R". In: *Advances in Methods and Practices in Psychological Science* 4.1 (Jan. 2021), pp. 1–10. DOI: 10.1177/2515245920911881.

Abstract: The percentile bootstrap is the Swiss Army knife of statistics: It is a nonparametric method based on data-driven simulations. It can be applied to many statistical problems, as a substitute to standard parametric approaches, or in situations for which parametric methods do not exist. In this Tutorial, we cover R code to implement the percentile bootstrap to make inferences about central tendency (e.g., means and trimmed means) and spread in a one-sample example and in an example comparing two independent groups. For each example, we explain how to derive a

bootstrap distribution and how to get a confidence interval and a p value from that distribution. We also demonstrate how to run a simulation to assess the behavior of the bootstrap. For some purposes, such as making inferences about the mean, the bootstrap performs poorly. But for other purposes, it is the only known method that works well over a broad range of situations. More broadly, combining the percentile bootstrap with robust estimators (i.e., estimators that are not overly sensitive to outliers) can help users gain a deeper understanding of their data than they would using conventional methods.

Rubin: Inference and missing data   Rubin-1976

Abstract: When making sampling distribution inferences about the parameter of the data, θ, it is appropriate to ignore the process that causes missing data if the missing data are 'missing at random' and the observed data are 'observed at random', but these inferences are generally conditional on the observed pattern of missing data. When making direct-likelihood or Bayesian inferences about θ, it is appropriate to ignore the process that causes missing data if the missing data are missing at random and the parameter of the missing data process is 'distinct' from θ. These conditions are the weakest general conditions under which ignoring the process that causes missing data always leads to correct inferences.

Rubin: Multiple imputation for nonresponse in surveys   Rubin-1987

Abstract: Demonstrates how nonresponse in sample surveys and censuses can be handled by replacing each missing value with two or more multiple imputations. Clearly illustrates the advantages of modern computing to such handle surveys, and demonstrates the benefit of this statistical technique

for researchers who must analyze them. Also presents the background for Bayesian and frequentist theory. After establishing that only standard complete-data methods are needed to analyze a multiply-imputed set, the text evaluates procedures in general circumstances, outlining specific procedures for creating imputations in both the ignorable and nonignorable cases. Examples and exercises reinforce ideas, and the interplay of Bayesian and frequentist ideas presents a unified picture of modern statistics.

Russell et al.: Affect relative to day-level drinking initiation: Analyzing ecological momentary assessment data with multilevel spline modeling

Russell-LindenCarmichael-Lanza-etal-2020

Abstract: Affect regulation models state that affect both motivates and reinforces alcohol use. We aimed to examine whether affect levels and rates of change differed across drinking versus nondrinking days in a manner consistent with affect regulation models. Four hundred four regularly drinking adults, aged 18-70 years, completed ecological momentary assessments over 3 weeks. Participants provided positive affect (PA; enthusiastic, excited, happy) and negative affect (NA; distressed, sad) reports during all prompts; alcohol consumption reports were also provided. Multilevel spline models revealed that on drinking days, PA was higher and NA was lower both before and after drinking compared to matched times on nondrinking days. PA and NA were also higher and lower, respectively, both before and after drinking, when heavy drinking days were compared to moderate drinking days. Examination of affect rates of change revealed that (a) accelerating increases in PA and accelerating decreases in NA preceded drinking initiation, (b) PA increases and NA decreases

were seen up to 2 hr after drinking initiation, and (c) pre- and postdrinking PA increases were larger on heavy versus moderate drinking days, whereas only postdrinking NA decreases were larger on heavy drinking days. Results supported affect regulation models while adding nuance, showing accelerating changes in predrinking affect on drinking days and pre- and postdrinking differences in affect levels and rates of change across days of varying drinking intensity. Beyond theory, our results suggest that accelerating changes in affect may provide a clue to future commencement of heavy drinking, which may aid momentary intervention development.

Russell et al.: Adolescents' subjective social status predicts day-to-day mental health and future substance use   Russell-Odgers-2019

Abstract: Adolescents' subjective social status (SSS) is associated with mental and behavioral health outcomes, independent of socioeconomic status (SES). Many previous findings, however, come from cross-sectional studies. We report results from a longitudinal study with 151 adolescents identified as at risk for early substance use and behavioral problems sampled from low-SES neighborhoods. We examined whether adolescent's SSS predicted mental health (depression, anxiety, and inattention/impulsivity) measured over 30 days via ecological momentary assessment and risk for substance use at an 18-month follow-up. Results showed that with each perceived step "up" the SSS ladder, adolescents experienced fewer mental health symptoms in daily life and lower future substance use risk after adjusting for objective SES and previous psychopathology. Implications of these findings are discussed.

Abstract: Objective: Transdermal alcohol concentration (TAC) sensors capture aspects of drinking events that self-reports cannot. The multidimensional nature of TAC data allows novel classification of drinking days and identification of associated behavioral and contextual risks. We used multilevel latent profile analysis (MLPA) to create day-level profiles of TAC features and test their associations with (a) daily behaviors and contexts and (b) risk for alcohol use disorders at baseline. Method: Two hundred twenty-two regularly heavy-drinking young adults (Mage = 22.3) completed the Alcohol Use Disorders Identification Test (AUDIT) at baseline and then responded to mobile phone surveys and wore TAC sensors for six consecutive days. MLPA identified day-level profiles using four TAC features (peak, rise rate, fall rate, and duration). TAC profiles were tested as correlates of daily drinking behaviors, contexts, and baseline AUDIT. Results: Four profiles emerged: (a) high-fast (8.5% of days), (b) moderate-fast (12.8%), (c) low-slow (20.4%), and (d) little-to-no drinking days (58.2%). Profiles differed in the odds of risky drinking behaviors and contexts. The highest risk occurred on high-fast days, followed by moderate-fast, low-slow, and little-to-no drinking days. Higher baseline AUDIT predicted higher odds of high-fast and moderate-fast days. Conclusions: Days with high and fast intoxication are reflective of high-risk drinking behaviors and were most frequent among those at risk for alcohol use disorders. TAC research using MLPA may offer novel and important insights to intervention efforts.

Abstract: Objective: Test whether frequent protective behavioral strategies (PBS) users report (a) fewer alcohol-related consequences and (b) less risky alcohol intoxication dynamics (measured via transdermal alcohol concentration [TAC] sensor "features") in daily life. Method: Two hundred twenty-two frequently heavy-drinking young adults (M_{age} = 22.3 years) wore TAC sensors for 6 consecutive days. TAC features peak (maximum TAC), rise rate (speed of TAC increase), and area under the curve (AUC) were derived for each day. Negative alcohol-related consequences were measured in the morning after each self-reported drinking day. Past-year PBS use was measured at baseline. Results: Young adults reporting more frequent baseline PBS use showed (a) fewer alcohol-related consequences and (b) lower intoxication dynamics on average (less AUC, lower peaks, and slower rise rates). Limiting/stopping and manner of drinking PBS showed the same pattern of findings as the total score. Serious harm reduction PBS predicted fewer negative alcohol-related consequences, but not TAC features. Multilevel path models showed that TAC features peak and rise rate partially explained associations between PBS (total, limiting/stopping, and manner of drinking) and consequences. Independent contributions of PBS subscales were small and nonsignificant, suggesting that total PBS use was a more important predictor of risk/protection than the specific types of PBS used. Conclusions: Young adults using more total PBS may experience fewer alcohol-related consequences during real-world drinking episodes in part through less risky intoxication dynamics (TAC features). Future research measuring PBS at the daily level is needed to formally test TAC features as day-level mechanisms of protection from acute alcohol-related consequences.

Abstract: Background: Wearable transdermal alcohol concentration (TAC) sensors allow passive monitoring of alcohol concentration in natural settings and measurement of multiple features from drinking episodes, including peak intoxication level, speed of intoxication (absorption rate) and elimination, and duration. These passively collected features extend commonly used self-reported drink counts and may facilitate the prediction of alcohol-related consequences in natural settings, aiding risk stratification and prevention efforts. Method: A total of 222 young adults aged 21-29 ($M_{\text{age}} = 22.3$, 64 female, 79% non-Hispanic white, 84% undergraduates) who regularly drink heavily participated in a 5-day study that included the ecological momentary assessment (EMA) of alcohol consumption (daily morning reports and participant-initiated episodic EMA sequences) and the wearing of TAC sensors (SCRAM-CAM anklets). The analytic sample contained 218 participants and 1274 days (including 554 self-reported drinking days). Five features—area under the curve (AUC), peak TAC, rise rate (rate of absorption), fall rate (rate of elimination), and duration—were extracted from TAC-positive trajectories for each drinking day. Day- and person-level associations of TAC features with drink counts (morning and episodic EMA) and alcohol-related consequences were tested using multilevel modeling. Results: TAC features were strongly associated with morning drink reports ($r = 0.60.7$) but only moderately associated with episodic EMA drink counts ($r = 0.30.5$) at both day and person levels. Higher peaks, larger AUCs, faster rise rates, and faster fall rates were significantly predictive of day-level alcohol-related consequences after adjusting for both morning and episodic EMA drink counts in separate models. Person means of TAC features added little above daily scores to the prediction of alcohol-related consequences. Conclusions: These

results support the utility of TAC sensors in studies of alcohol misuse among young adults in natural settings and outline the specific TAC features that contribute to the day-level prediction of alcohol-related consequences. TAC sensors provide a passive option for obtaining valid and unique information predictive of drinking risk in natural settings.

Russell et al.: Witnessing substance use increases same-day antisocial behavior among at-risk adolescents: Gene-environment interaction in a 30-day ecological momentary assessment study   Russell-Wang-Odgers-2015

Abstract: Many young adolescents are embedded in neighborhoods, schools, and homes where alcohol and drugs are frequently used. However, little is known about (a) how witnessing others' substance use affects adolescents in their daily lives and (b) which adolescents will be most affected. The current study used ecological momentary assessment with 151 young adolescents (ages 11-15) to examine the daily association between witnessing substance use and antisocial behavior across 38 consecutive days. Results from multilevel logistic regression models indicated that adolescents were more likely to engage in antisocial behavior on days when they witnessed others using substances, an association that held when substance use was witnessed inside the home as well as outside the home (e.g., at school or in their neighborhoods). A significant Gene × Environment interaction suggested that the same-day association between witnessing substance use and antisocial behavior was significantly stronger among adolescents with, versus without, the dopamine receptor D4 seven repeat (DRD4-7R) allele. The implications of the findings for theory and research related to adolescent antisocial behavior are discussed.

Ryan et al.: Time to intervene: A continuous-time approach to network analysis and centrality **Ryan-Hamaker-2021**

Abstract: Network analysis of ESM data has become popular in clinical psychology. In this approach, discrete-time (DT) vector auto-regressive (VAR) models define the network structure with centrality measures used to identify intervention targets. However, VAR models suffer from time-interval dependency. Continuous-time (CT) models have been suggested as an alternative but require a conceptual shift, implying that DT-VAR parameters reflect total rather than direct effects. In this paper, we propose and illustrate a CT network approach using CT-VAR models. We define a new network representation and develop centrality measures which inform intervention targeting. This methodology is illustrated with an ESM dataset.

Ryan et al.: A continuous-time approach to intensive longitudinal data: What, why, and how? **Ryan-Kuiper-Hamaker-2018**

Abstract: The aim of this chapter is to (a) provide a broad didactical treatment of the first-order stochastic differential equation model—also known as the continuous-time (CT) first-order vector autoregressive (VAR(1)) model—and (b) argue for and illustrate the potential of this model for the study of psychological processes using intensive longitudinal data. We begin by describing what the CT-VAR(1) model is and how it relates to the more commonly used discrete-time VAR(1) model. Assuming no prior knowledge on the part of the reader, we introduce important concepts

for the analysis of dynamic systems, such as stability and fixed points. In addition we examine why applied researchers should take a continuous-time approach to psychological phenomena, focusing on both the practical and conceptual benefits of this approach. Finally, we elucidate how researchers can interpret CT models, describing the direct interpretation of CT model parameters as well as tools such as impulse response functions, vector fields, and lagged parameter plots. To illustrate this methodology, we reanalyze a single-subject experience-sampling dataset with the R package ctsem; for didactical purposes, R code for this analysis is included, and the dataset itself is publicly available.

Sacks et al.: 2010 national and state costs of excessive alcohol consumption

Sacks-Gonzales-Bouchery-etal-2015

Abstract: Introduction: Excessive alcohol use cost the U.S. $223.5 billion in 2006. Given economic shifts in the U.S. since 2006, current estimates are needed to help inform the planning of prevention strategies. Methods: From March 2012 to March 2014 attributable deaths) and price (e.g., inflation rate in cost of medical care). The total cost, cost to government, and costs for binge Excessive drinking cost the U.S. 249.0 billion in 2010, or about 2.05 per drink. Government paid for 100.7 billion (40.4%) of these costs. Binge drinking accounted for $191.1 billion (76.7\%) of costs$; underage drinking 24.3 billion (9.7%) of costs; and drinking while pregnant $5.5 billion (2.2\%) of costs. The median cost per state was 3.5 billion. Binge drinking was responsible for ¿70% of these costs in all states, and ¿40% of the binge drinking-related costs were paid by government. Conclusions: Excessive drinking cost the nation almost $250 billion in 2010. Two of every $5 of the total cost was paid by government, and three quarters of the costs were due to binge drinking. Several evidence-based strategies can help reduce excessive drinking and related costs, including increasing alcohol excise taxes, limiting alcohol outlet density, and commercial host liability.

Abstract: Background: To study the validity of transdermal assessment of alcohol concentration measured by a lightweight, noninvasive device. Methods: Subjects wore a 227-g anklet that sensed transdermal alcohol concentrations (TACs) every 15 to 30 minutes, downloading results to a remote computer each day. Twenty-four subjects entered a laboratory and received a dose of 0, 0.28, or 0.56 g/kg of ethanol. Breath alcohol concentrations (BrAC) and TAC were measured every 15 to 30 minutes Twenty others [10 alcohol dependent (AD) and 10 not (NAD)] in the community who wore the anklet for 8 days kept a drinking log and provided a BrAC sample each day. Results: In the laboratory, no zero-dose subject, and every subject receiving alcohol, had alcohol-positive TACs. The device distinguished low- and high-alcohol-dosing groups using peak ($t_{14} = 3.37$; $p < 0.01$) and area under the curve ($t_{14} = 3.42$; $p < 0.01$) of TACs. Within dosing groups, average TAC curves were broader (right-shifted) and had lower peaks than average BrAC curves. For community participants, self-reported number of drinks ($t_{18} = -3.77$; $p < 0.01$), area under the TAC curve ($t_{9.5} = -3.56$; $p < 0.01$), and mean TAC ($t_{9.9} = -3.35$; $p < 0.01$) all significantly distinguished the AD and NAD groups. However, individual transdermal readings were not reliably quantitatively equivalent to simultaneously obtained breath results. Conclusions: Within the limits of the laboratory study, the device consistently detected consumption of approximately 2 standard drinks. On average, the device shows discriminative validity as a semiquantitative measure of alcohol consumption but individual readings often are not equivalent to simultaneous BrACs.

SAMHSA: Key substance use and mental health indicators in the United States: Results from the 2019 National Survey on Drug Use and Health (HHS Publication No. PEP20-07-01-001, NSDUH Series H-55) SAMHSA-2020

SAMHSA. *Key substance use and mental health indicators in the United States: Results from the 2019 National Survey on Drug Use and Health (HHS Publication No. PEP20-07-01-001, NSDUH Series H-55)*. Rockville, MD: Center for Behavioral Health Statistics, Quality, Substance Abuse, and Mental Health Services Administration, 2020. URL: https://www.samhsa.gov/data/.

SAMHSA: Key substance use and mental health indicators in the United States: Results from the 2022 National Survey on Drug Use and Health (HHS Publication No. PEP23-07-01-006, NSDUH Series H-58) SAMHSA-2023

SAMHSA. *Key substance use and mental health indicators in the United States: Results from the 2022 National Survey on Drug Use and Health (HHS Publication No. PEP23-07-01-006, NSDUH Series H-58)*. Rockville, MD: Center for Behavioral Health Statistics, Quality, Substance Abuse, and Mental Health Services Administration, 2023. URL: https://www.samhsa.gov/data/report/2022-nsduh-annual-national-report.

Saunders et al.: Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO Collaborative Project on Early Detection of Persons with Harmful Alcohol Consumption-II Saunders-Assland-Babor-etal-1993

John B. Saunders et al. "Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO Collaborative Project on Early Detection of Persons with Harmful Alcohol Consumption-II". In: *Addiction* 88.6 (June 1993), pp. 791–804. ISSN: 1360-0443. DOI: 10.1111/j.1360-0443.1993.tb02093.x.

Abstract: The Alcohol Use Disorders Identification Test (AUDIT) has been developed from a six-country WHO collaborative project as a screening instrument for hazardous and harmful alcohol

consumption. It is a 10-item questionnaire which covers the domains of alcohol consumption, drinking behaviour, and alcohol-related problems. Questions were selected from a 150-item assessment schedule (which was administered to 1888 persons attending representative primary health care facilities) on the basis of their representativeness for these conceptual domains and their perceived usefulness for intervention. Responses to each question are scored from 0 to 4, giving a maximum possible score of 40. Among those diagnosed as having hazardous or harmful alcohol use, 92% had an AUDIT score of 8 or more, and 94% of those with non-hazardous consumption had a score of less than 8. AUDIT provides a simple method of early detection of hazardous and harmful alcohol use in primary health care settings and is the first instrument of its type to be derived on the basis of a cross-national study.

Savalei et al.: Computational options for standard errors and test statistics with incomplete normal and nonnormal data in SEM   Savalei-Rosseel-2021

Abstract: This article provides an overview of different computational options for inference following normal theory maximum likelihood (ML) estimation in structural equation modeling (SEM) with incomplete normal and nonnormal data. Complete data are covered as a special case. These computational options include whether the information matrix is observed or expected, whether the observed information matrix is estimated numerically or using an analytic asymptotic approximation, and whether the information matrix and the outer product matrix of the score vector are evaluated at the saturated or at the structured estimates. A variety of different standard errors and robust test statistics become possible by varying these options. We review the asymptotic properties of these computational variations, and we show how to obtain them using lavaan in R. We hope that this article will encourage methodologists to study the impact of the available computational options on the performance of standard errors and test statistics in SEM.

Abstract: The last two decades have seen enormous developments in statistical methods for incomplete data. The EM algorithm and its extensions, multiple imputation, and Markov Chain Monte Carlo provide a set of flexible and reliable tools from inference in large classes of missing-data problems. Yet, in practical terms, those developments have had surprisingly little impact on the way most data analysts handle missing values on a routine basis. Analysis of Incomplete Multivariate Data helps bridge the gap between theory and practice, making these missing-data tools accessible to a broad audience. It presents a unified, Bayesian approach to the analysis of incomplete multivariate data, covering datasets in which the variables are continuous, categorical, or both. The focus is applied, where necessary, to help readers thoroughly understand the statistical properties of those methods, and the behavior of the accompanying algorithms. All techniques are illustrated with real data examples, with extended discussion and practical advice. All of the algorithms described in this book have been implemented by the author for general use in the statistical languages S and S Plus. The software is available free of charge on the Internet.

Abstract: Statistical procedures for missing data have vastly improved, yet misconception and unsound practice still abound. The authors frame the missing-data problem, review methods, offer advice, and raise issues that remain unresolved. They clear up common misunderstandings regarding the missing at random (MAR) concept. They summarize the evidence against older procedures and, with few exceptions, discourage their use. They present, in both technical and practical language, 2

general approaches that come highly recommended: maximum likelihood (ML) and Bayesian multiple imputation (MI). Newer developments are discussed, including some for dealing with missing data that are not MAR. Although not yet in the mainstream, these procedures may eventually extend the ML and MI methods that currently represent the state of the art.

Schenker: Better bootstrap confidence intervals: Comment   Schenker-1987

Nathaniel Schenker. "Better bootstrap confidence intervals: Comment". In: *Journal of the American Statistical Association* 82.397 (Mar. 1987), p. 192. DOI: 10.2307/2289150.

Schermerhorn et al.: Developmental family processes and interparental conflict: Patterns of microlevel influences.   Schermerhorn-Chow-Cummings-2010

Alice C. Schermerhorn, Sy-Miin Chow, and E. Mark Cummings. "Developmental family processes and interparental conflict: Patterns of microlevel influences." In: *Developmental Psychology* 46.4 (2010), pp. 869–885. ISSN: 0012-1649. DOI: 10.1037/a0019662.

Abstract: Although there are frequent calls for the study of effects of children on families and mutual influence processes within families, little empirical progress has been made. We address these questions at the level of microprocesses during marital conflict, including children's influence on marital conflict and parents' influence on each other. Participants were 111 cohabiting couples with a child (55 male, 56 female) age 8-16 years. Data were drawn from parents' diary reports of interparental conflict over 15 days and were analyzed with dynamic systems modeling tools. Child emotions and behavior during conflicts were associated with interparental positivity, negativity, and resolution at the end of the same conflicts. For example, children's agentic behavior was associated with more marital conflict resolution, whereas child negativity was linked with more marital negativity. Regarding parents' influence on each other, among the findings, husbands' and wives' influence on themselves from one conflict to the next was indicated, and total number of conflicts predicted greater influence of wives' positivity on husbands' positivity. Contributions of these findings to the

understanding of developmental family processes are discussed, including implications for advanced understanding of interrelations between child and adult functioning and development.

Schouten et al.: Generating missing values for simulation purposes: A multivariate amputation procedure Schouten-Lugtig-Vink-2018

Abstract: Missing data form a ubiquitous problem in scientific research, especially since most statistical analyses require complete data. To evaluate the performance of methods dealing with missing data, researchers perform simulation studies. An important aspect of these studies is the generation of missing values in a simulated, complete data set: the amputation procedure. We investigated the methodological validity and statistical nature of both the current amputation practice and a newly developed and implemented multivariate amputation procedure. We found that the current way of practice may not be appropriate for the generation of intuitive and reliable missing data problems. The multivariate amputation procedure, on the other hand, generates reliable amputations and allows for a proper regulation of missing data problems. The procedure has additional features to generate any missing data scenario precisely as intended. Hence, the multivariate amputation procedure is an efficient method to accurately evaluate missing data methodology.

Schulenberg et al.: How and why the understanding of developmental continuity and discontinuity is important Schulenberg-Maggs-OMalley-2003

Abstract: No story is a straight line. The geometry of human life is too imperfect and too complex, too distorted by the laughter of time and the bewildering intricacies of fate to admit the straight line into its system of laws. (Pat Conroy, 1995, p. 104).

Schulenberg et al.: Getting drunk and growing up: trajectories of frequent binge drinking during the transition to young adulthood

Schulenberg-OMalley-Bachman-etal-1996

Abstract: Objective: The purpose of this study was: (1) to identify different trajectories of frequent binge drinking during the transition to young adulthood; (2) to validate the trajectories by relating them to behaviors and attitudes concerning alcohol and other drug use; and (3) to distinguish among the trajectories according to demographic characteristics and lifestyle experiences typical of the transition to young adulthood. Method: Four waves of national panel data were obtained from the Monitoring the Future project; 9,945 weighted cases from the 1976-85 high school senior year cohorts were surveyed at biennial intervals between ages 18 and 24. Frequent binge drinking was defined as having five or more drinks in a row at least twice in the past two weeks. Results: Six distinct frequent binge drinking trajectory groups were specified a priori and confirmed with cluster analysis: Never, Rare, Chronic, Decreased, Increased and "Fling." Repeated measures ANOVAS revealed that the trajectories corresponded to patterns of change and stability in problems with alcohol, attitudes about heavy drinking, peer heavy drinking and illicit drug use. Results from logistic regression analyses predicting diverging and converging trajectories provided some support for the general hypothesis that trajectories of Chronic and Increased frequent binge drinking over time are associated with difficulties in negotiating the transition to young adulthood. Conclusions: The findings provide strong evidence for wide developmental variation in drinking patterns in the population, variation that is obscured by more aggregate-level considerations. The developmental

variation in frequent binge drinking during the transition to young adulthood reflects systematic variation in success and difficulties with negotiating the transition.

Schulenberg et al.: Monitoring the Future national survey results on drug use, 1975-2020: Volume II, College students and adults ages 19-60

Schulenberg-Patrick-Johnston-etal-2021

John E. Schulenberg, Megan E. Patrick, et al. *Monitoring the Future national survey results on drug use, 1975-2020: Volume II, College students and adults ages 19-60.* Ann Arbor, MI: Institute for Social Research, The University of Michigan, 2021.

Schultzberg et al.: Number of subjects and time points needed for multilevel time-series analysis: A simulation study of dynamic structural equation modeling

Schultzberg-Muthen-2017

Mårten Schultzberg and Bengt Muthén. "Number of subjects and time points needed for multilevel time-series analysis: A simulation study of dynamic structural equation modeling". In: *Structural Equation Modeling: A Multidisciplinary Journal* 25.4 (Dec. 2017), pp. 495–515. ISSN: 1532-8007. DOI: 10.1080/10705511.2017.1392862.

Abstract: Dynamic structural equation modeling (DSEM) is a novel, intensive longitudinal data (ILD) analysis framework. DSEM models intraindividual changes over time on Level 1 and allows the parameters of these processes to vary across individuals on Level 2 using random effects. DSEM merges time series, structural equation, multilevel, and time-varying effects models. Despite the well-known properties of these analysis areas by themselves, it is unclear how their sample size requirements and recommendations transfer to the DSEM framework. This article presents the results of a simulation study that examines the estimation quality of univariate 2-level autoregressive models of order 1, AR(1), using Bayesian analysis in Mplus Version 8. Three features are varied in the simulations: complexity of the model, number of subjects, and number of time points per

subject. Samples with many subjects and few time points are shown to perform substantially better than samples with few subjects and many time points.

Schuurman et al.: How to compare cross-lagged associations in a multilevel autoregressive model   Schuurman-Ferrer-deBoerSonnenschein-etal-2016

Abstract: By modeling variables over time it is possible to investigate the Granger-causal cross-lagged associations between variables. By comparing the standardized cross-lagged coefficients, the relative strength of these associations can be evaluated in order to determine important driving forces in the dynamic system. The aim of this study was twofold: first, to illustrate the added value of a multilevel multivariate autoregressive modeling approach for investigating these associations over more traditional techniques; and second, to discuss how the coefficients of the multilevel autoregressive model should be standardized for comparing the strength of the cross-lagged associations. The hierarchical structure of multilevel multivariate autoregressive models complicates standardization, because subject-based statistics or group-based statistics can be used to standardize the coefficients, and each method may result in different conclusions. We argue that in order to make a meaningful comparison of the strength of the cross-lagged associations, the coefficients should be standardized within persons. We further illustrate the bivariate multilevel autoregressive model and the standardization of the coefficients, and we show that disregarding individual differences in dynamics can prove misleading, by means of an empirical example on experienced competence and exhaustion in persons diagnosed with burnout.

Abstract: An increasing number of researchers in psychology are collecting intensive longitudinal data in order to study psychological processes on an intraindividual level. An increasingly popular way to analyze these data is autoregressive time series modeling; either by modeling the repeated measures for a single individual using classic n = 1 autoregressive models, or by using multilevel extensions of these models, with the dynamics for each individual modeled at Level 1 and interindividual differences in these dynamics modeled at Level 2. However, while it is widely accepted in psychology that psychological measurements usually contain a certain amount of measurement error, the issue of measurement error is largely neglected in applied psychological (autoregressive) time series modeling: The regular autoregressive model incorporates innovations, or "dynamic errors," but not measurement error. In this article we discuss the concepts of reliability and measurement error in the context of dynamic (VAR(1)) models, and the consequences of disregarding measurement error variance in the data. For this purpose, we present a preliminary model that accounts for measurement error for constructs that are measured with a single indicator. We further discuss how this model could be used to investigate the between-person reliability of the measurements, as well as the (person-specific) within-person reliabilities and any individual differences in these reliabilities. We illustrate the consequences of assuming perfect reliability, the preliminary model, and reliabilities, using an empirical application in which we relate women's general positive affect to their positive affect concerning their romantic relationship.

Abstract: Measurement error is omnipresent in psychological data. However, the vast majority of applications of autoregressive time series analyses in psychology do not take measurement error into account. Disregarding measurement error when it is present in the data results in a bias of the autoregressive parameters. We discuss two models that take measurement error into account: An autoregressive model with a white noise term (AR+WN), and an autoregressive moving average (ARMA) model. In a simulation study we compare the parameter recovery performance of these models, and compare this performance for both a Bayesian and frequentist approach. We find that overall, the AR+WN model performs better. Furthermore, we find that for realistic (i.e., small) sample sizes, psychological research would benefit from a Bayesian approach in fitting these models. Finally, we illustrate the effect of disregarding measurement error in an AR(1) model by means of an empirical application on mood data in women. We find that, depending on the person, approximately 30-50% of the total variance was due to measurement error, and that disregarding this measurement error results in a substantial underestimation of the autoregressive parameters.

Selig et al.: Mediation models for longitudinal data in developmental research

Abstract: Mediation models are used to describe the mechanism(s) by which one variable influences another. These models can be useful in developmental research to explicate the relationship between

variables, developmental processes, or combinations of variables and processes. In this article we describe aspects of mediation effects specific to developmental research. We focus on three central issues in longitudinal mediation models: the theory of change for variables in the model, the role of time in the model, and the types of indirect effects in the model. We use these themes as we describe three different models for examining mediation in longitudinal data.

Serlin: Testing for robustness in Monte Carlo studies Serlin-2000

Abstract: Monte Carlo studies provide the information needed to help researchers select appropriate analytical procedures under design conditions in which the underlying assumptions of the procedures are not met. In Monte Carlo studies, the 2 errors that one could commit involve (a) concluding that a statistical procedure is robust when it is not or (b) concluding that it is not robust when it is. In previous attempts to apply standard statistical design principles to Monte Carlo studies, the less severe of these errors has been wrongly designated the Type I error. In this article, a method is presented for controlling the appropriate Type I error rate; the determination of the number of iterations required in a Monte Carlo study to achieve desired power is described; and a confidence interval for a test's true Type I error rate is derived. A robustness criterion is also proposed that is a compromise between W. G. Cochran's (1952) and J. V. Bradley's (1978) criteria.

Serlin et al.: Rationality in psychological research: The good-enough principle Serlin-Lapsley-1985

Abstract: Reexamines methodological and procedural issues raised by P. Meehl (1967; see also PA, Vol 62:5042) that question the rationality of psychological inquiry. Issues concern the asymmetry

in theory testing between psychology and physics and the slow progress observed in psychological research. A good-enough principle is proposed to resolve Meehl's methodological paradox, and a more powerful reconstruction of science developed by I. Lakatos (1978) is suggested to account for the actual practice of psychological researchers.

Shao: Impact of the bootstrap on sample surveys — Shao-2003

Abstract: This article discusses the impact of the bootstrap on sample surveys and introduces some of the main developments of the bootstrap methodology for sample surveys in the last twenty five years.

Shapiro et al.: On the treatment of correlation structures as covariance structures — Shapiro-Browne-1990

Abstract: Necessary and sufficient conditions are provided for minimum discrepancy methods, intended for covariance structures, to retain their asymptotic properties in the analysis of correlation structures. Examples of correlation structures satisfying these conditions are considered, and alternative discrepancy functions, which are always appropriate for correlation structures under normality assumptions, are discussed.

Abstract: Background: Pediatric chronic pain is prevalent and disabling. The present study aimed to assess the prevalence of chronic pain among adolescents in Shiraz, Iran. We also compared emotional intelligence, self-esteem and parenting style between adolescents with chronic pain and healthy adolescents. Finally, we examined the predicting role of these variables regarding chronic pain in adolescents. Methods: This cross-sectional study, from January to June 2018, was conducted on 734 adolescents in Shiraz. A clustering sampling method was used. Adolescents with chronic pain were identified by affirmative answers to screening questions based on the International Classification of Diseases 11th Revision (ICD-11) criteria. Participants completed three validated self-report questionnaires: Trait Emotional Intelligence Questionnaire, Rosenberg self-esteem scale and Baumrind parenting style questionnaire. The data were analyzed through SPSS v.22 software using Mann-Whitney and binary logistic regression tests. P¡0.05 was considered significant. Results: There were 221(30.1%) adolescents who met the ICD-11 criteria of chronic pain. Mann-Whitney tests showed that emotional intelligence (P¡0.001), self-esteem (P¡0.001), authoritative parenting style (P=0.004), and authoritarian parenting style (P=0.006) were significantly different in adolescents with chronic pain compared to healthy adolescents. Binary logistic regression revealed that emotional intelligence (P¡0.001), self-esteem (P¡0.001), authoritarian parenting style (P=0.04) and authoritative parenting style (P=0.01) were significantly correlated with chronic pain after controlling for demographic variables. Conclusion: Our findings indicate that emotional intelligence, self-esteem and parenting styles could be important factors in development or maintenance of chronic pain in adolescents. The results have potential to be extended to future interventions for adolescents with chronic pain.

Saul Shiffman. "Ecological momentary assessment (EMA) in studies of substance use". In: *Psychological Assessment* 21.4 (Dec. 2009), pp. 486–497. DOI: 10.1037/a0017074.

Abstract: Ecological momentary assessment (EMA) is particularly suitable for studying substance use, because use is episodic and thought to be related to mood and context. This article reviews EMA methods in substance use research, focusing on tobacco and alcohol use and relapse, where EMA has been most applied. Common EMA designs combine event-based reports of substance use with time-based assessments. Approaches to data organization and analysis have been very diverse, particularly regarding their treatment of time. Compliance with signaled assessments is often high. Compliance with recording of substance use appears good but is harder to validate. Treatment applications of EMA are emerging. EMA captures substance use patterns not measured by questionnaires or retrospective data and holds promise for substance use research.

Saul Shiffman, Arthur A. Stone, and Michael R. Hufford. "Ecological momentary assessment". In: *Annual Review of Clinical Psychology* 4.1 (Apr. 2008), pp. 1–32. DOI: 10.1146/annurev.clinpsy.3.022806.091415.

Abstract: Assessment in clinical psychology typically relies on global retrospective self-reports collected at research or clinic visits, which are limited by recall bias and are not well suited to address how behavior changes over time and across contexts. Ecological momentary assessment (EMA) involves repeated sampling of subjects' current behaviors and experiences in real time, in subjects' natural environments. EMA aims to minimize recall bias, maximize ecological validity, and allow study of microprocesses that influence behavior in real-world contexts. EMA studies assess particular events in subjects' lives or assess subjects at periodic intervals, often by random time sampling,

using technologies ranging from written diaries and telephones to electronic diaries and physiological sensors. We discuss the rationale for EMA, EMA designs, methodological and practical issues, and comparisons of EMA and recall data. EMA holds unique promise to advance the science and practice of clinical psychology by shedding light on the dynamics of behavior in real-world settings.

Shrout: Commentary: Mediation analysis, causal process, and cross-sectional data
Shrout-2011

Abstract: Maxwell, Cole, and Mitchell (2011) extended the work of Maxwell and Cole (2007), which raised important questions about whether mediation analyses based on cross-sectional data can shed light on longitudinal mediation process. The latest article considers longitudinal processes that can only be partially explained by an intervening variable, and Maxwell et al. showed that the same general conclusions are obtained, namely that analyses of cross-sectional data will not reveal the longitudinal mediation process. While applauding the advances of the target article, this comment encourages the detailed exploration of alternate causal models in psychology beyond the autoregressive model considered by Maxwell et al. When inferences based on cross-sectional analyses are compared to alternate models, different patterns of bias are likely to be observed. I illustrate how different models of the causal process can be derived using examples from research on psychopathology.

Shrout et al.: Mediation in experimental and nonexperimental studies: New procedures and recommendations
Shrout-Bolger-2002

Patrick E. Shrout and Niall Bolger. "Mediation in experimental and nonexperimental studies: New

procedures and recommendations". In: *Psychological Methods* 7.4 (2002), pp. 422–445. DOI: `10.1037/1082-989x.7.4.422`.

Abstract: Mediation is said to occur when a causal effect of some variable X on an outcome Y is explained by some intervening variable M. The authors recommend that with small to moderate samples, bootstrap methods (B. Efron & R. Tibshirani, 1993) be used to assess mediation. Bootstrap tests are powerful because they detect that the sampling distribution of the mediated effect is skewed away from 0. They argue that R. M. Baron and D. A. Kenny's (1986) recommendation of first testing the $X \rightarrow Y$ association for statistical significance should not be a requirement when there is a priori belief that the effect size is small or suppression is a possibility. Empirical examples and computer setups for bootstrap analyses are provided.

Shumway et al.: Time series analysis and its applications: With R examples

Shumway-Stoffer-2017

Robert H. Shumway and David S. Stoffer. *Time series analysis and its applications: With R examples.* Springer International Publishing, 2017. ISBN: 978-3-319-52452-8. DOI: `10.1007/978-3-319-52452-8`.

Abstract: The fourth edition of this popular graduate textbook, like its predecessors, presents a balanced and comprehensive treatment of both time and frequency domain methods with accompanying theory. Numerous examples using nontrivial data illustrate solutions to problems such as discovering natural and anthropogenic climate change, evaluating pain perception experiments using functional magnetic resonance imaging, and monitoring a nuclear test ban treaty. The book is designed as a textbook for graduate level students in the physical, biological, and social sciences and as a graduate level text in statistics. Some parts may also serve as an undergraduate introductory course. Theory and methodology are separated to allow presentations on different levels. In addition to coverage of classical methods of time series regression, ARIMA models, spectral analysis and state-space models, the text includes modern developments including categorical time series

analysis, multivariate spectral methods, long memory series, nonlinear models, resampling techniques, GARCH models, ARMAX models, stochastic volatility, wavelets, and Markov chain Monte Carlo integration methods. This edition includes R code for each numerical example in addition to Appendix R, which provides a reference for the data sets and R scripts used in the text in addition to a tutorial on basic R commands and R time series. An additional file is available on the book's website for download, making all the data sets and scripts easy to load into R.

Simons et al.: Event-level associations between affect, alcohol intoxication, and acute depenence symptoms: Effects of urgency, self-control, and drinking experience

Simons-Dvorak-Batien-etal-2010

Abstract: This study used experience sampling to examine within-person associations between positive affect, anxiety, sadness, and hostility and two outcomes: alcohol intoxication and acute dependence symptoms. We examined the role of urgency, premeditation, and perseverance in predicting the alcohol outcomes and tested whether the affective associations varied as a function of urgency. Participants completed baseline assessments and 21 days of experience sampling on PDAs. Hypotheses were partially confirmed. Positive affect was positively, and sadness inversely, associated with intoxication. Hostility was associated with intoxication for men but not women. Negative urgency moderated the association between anxiety and intoxication, making it stronger. However, positive urgency did not moderate the effect of positive affect. Heavier drinkers exhibited the greatest number of symptoms, yet the association between intoxication and acute signs of alcohol

disorder was attenuated among these individuals. Results support the use of experience sampling to study acute signs and symptoms of high risk drinking and dependence.

Abstract: Research on alcohol use depends heavily on the validity of self-reported drinking. The present paper presents data from 647 days of self-monitoring with a transdermal alcohol sensor by 60 young adults. We utilized a biochemical measure, transdermal alcohol assessment with the WrisTAS, to examine the convergent validity of three approaches to collecting daily self-report drinking data: experience sampling, daily morning reports of the previous night, and 1-week timeline follow-back (TLFB) assessments. We tested associations between three pharmacokinetic indices (peak concentration, area under the curve (AUC), and time to reach peak concentration) derived from the transdermal alcohol signal and within- and between- person variation in alcohol dependence symptoms. The WrisTAS data corroborated 85.74% of self-reported drinking days based on the experience sampling data. The TLFB assessment and combined experience sampling and morning reports agreed on 87.27% of drinking days. Drinks per drinking day did not vary as a function of wearing or not wearing the sensor; this indicates that participants provided consistent reports of their drinking regardless of biochemical verification. In respect to self-reported alcohol dependence symptoms, the AUC of the WrisTAS alcohol signal was associated with dependence symptoms at both the within- and between- person level. Furthermore, alcohol dependence symptoms at baseline predicted drinking episodes characterized in biochemical data by both higher peak alcohol concentration and faster time to reach peak concentration. The results support the validity of self-report alcohol data, provide empirical data useful for optimal design of daily process sampling,

and provide an initial demonstration of the use of transdermal alcohol assessment to characterize drinking dynamics associated with risk for alcohol dependence.

Singer: SEM modeling with singular moment matrices part II: ML-estimation of sampled stochastic differential equations   Singer-2012

Abstract: Linear stochastic differential equations are expressed as an exact discrete model (EDM) and estimated with structural equation models (SEMs) and the Kalman filter (KF) algorithm. The oversampling approach is introduced in order to formulate the EDM on a time grid which is finer than the sampling intervals. This leads to a simple computation of the nonlinear parameter functionals of the EDM. For small discretization intervals, the functionals can be linearized, and standard software permitting only linear parameter restrictions can be used. However, in this case the SEM approach must handle large matrices leading to degraded performance and possible numerical problems. The methods are compared using coupled linear random oscillators with time-varying parameters and irregular sampling times.

Singh: On the asymptotic accuracy of Efron's bootstrap   Singh-1981

Abstract: In the non-lattice case it is shown that the bootstrap approximation of the distribution of the standardized sample mean is asymptotically more accurate than approximation by the limiting normal distribution. The exact convergence rate of the bootstrap approximation of the distributions of sample quantiles is obtained. A few other convergence rates regarding the bootstrap method are also studied.

Abstract: Substance dependence is characterized by compulsive drug-taking despite negative consequences. Animal research suggests an underlying imbalance between goal-directed and habitual action control with chronic drug use. However, this imbalance, and its associated neurophysiological mechanisms, has not yet been experimentally investigated in human drug abusers. The aim of the present study therefore was to assess the balance between goal-directed and habit-based learning and its neural correlates in abstinent alcohol-dependent (AD) patients. A total of 31 AD patients and 19 age, gender and education matched healthy controls (HC) underwent functional magnetic resonance imaging (fMRI) during completion of an instrumental learning task designed to study the balance between goal-directed and habit learning. Task performance and task-related blood oxygen level-dependent activations in the brain were compared between AD patients and healthy matched controls. Findings were additionally associated with duration and severity of alcohol dependence. The results of this study provide evidence for an overreliance on stimulus-response habit learning in AD compared with HC, which was accompanied by decreased engagement of brain areas implicated in goal-directed action (ventromedial prefrontal cortex and anterior putamen) and increased recruitment of brain areas implicated in habit learning (posterior putamen) in AD patients. In conclusion, this is the first human study to provide experimental evidence for a disturbed balance between goal-directed and habitual control by use of an instrumental learning task, and to directly implicate cortical dysfunction to overreliance on inflexible habits in AD patients.

Abstract: Purpose of Review: Ambulatory assessment methods, including ecological momentary assessment (EMA), have often been used in eating disorders (EDs) to assess the type, frequency, and temporal sequencing of ED symptoms occurring in naturalistic environments. Relatedly, growing research in EDs has explored the utility of ecological momentary interventions (EMIs) to target ED symptoms. The aims of the present review were to (1) synthesize recent literature pertaining to ambulatory assessment/EMA and EMI in EDs, and (2) identify relevant limitations and future directions in these domains. Recent Findings: With respect to ambulatory assessment and EMA, there has been substantial growth in the expansion of constructs assessed with EMA, the exploration of state- vs. trait-level processes, integration of objective and passive assessment approaches, and consideration of methodological issues. The EMI literature in EDs also continues to grow, though most of the recent research focuses on mobile health (mHealth) technologies with relatively minimal EMI components that adapt to momentary contextual information. Summary: Despite these encouraging advances, there remain several promising areas of ambulatory assessment research and clinical applications in EDs going forward. These include integration of passive data collection, use of EMA in treatment evaluation and design, evaluation of dynamic system processes, inclusion of diverse samples, and development and evaluation of adaptive, tailored EMIs such as just-in-time adaptive interventions. While much remains to be learned in each of these domains, the continual growth in mobile technology has potential to facilitate and refine our understanding of the nature of ED psychopathology and ultimately improve intervention approaches.

Sobel: Asymptotic confidence intervals for indirect effects in structural equation models   **Sobel-1982**

Michael E. Sobel. "Asymptotic confidence intervals for indirect effects in structural equation models". In: *Sociological Methodology* 13 (1982), p. 290. DOI: 10.2307/270723.

Sobel: Some new results on indirect effects and their standard errors in covariance structure models   **Sobel-1986**

Michael E. Sobel. "Some new results on indirect effects and their standard errors in covariance structure models". In: *Sociological Methodology* 16 (1986), p. 159. DOI: 10.2307/270922.

Sobel: Direct and indirect effects in linear structural equation models   **Sobel-1987**

Michael E. Sobel. "Direct and indirect effects in linear structural equation models". In: *Sociological Methods & Research* 16.1 (Aug. 1987), pp. 155–176. DOI: 10.1177/0049124187016001006.

Abstract: This article discusses total indirect effects in linear structural equation models. First, I define these effects. Second, I show how the delta method may be used to obtain the standard errors of the sample estimates of these effects and test hypotheses about the magnitudes of the indirect effects. To keep matters simple, I focus throughout on a particularly simple linear structural equation system; for a treatment of the general case, see Sobel (1986). To illustrate the ideas and results, a detailed example is presented.

P. S. Soltis et al.: Applying the Bootstrap in Phylogeny Reconstruction

Soltis-Soltis-2003

Pamela S. Soltis and Douglas E. Soltis. "Applying the Bootstrap in Phylogeny Reconstruction". In: *Statistical Science* 18.2 (May 2003). DOI: 10.1214/ss/1063994980.

Abstract: With the increasing emphasis in biology on reconstruction of phylogenetic trees, questions have arisen as to how confident one should be in a given phylogenetic tree and how support for phylogenetic trees should be measured. Felsenstein suggested that bootstrapping be applied across characters of a taxon-by-character data matrix to produce replicate "bootstrap data sets," each of which is then analyzed phylogenetically, with a consensus tree constructed to summarize the results of all replicates. The proportion of trees/replicates in which a grouping is recovered is presented as a measure of support for that group. Bootstrapping has become a common feature of phylogenetic analysis. However, the interpretation of bootstrap values remains open to discussion, and phylogeneticists have used these values in multiple ways. The usefulness of phylogenetic bootstrapping is potentially limited by a number of features, such as the size of the data matrix and the underlying assumptions of the phylogeny reconstruction program. Recent studies have explored the application of bootstrapping to large data sets and the relative performance of bootstrapping and jackknifing.

Song et al.: Bayesian estimation of random coefficient dynamic factor models

Song-Ferrer-2012

Abstract: Dynamic factor models (DFMs) have typically been applied to multivariate time series data collected from a single unit of study, such as a single individual or dyad. The goal of DFMs application is to capture dynamics of multivariate systems. When multiple units are available, however, DFMs are not suited to capture variations in dynamics across units. The aims of this study are (a) to propose a random coefficient DFM (RC-DFM) to statistically model variations of dynamics across multiple units using the Bayesian method, (b) to illustrate the use of the proposed procedure by applying RC-DFMs to affect data collected from multiple dyads in romantic relationships, and (c) to evaluate the performance of the RC-DFMs with Bayesian estimation through simulation analyses. The results from the simulation analyses show that the Bayesian estimation of RC-DFMs

works well in recovering parameters including both fixed and random effects. A number of practical considerations are provided to guide future research on using Bayesian methods for estimating multivariate time series from multiple units.

Squeglia et al.: The influence of substance use on adolescent brain development
Squeglia-Jacobus-Tapert-2009

Abstract: Adolescence is a unique period in neurodevelopment. Alcohol and marijuana use are common. Recent research has indicated that adolescent substance users show abnormalities on measures of brain functioning, which is linked to changes in neurocognition over time. Abnormalities have been seen in brain structure volume, white matter quality, and activation to cognitive tasks, even in youth with as little as 1-2 years of heavy drinking and consumption levels of 20 drinks per month, especially if ¿4-5 drinks are consumed on a single occasion. Heavy marijuana users show some subtle anomalies too, but generally not the same degree of divergence from demographically similar non-using adolescents. This article reviews the extant literature on neurocognition, brain structure, and brain function in adolescent substance users with an emphasis on the most commonly used substances, and in the context of ongoing neuromaturational processes. Methodological and treatment implications are provided.

Stattin et al.: Parental monitoring: A reinterpretation
Stattin-Kerr-2000

Abstract: Monitoring (tracking and surveillance) of children's behavior is considered an essential parenting skill. Numerous studies show that well-monitored youths are less involved in delinquency

and other normbreaking behaviors, and scholars conclude that parents should track their children more carefully. This study questions that conclusion. We point out that monitoring measures typically assess parents' knowledge but not its source, and parents could get knowledge from their children's free disclosure of information as well as their own active surveillance efforts. In our study of 703 14-year-olds in central Sweden and their parents, parental knowledge came mainly from child disclosure, and child disclosure was the source of knowledge that was most closely linked to broad and narrow measures of delinquency (normbreaking and police contact). These results held for both children's and parents' reports, for both sexes, and were independent of whether the children were exhibiting problem behavior or not. We conclude that tracking and surveillance is not the best prescription for parental behavior and that a new prescription must rest on an understanding of the factors that determine child disclosure.

Staudenmayer et al.: Measurement error in linear autoregressive models

Abstract: Time series data are often subject to measurement error, usually the result of needing to estimate the variable of interest. Although it is often reasonable to assume that the measurement error is additive (i.e., the estimator is conditionally unbiased for the missing true value), the measurement error variances often vary as a result of changes in the population/process over time and/or changes in sampling effort. In this article we address estimation of the parameters in linear autoregressive models in the presence of additive and uncorrelated measurement errors, allowing heteroscedasticity in the measurement error variances. We establish the asymptotic properties of naive estimators that ignore measurement error and propose an estimator based on correcting the Yule-Walker estimating equations. We also examine a pseudo-likelihood method based on normality assumptions and computed using the Kalman filter. We review other techniques that have been

proposed, including two that require no information about the measurement error variances, and compare the various estimators both theoretically and via simulations. The estimator based on corrected estimating equations is easy to obtain and readily accommodates (and is robust to) unequal measurement error variances. Asymptotic calculations and finite-sample simulations show that it is often relatively efficient.

Stoffer et al.: Bootstrapping state-space models: Gaussian maximum likelihood estimation and the Kalman filter   Stoffer-Wall-1991

Abstract: The bootstrap is proposed as a method for assessing the precision of Gaussian maximum likelihood estimates of the parameters of linear state-space models. Our results also apply to autoregressive moving average models, since they are a special case of state-space models. It is shown that for a time-invariant, stable system, the bootstrap applied to the innovations yields asymptotically consistent standard errors. To investigate the performance of the bootstrap for finite sample lengths, simulation results are presented for a two-state model with 50 and 100 observations; two cases are investigated, one with real characteristic roots and one with complex characteristic roots. The bootstrap is then applied to two real data sets, one used in a test for efficient capital markets and one used to develop an autoregressive integrated moving average model for quarterly earnings data. We find the bootstrap to be of definite value over the conventional asymptotics.

Swift: Transdermal alcohol measurement for estimation of blood alcohol concentration
$\hfill$ **Swift-2000**

Robert Swift. "Transdermal alcohol measurement for estimation of blood alcohol concentration". In: *Alcoholism: Clinical and Experimental Research* 24.4 (Apr. 2000), pp. 422–423. ISSN: 1530-0277. DOI: 10.1111/j.1530-0277.2000.tb02006.x.

Tange: GNU Parallel 20210922 ('Vindelev') [stable] $\hfill$ **Tange-2021**

Ole Tange. *GNU Parallel 20210922 ('Vindelev') [stable]*. 2021. DOI: 10.5281/ZENODO.5523272.

Tange: GNU Parallel 20241222 ('Bashar') [stable] $\hfill$ **Tange-2024**

Ole Tange. *GNU Parallel 20241222 ('Bashar') [stable]*. 2024. DOI: 10.5281/ZENODO.14550073.

A. B. Taylor et al.: Four applications of permutation methods to testing a single-mediator model $\hfill$ **Taylor-MacKinnon-2012**

Aaron B. Taylor and David P. MacKinnon. "Four applications of permutation methods to testing a single-mediator model". In: *Behavior Research Methods* 44.3 (Feb. 2012), pp. 806–844. DOI: 10.3758/s13428-011-0181-x.

Abstract: Four applications of permutation tests to the single-mediator model are described and evaluated in this study. Permutation tests work by rearranging data in many possible ways in order to estimate the sampling distribution for the test statistic. The four applications to mediation evaluated here are the permutation test of ab, the permutation joint significance test, and the noniterative and iterative permutation confidence intervals for ab. A Monte Carlo simulation study was used to compare these four tests with the four best available tests for mediation found in previous research: the joint significance test, the distribution of the product test, and the percentile and bias-corrected bootstrap tests. We compared the different methods on Type I error, power, and

confidence interval coverage. The noniterative permutation confidence interval for ab was the best performer among the new methods. It successfully controlled Type I error, had power nearly as good as the most powerful existing methods, and had better coverage than any existing method. The iterative permutation confidence interval for ab had lower power than do some existing methods, but it performed better than any other method in terms of coverage. The permutation confidence interval methods are recommended when estimating a confidence interval is a primary concern. SPSS and SAS macros that estimate these confidence intervals are provided.

A. B. Taylor et al.: Tests of the three-path mediated effect

<div align="right">

Taylor-MacKinnon-Tein-2007

</div>

Aaron B. Taylor, David P. MacKinnon, and Jenn-Yun Tein. "Tests of the three-path mediated effect". In: *Organizational Research Methods* 11.2 (July 2007), pp. 241–269. DOI: 10.1177/1094428107300344.

Abstract: In a three-path mediational model, two mediators intervene in a series between an independent and a dependent variable. Methods of testing for mediation in such a model are generalized from the more often used single-mediator model. Six such methods are introduced and compared in a Monte Carlo study in terms of their Type I error, power, and coverage. Based on its results, the joint significance test is preferred when only a hypothesis test is of interest. The percentile bootstrap and bias-corrected bootstrap are preferred when a confidence interval on the mediated effect is desired, with the latter having more power but also slightly inflated Type I error in some conditions.

R. Tibshirani: Regression shrinkage and selection via the lasso Tibshirani-1996

Robert Tibshirani. "Regression shrinkage and selection via the lasso". In: *Journal of the Royal Statistical Society Series B: Statistical Methodology* 58.1 (Jan. 1996), pp. 267–288. ISSN: 1467-9868. DOI: 10.1111/j.2517-6161.1996.tb02080.x.

Abstract: We propose a new method for estimation in linear models. The 'lasso' minimizes the residual sum of squares subject to the sum of the absolute value of the coefficients being less than a constant. Because of the nature of this constraint it tends to produce some coefficients that are exactly 0 and hence gives interpretable models. Our simulation studies suggest that the lasso enjoys some of the favourable properties of both subset selection and ridge regression. It produces interpretable models like subset selection and exhibits the stability of ridge regression. There is also an interesting relationship with recent work in adaptive function estimation by Donoho and Johnstone. The lasso idea is quite general and can be applied in a variety of statistical models: extensions to generalized regression models and tree-based models are briefly described.

R. Tibshirani: Regression shrinkage and selection via the lasso: A retrospective

Tibshirani-2011

Abstract: In the paper I give a brief review of the basic idea and some history and then discuss some developments since the original paper on regression shrinkage and selection via the lasso.

Tofighi et al.: Indirect effects in sequential mediation models: Evaluating methods for hypothesis testing and confidence interval formation

Tofighi-Kelley-2019

Abstract: Complex mediation models, such as a two-mediator sequential model, have become more prevalent in the literature. To test an indirect effect in a two-mediator model, we conducted a large-scale Monte Carlo simulation study of the Type I error, statistical power, and confidence interval

coverage rates of 10 frequentist and Bayesian confidence/credible intervals (CIs) for normally and nonnormally distributed data. The simulation included never-studied methods and conditions (e.g., Bayesian CI with flat and weakly informative prior methods, two model-based bootstrap methods, and two nonnormality conditions) as well as understudied methods (e.g., profile-likelihood, Monte Carlo with maximum likelihood standard error [MC-ML] and robust standard error [MC-Robust]). The popular BC bootstrap showed inflated Type I error rates and CI under-coverage. We recommend different methods depending on the purpose of the analysis. For testing the null hypothesis of no mediation, we recommend MC-ML, profile-likelihood, and two Bayesian methods. To report a CI, if data has a multivariate normal distribution, we recommend MC-ML, profile-likelihood, and the two Bayesian methods; otherwise, for multivariate nonnormal data we recommend the percentile bootstrap. We argue that the best method for testing hypotheses is not necessarily the best method for CI construction, which is consistent with the findings we present.

Tofighi et al.: Improved inference in mediation analysis: Introducing the model-based constrained optimization procedure   Tofighi-Kelley-2020

Abstract: Mediation analysis is an important approach for investigating causal pathways. One approach used in mediation analysis is the test of an indirect effect, which seeks to measure how the effect of an independent variable impacts an outcome variable through one or more mediators. However, in many situations the proposed tests of indirect effects, including popular confidence interval-based methods, tend to produce poor Type I error rates when mediation does not occur and, more generally, only allow dichotomous decisions of "not significant" or "significant" with regards to the statistical conclusion. To remedy these issues, we propose a new method, a likelihood ratio test (LRT), that uses non-linear constraints in what we term the model-based constrained optimization (MBCO) procedure. The MBCO procedure (a) offers a more robust Type I error rate

than existing methods; (b) provides a p-value, which serves as a continuous measure of compatibility of data with the hypothesized null model (not just a dichotomous reject or fail-to-reject decision rule); (c) allows simple and complex hypotheses about mediation (i.e., one or more mediators; different mediational pathways), and (d) allows the mediation model to use observed or latent variables. The MBCO procedure is based on a structural equation modeling framework (even if latent variables are not specified) with specialized fitting routines, namely with the use of non-linear constraints. We advocate using the MBCO procedure to test hypotheses about an indirect effect in addition to reporting a confidence interval to capture uncertainty about the indirect effect because this combination transcends existing methods.

Tofighi et al.: Monte Carlo confidence intervals for complex functions of indirect effects
Tofighi-MacKinnon-2015

Abstract: One challenge in mediation analysis is to generate a confidence interval (CI) with high coverage and power that maintains a nominal significance level for any well-defined function of indirect and direct effects in the general context of structural equation modeling (SEM). This study discusses a proposed Monte Carlo extension that finds the CIs for any well-defined function of the coefficients of SEM such as the product of k coefficients and the ratio of the contrasts of indirect effects, using the Monte Carlo method. Finally, we conduct a small-scale simulation study to compare CIs produced by the Monte Carlo, nonparametric bootstrap, and asymptotic-delta methods. Based on our simulation study, we recommend researchers use the Monte Carlo method to test a complex function of indirect effects.

Dennis C. Turk and Elena S. Monarch. "Biopsychosocial perspective on chronic pain". In: *Psychological approaches to pain management: A practitioner's handbook*. Ed. by Dennis C. Turk and Robert J. Gatchel. 3rd ed. New York: The Guilford Press, 2018. ISBN: 9781462535620.

Uhlenbeck et al.: On the Theory of the Brownian Motion   Uhlenbeck-Ornstein-1930

G. E. Uhlenbeck and L. S. Ornstein. "On the Theory of the Brownian Motion". In: *Physical Review* 36.5 (Sept. 1930), pp. 823–841. DOI: 10.1103/physrev.36.823.

Abstract: With a method first indicated by Ornstein the mean values of all the powers of the velocity u and the displacement s of a free particle in Brownian motion are calculated. It is shown that $u - u_0 \exp(-\beta t)$ and $s - u_0 \beta[1 - \exp(-\beta t)]$ where u_0 is the initial velocity and β the friction coefficient divided by the mass of the particle, follow the normal Gaussian distribution law. For s this gives the exact frequency distribution corresponding to the exact formula for s^2 of Ornstein and Fürth. Discussion is given of the connection with the Fokker-Planck partial differential equation. By the same method exact expressions are obtained for the square of the deviation of a harmonically bound particle in Brownian motion as a function of the time and the initial deviation. Here the periodic, aperiodic and overdamped cases have to be treated separately. In the last case, when β is much larger than the frequency and for values of $t >> \beta^{-1}$, the formula takes the form of that previously given by Smoluchowski.

Usami: On the differences between general cross-lagged panel model and random-intercept cross-lagged panel model: Interpretation of cross-lagged parameters and model choice   Usami-2020

Satoshi Usami. "On the differences between general cross-lagged panel model and random-intercept cross-lagged panel model: Interpretation of cross-lagged parameters and model choice". In: *Struc-*

tural Equation Modeling: A Multidisciplinary Journal 28.3 (Oct. 2020), pp. 331–344. ISSN: 1532-8007. DOI: [10.1080/10705511.2020.1821690](10.1080/10705511.2020.1821690).

Abstract: Many methods have been developed to infer reciprocal relations between longitudinally observed variables. Among them, the general cross-lagged panel model (GCLM) is the most recent development as a variant of the cross-lagged panel model (CLPM), while the random-intercept CLPM (RI-CLPM) has rapidly become a popular approach. In this article, we describe how common factors and cross-lagged parameters included in these models can be interpreted, using a unified framework that was recently developed. Because common factors are modeled with lagged effects in the GCLM, they have both direct and indirect influences on observed scores, unlike stable trait factors included in the RI-CLPM. This indicates that the GCLM does not control for stable traits as the RI-CLPM does, and that there are interpretative differences in cross-lagged parameters between these models. We also explain that including such common factors as well as moving-average terms in the GCLM makes this interpretation very complicated.

Usami: Within-person variability score-based causal inference: A two-step estimation for joint effects of time-varying treatments   Usami-2022

Satoshi Usami. "Within-person variability score-based causal inference: A two-step estimation for joint effects of time-varying treatments". In: *Psychometrika* 88.4 (Aug. 2022), pp. 1466–1494. ISSN: 1860-0980. DOI: [10.1007/s11336-022-09879-1](10.1007/s11336-022-09879-1).

Abstract: Behavioral science researchers have shown strong interest in disaggregating within-person relations from between-person differences (stable traits) using longitudinal data. In this paper, we propose a method of within-person variability score-based causal inference for estimating joint effects of time-varying continuous treatments by controlling for stable traits of persons. After explaining the assumed data-generating process and providing formal definitions of stable trait factors, within-person variability scores, and joint effects of time-varying treatments at the within-person level, we introduce the proposed method, which consists of a two-step analysis. Within-person vari-

ability scores for each person, which are disaggregated from stable traits of that person, are first calculated using weights based on a best linear correlation preserving predictor through structural equation modeling (SEM). Causal parameters are then estimated via a potential outcome approach, either marginal structural models (MSMs) or structural nested mean models (SNMMs), using calculated within-person variability scores. Unlike the approach that relies entirely on SEM, the present method does not assume linearity for observed time-varying confounders at the within-person level. We emphasize the use of SNMMs with G-estimation because of its property of being doubly robust to model misspecifications in how observed time-varying confounders are functionally related to treatments/predictors and outcomes at the within-person level. Through simulation, we show that the proposed method can recover causal parameters well and that causal estimates might be severely biased if one does not properly account for stable traits. An empirical application using data regarding sleep habits and mental health status from the Tokyo Teen Cohort study is also provided.

Usami et al.: A unified framework of longitudinal models to examine reciprocal relations
Usami-Murayama-Hamaker-2019

Abstract: Inferring reciprocal effects or causality between variables is a central aim of behavioral and psychological research. To address reciprocal effects, a variety of longitudinal models that include cross-lagged relations have been proposed in different contexts and disciplines. However, the relations between these cross-lagged models have not been systematically discussed in the literature. This lack of insight makes it difficult for researchers to select an appropriate model when analyzing longitudinal data, and some researchers do not even think about alternative cross-lagged models. The present research provides a unified framework that clarifies the conceptual and mathematical similarities and differences between these models. The unified framework shows that

existing longitudinal models can be effectively classified based on whether the model posits unique factors and/or dynamic residuals and what types of common factors are used to model changes. The latter is essential to understand how cross-lagged parameters are interpreted. We also present an example using empirical data to demonstrate that there is great risk of drawing different conclusions depending on the cross-lagged models used.

van Buuren: Flexible imputation of missing data   vanBuuren-2018

Abstract: Missing data pose challenges to real-life data analysis. Simple ad-hoc fixes, like deletion or mean imputation, only work under highly restrictive conditions, which are often not met in practice. Multiple imputation replaces each missing value by multiple plausible values. The variability between these replacements reflects our ignorance of the true (but missing) value. Each of the completed data set is then analyzed by standard methods, and the results are pooled to obtain unbiased estimates with correct confidence intervals. Multiple imputation is a general approach that also inspires novel solutions to old problems by reformulating the task at hand as a missing-data problem. This is the second edition of a popular book on multiple imputation, focused on explaining the application of methods through detailed worked examples using the MICE package as developed by the author. This new edition incorporates the recent developments in this fast-moving field. This class-tested book avoids mathematical and technical details as much as possible: formulas are accompanied by verbal statements that explain the formula in accessible terms. The book sharpens the reader's intuition on how to think about missing data, and provides all the tools needed to execute a well-grounded quantitative analysis in the presence of missing data.

Library: QA278.

Stef van Buuren, J. P. L. Brand, et al. "Fully conditional specification in multivariate imputation". In: *Journal of Statistical Computation and Simulation* 76.12 (Dec. 2006), pp. 1049–1064. DOI: `10.1080/10629360600810434`.

Abstract: The use of the Gibbs sampler with fully conditionally specified models, where the distribution of each variable given the other variables is the starting point, has become a popular method to create imputations in incomplete multivariate data. The theoretical weakness of this approach is that the specified conditional densities can be incompatible, and therefore the stationary distribution to which the Gibbs sampler attempts to converge may not exist. This study investigates practical consequences of this problem by means of simulation. Missing data are created under four different missing data mechanisms. Attention is given to the statistical behavior under compatible and incompatible models. The results indicate that multiple imputation produces essentially unbiased estimates with appropriate coverage in the simple cases investigated, even for the incompatible models. Of particular interest is that these results were produced using only five Gibbs iterations starting from a simple draw from observed marginal distributions. It thus appears that, despite the theoretical weaknesses, the actual performance of conditional model specification for multivariate imputation can be quite good, and therefore deserves further study.

Stef van Buuren and Karin Groothuis-Oudshoorn. "mice: Multivariate Imputation by Chained Equations in R". In: *Journal of Statistical Software* 45.3 (2011). DOI: `10.18637/jss.v045.i03`.

Abstract: The R package mice imputes incomplete multivariate data by chained equations. The software mice 1.0 appeared in the year 2000 as an S-PLUS library, and in 2001 as an R package.

mice 1.0 introduced predictor selection, passive imputation and automatic pooling. This article documents mice, which extends the functionality of mice 1.0 in several ways. In mice, the analysis of imputed data is made completely general, whereas the range of models under which pooling works is substantially extended. mice adds new functionality for imputing multilevel data, automatic predictor selection, data handling, post-processing imputed values, specialized pooling routines, model selection tools, and diagnostic graphs. Imputation of categorical data is improved in order to bypass problems caused by perfect prediction. Special attention is paid to transformations, sum scores, indices and interactions using passive imputation, and to the proper setup of the predictor matrix. mice can be downloaded from the Comprehensive R Archive Network. This article provides a hands-on, stepwise approach to solve applied incomplete data problems.

van Egmond et al.: Wearable transdermal alcohol monitors: A systematic review of detection validity, and relationship between transdermal and breath alcohol concentration and influencing factors   vanEgmond-Wright-Livingston-etal-2020

Abstract: Background Research on alcohol consumption mostly relies on self-reported data, which are subject to recall bias. Wearable transdermal alcohol concentration (TAC) monitors address this limitation by continuously measuring the ethanol excreted via the skin. This systematic review aims to provide an overview of TAC monitors' reliability to detect alcohol consumption and methods to estimate breath alcohol concentration (BrAC) and number of standard drinks consumed in a given time frame. Methods The databases MEDLINE, PsycINFO, SCOPUS, Engineering Village, and CINAHL were systematically searched to identify 1,048 empirical research papers published from 2013 onwards, of which 13 were included after full-text screening. The selected studies included 3 TAC monitors: SCRAM, WristTAS, and Skyn. Results TAC measures of SCRAM, WrisTAS,

and Skyn are found to be positively correlated with BrAC ($r = 0.56$ to 0.79) and/or self-reports ($r = 0.62$). Using the AMS criteria for detection results in low sensitivity, adjusted criteria can increase the sensitivity of the SCRAM from 39.9 to 68.5%. The WrisTAS and an early prototype of the Skyn showed high failure rates (17 to 38%). Recent advances toward transforming the TAC data into more clinically relevant measures have led to the development of mathematical models and the BrAC Estimator Software. Using TAC data, both approaches produce estimates explaining 70 to 82% of actual BrAC and self-reported drinking or to highly correlate with the actual BrAC measures ($\beta = 0.90$ to 0.91). Conclusions Transdermal alcohol monitors offer an opportunity to measure alcohol consumption in a valid and continuous way with mathematical models and software estimating BrAC values improving interpretation of TAC data. However, the SCRAM seems unable to detect low-to-moderate drinking levels using the thresholds and criteria set by the manufacturer. Moreover, the WrisTAS and the Skyn prototype show a high failure rate, raising questions about reliability. Future studies will assess the validity of new-generation wristbands, including the next Skyn generations.

van Montfort et al.: Longitudinal research with latent variables

<div align="right">

vanMontfort-Oud-Satorra-2010

</div>

Kees van Montfort, Johan H. L. Oud, and A. Satorra, eds. *Longitudinal research with latent variables*. New York: Springer, 2010. 301 pp. ISBN: 9783642117602.

van Montfort et al.: Continuous time modeling in the behavioral and related sciences

<div align="right">

vanMontfort-Oud-Voelkle-2018

</div>

Kees van Montfort, Johan H. L. Oud, and Manuel C. Voelkle, eds. *Continuous time modeling in the behavioral and related sciences*. Springer International Publishing, 2018. DOI: 10.1007/978-3-319-77219-6.

Vanhasbroeck et al.: Computational models for affect dynamics

Vanhasbroeck-Ariens-Tuerlinckx-etal-2021

Niels Vanhasbroeck et al. "Computational models for affect dynamics". In: *Affect dynamics*. Springer International Publishing, 2021, pp. 213–260. ISBN: 9783030829650. DOI: 10.1007/978-3-030-82965-0_10.

Abstract: Computational models of affect dynamics are ubiquitous. These models are appropriate for either exploring intensive longitudinal data or testing theories about affect dynamics. In this chapter, we give a brief overview of some of the computational models that have been applied in the field of affect dynamics, focusing on both discrete-time and continuous-time models. The emphasis of this chapter lies on describing the core ideas of the models and how they can be interpreted. At the end, we provide references to other important topics for the interested reader.

Venables et al.: Modern applied statistics with S   **Venables-Ripley-2002**

W. N. Venables and B. D. Ripley. *Modern applied statistics with S*. Springer New York, 2002. DOI: 10.1007/978-0-387-21706-2.

Venzon et al.: A method for computing profile-likelihood-based confidence intervals

Venzon-Moolgavkar-1988

D. J. Venzon and S. H. Moolgavkar. "A method for computing profile-likelihood-based confidence intervals". In: *Applied Statistics* 37.1 (1988), p. 87. DOI: 10.2307/2347496.

Abstract: The method of constructing confidence regions based on the generalised likelihood ratio statistic is well known for parameter vectors. A similar construction of a confidence interval for a single entry of a vector can be implemented by repeatedly maximising over the other parameters. We present an algorithm for finding these confidence interval endpoints that requires less computation.

It employs a modified Newton-Raphson iteration to solve a system of equations that defines the endpoints.

Ver Hoef: Who invented the delta method?   VerHoef-2012

Abstract: Many statisticians and other scientists use what is commonly called the "delta method." However, few people know who proposed it. The earliest article was found in an obscure journal, and the author is rarely cited for his contribution. This article briefly reviews three modern versions of the delta method and how they are used. Then, some history on the author and the journal of the first known article on the delta method is given. The original author's specific contribution is reproduced, along with a discussion on possible reasons that it has been overlooked.

Viechtbauer: Confidence intervals for the amount of heterogeneity in meta-analysis
Viechtbauer-2007

Abstract: Effect size estimates to be combined in a systematic review are often found to be more variable than one would expect based on sampling differences alone. This is usually interpreted as evidence that the effect sizes are heterogeneous. A random-effects model is then often used to account for the heterogeneity in the effect sizes. A novel method for constructing confidence intervals for the amount of heterogeneity in the effect sizes is proposed that guarantees nominal coverage probabilities even in small samples when model assumptions are satisfied. A variety of existing approaches for constructing such confidence intervals are summarized and the various methods are applied to an example to illustrate their use. A simulation study reveals that the newly proposed method yields the most accurate coverage probabilities under conditions more analogous to practice,

where assumptions about normally distributed effect size estimates and known sampling variances only hold asymptotically.

Voelkle et al.: Continuous time modelling with individually varying time intervals for oscillating and non-oscillating processes — Voelkle-Oud-2012

Abstract: When designing longitudinal studies, researchers often aim at equal intervals. In practice, however, this goal is hardly ever met, with different time intervals between assessment waves and different time intervals between individuals being more the rule than the exception. One of the reasons for the introduction of continuous time models by means of structural equation modelling has been to deal with irregularly spaced assessment waves (e.g., Oud & Delsing, 2010). In the present paper we extend the approach to individually varying time intervals for oscillating and non-oscillating processes. In addition, we show not only that equal intervals are unnecessary but also that it can be advantageous to use unequal sampling intervals, in particular when the sampling rate is low. Two examples are provided to support our arguments. In the first example we compare a continuous time model of a bivariate coupled process with varying time intervals to a standard discrete time model to illustrate the importance of accounting for the exact time intervals. In the second example the effect of different sampling intervals on estimating a damped linear oscillator is investigated by means of a Monte Carlo simulation. We conclude that it is important to account for individually varying time intervals, and encourage researchers to conceive of longitudinal studies with different time intervals within and between individuals as an opportunity rather than a problem.

Abstract: Panel studies, in which the same subjects are repeatedly observed at multiple time points, are among the most popular longitudinal designs in psychology. Meanwhile, there exists a wide range of different methods to analyze such data, with autoregressive and cross-lagged models being 2 of the most well known representatives. Unfortunately, in these models time is only considered implicitly, making it difficult to account for unequally spaced measurement occasions or to compare parameter estimates across studies that are based on different time intervals. Stochastic differential equations offer a solution to this problem by relating the discrete time model to its underlying model in continuous time. It is the goal of the present article to introduce this approach to a broader psychological audience. A step-by-step review of the relationship between discrete and continuous time modeling is provided, and we demonstrate how continuous time parameters can be obtained via structural equation modeling. An empirical example on the relationship between authoritarianism and anomia is used to illustrate the approach.

Abstract: Empirical results from epidemiological studies on pain-depression comorbidity in primary care and population samples have shown that: (a) pain is as strongly associated with anxiety as with depressive disorders; (b) characteristics that most strongly predict depression are diffuseness of pain and the extent to which pain interferes with activities; (c) certain psychological symptoms (low energy, disturbed sleep, worry) are prominent among pain patients, while others (guilt, loneliness)

are not; (d) depression and pain dysfunction are evident early in the natural history of pain, but dysfunction and distress are often transient; and (e) among initially dysfunctional pain patients whose dysfunction is chronic, depression levels do not improve but neither do they increase over time with chronicity alone. These results seem consistent with these mechanisms of pain-depression comorbidity; (1) a trait of susceptibility to both dysphoric physical symptoms (including pain) and psychological symptoms (including depression), and a state of somatosensory amplification in which psychological distress amplifies dysphoric physical sensations (including pain); (2) psychological illness and behavioural dysfunction being interrelated features of a maladaptive response to pain evident early in the natural history of the condition, and often resolving during an early recovery phase; (3) pain constituting a significant physical and psychological stressor that may induce or exacerbate psychological distress. Thus, pain and psychological illness should be viewed as having reciprocal psychological and behavioural effects involving both processes of illness expression and adaption, as well as pain having specific effects on emotional state and behavioural function.

Vuorre et al.: Within-subject mediation analysis for experimental data in cognitive psychology and neuroscience   Vuorre-Bolger-2017

Abstract: Statistical mediation allows researchers to investigate potential causal effects of experimental manipulations through intervening variables. It is a powerful tool for assessing the presence and strength of postulated causal mechanisms. Although mediation is used in certain areas of psychology, it is rarely applied in cognitive psychology and neuroscience. One reason for the scarcity of applications is that these areas of psychology commonly employ within-subjects designs, and mediation models for within-subjects data are considerably more complicated than for between-subjects data. Here, we draw attention to the importance and ubiquity of mediational hypotheses in within-subjects designs, and we present a general and flexible software package for conducting Bayesian

288

within-subjects mediation analyses in the R programming environment. We use experimental data from cognitive psychology to illustrate the benefits of within-subject mediation for theory testing and comparison.

Waller: fungible: Psychometric functions from the Waller Lab Waller-2022

Niels G. Waller. *fungible: Psychometric functions from the Waller Lab*. The R Foundation, 2022. URL: https://CRAN.R-project.org/package=fungible.

K. Wang: Understanding power anomalies in mediation analysis Wang-2018

Kai Wang. "Understanding power anomalies in mediation analysis". In: *Psychometrika* 83.2 (June 2018), pp. 387–406. ISSN: 1860-0980. DOI: 10.1007/s11336-017-9598-1.

Abstract: Previous studies have found some puzzling power anomalies related to testing the indirect effect of a mediator. The power for the indirect effect stagnates and even declines as the size of the indirect effect increases. Furthermore, the power for the indirect effect can be much higher than the power for the total effect in a model where there is no direct effect and therefore the indirect effect is of the same magnitude as the total effect. In the presence of direct effect, the power for the indirect effect is often much higher than the power for the direct effect even when these two effects are of the same magnitude. In this study, the limiting distributions of related statistics and their non-centralities are derived. Computer simulations are conducted to demonstrate their validity. These theoretical results are used to explain the observed anomalies.

Lijuan Wang et al.: Investigating the impact of the time interval selection on autoregressive mediation modeling: Result interpretations, effect reporting, and temporal designs Wang-Zhang-2020

Lijuan Wang and Qian Zhang. "Investigating the impact of the time interval selection on autore-

gressive mediation modeling: Result interpretations, effect reporting, and temporal designs". In: *Psychological Methods* 25.3 (June 2020), pp. 271–291. ISSN: 1082-989X. DOI: 10.1037/met0000235.

Abstract: This study investigates the impact of the time interval (the time passed between 2 consecutive measurements) selection on autoregressive mediation modeling (AMM). For a widely used autoregressive mediation model, via analytical derivations, we explained why and how the conventionally reported time-specific coefficient estimates (e.g., $\hat{a}\hat{b}$ and $\hat{c}'$) and inference results in AMM provide limited information and can arrive in even misleading conclusions about direct and indirect effects over time. Furthermore, under the stationarity assumption, we proposed an approach to calculate the overall direct and indirect effect estimates over time and the time lag lengths at which they reach maxima, using AMM results. The derivation results revealed that the overall direct and indirect effect curves are asymptotically invariant to the time interval selection, under stationarity. With finite samples and thus sampling errors and potential computing problems, however, our simulation results revealed that the overall indirect effect curves were better recovered when the time interval is selected to be closer to half of the time lag length at which the overall indirect effect reaches its maximum. An R function and an R Shiny app were developed to obtain the overall direct and indirect effect curves over time and facilitate the time interval selection using AMM results. Our findings provide another look at the connections between AMM and continuous time mediation modeling and the connections are discussed.

Wechsler et al.: A gender-specific measure of binge drinking among college students
Wechsler-Dowdall-Davenport-etal-1995

Henry Wechsler et al. "A gender-specific measure of binge drinking among college students". In: *American Journal of Public Health* 85.7 (July 1995), pp. 982–985. ISSN: 1541-0048. DOI: 10.2105/ajph.85.7.982.

H. White: A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity   White-1980

Halbert White. "A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity". In: *Econometrica* 48.4 (May 1980), pp. 817–838. DOI: 10.2307/1912934.

Abstract: This paper presents a parameter covariance matrix estimator which is consistent even when the disturbances of a linear regression model are heteroskedastic. This estimator does not depend on a formal model of the structure of the heteroskedasticity. By comparing the elements of the new estimator to those of the usual covariance estimator, one obtains a direct test for heteroskedasticity, since in the absence of heteroskedasticity, the two estimators will be approximately equal, but will generally diverge otherwise. The test has an appealing least squares interpretation.

Wichers et al.: Critical slowing down as a personalized early warning signal for depression   Wichers-Groot-Psychosystems-2016

Marieke Wichers et al. "Critical slowing down as a personalized early warning signal for depression". In: *Psychotherapy and Psychosomatics* 85.2 (2016), pp. 114–116. ISSN: 1423-0348. DOI: 10.1159/000441458.

Thomas Ashby Wills et al.: Stress-Coping Model of Adolescent Substance Use

Wills-Filer-1996

Thomas Ashby Wills and Marnie Filer. "Stress-Coping Model of Adolescent Substance Use". In: *Advances in Clinical Child Psychology*. Springer US, 1996, pp. 91–132. ISBN: 9781461303237. DOI: 10.1007/978-1-4613-0323-7_3.

Abstract: The goal of this chapter is to discuss research on adolescent substance use from the perspective of a stress-coping model. In addition to the long-term health implications of cigarette smoking and alcohol use (e.g., Helzer, 1987; U.S. Department of Health and Human Services, 1988),

adolescent substance use is of concern to clinical psychology both because early onset of substance use has prognostic significance for later substance abuse problems (Robins & Przybeck, 1985) and because substance use tends to be correlated with other problem behaviors, including aggressive and depressive symptomatology (e.g., see Cole & Carpentieri, 1990; Loeber, 1988). Thus, research aimed at a better understanding of adolescent substance use has relevance for informing research on other types of child behavior problems.

Thomas Ashby Wills et al.: Role of parent support and peer support in adolescent substance use: A test of mediated effects   Wills-Resko-Ainette-etal-2004

Abstract: This research tested comparative effects of parent and peer support on adolescent substance use (tobacco, alcohol, and marijuana) with data from 2 assessments of a multiethnic sample of 1,826 adolescents, mean age 12.3 years. Multiple regression analyses indicated that parental support was inversely related to substance use and that peer support was positively related to substance use, as a suppression effect. Structural modeling analyses indicated that effects of support were mediated through pathways involving good self-control, poor self-control, and risk-taking tendency; parent and peer support had different patterns of relations to these mediators. The mediators had pathways to substance use through positive and negative recent events and through peer affiliations. Effects for gender and ethnicity were also noted. Mechanisms of operation for parent and peer support are discussed.

Wright: On the nature of size factors   Wright-1918

Wright: The relative importance of heredity and environment in determining the piebald pattern of guinea-pigs Wright-1920

Sewall Wright. "The relative importance of heredity and environment in determining the piebald pattern of guinea-pigs". In: *Proceedings of the National Academy of Sciences of the United States of America* 6.6 (1920), pp. 320–332. ISSN: 00278424. URL: http://www.jstor.org/stable/84353.

Wright: The method of path coefficients Wright-1934

Sewall Wright. "The method of path coefficients". In: *The Annals of Mathematical Statistics* 5.3 (Sept. 1934), pp. 161–215. DOI: 10.1214/aoms/1177732676.

W. Wu et al.: A new procedure to test mediation with missing data through nonparametric bootstrapping and multiple imputation Wu-Jia-2013

Wei Wu and Fan Jia. "A new procedure to test mediation with missing data through nonparametric bootstrapping and multiple imputation". In: *Multivariate Behavioral Research* 48.5 (Sept. 2013), pp. 663–691. DOI: 10.1080/00273171.2013.816235.

Abstract: This article proposes a new procedure to test mediation with the presence of missing data by combining nonparametric bootstrapping with multiple imputation (MI). This procedure performs MI first and then bootstrapping for each imputed data set. The proposed procedure is more computationally efficient than the procedure that performs bootstrapping first and then MI for each bootstrap sample. The validity of the procedure is evaluated using a simulation study under different sample size, missing data mechanism, missing data proportion, and shape of distribution conditions. The result suggests that the proposed procedure performs comparably to the procedure that combines bootstrapping with full information maximum likelihood under most conditions. However, caution needs to be taken when using this procedure to handle missing not-at-random or nonnormal data.

Yang: Evaluating latent class analysis models in qualitative phenotype identification

Yang-2006

Chih-Chien Yang. "Evaluating latent class analysis models in qualitative phenotype identification". In: *Computational Statistics & Data Analysis* 50.4 (Feb. 2006), pp. 1090–1104. ISSN: 0167-9473. DOI: 10.1016/j.csda.2004.11.004.

Abstract: The paper is aimed to investigate the performance of information criteria in selecting latent class analysis models which are often used in research of phenotype identification. Six information criteria and a sample size adjustment (Psychometrika 52 (1987) 333) are compared under various sample sizes and model dimensionalities. The simulation design is particularly meaningful for phenotypic research in practice. Results show that improvements by the sample size adjustment are considerable. In addition, the sample size and model dimensionality effects are found to be influential in the simulation study.

Yoo et al.: SLURM: Simple Linux Utility for Resource Management

Yoo-Jette-Grondona-2003

Andy B. Yoo, Morris A. Jette, and Mark Grondona. "SLURM: Simple Linux Utility for Resource Management". In: *Job scheduling strategies for parallel processing.* Ed. by Dror Feitelson, Larry Rudolph, and Uwe Schwiegelshohn. Springer Berlin Heidelberg, 2003, pp. 44–60. ISBN: 9783540397274. DOI: 10.1007/10968987_3.

Abstract: A new cluster resource management system called Simple Linux Utility Resource Management (SLURM) is described in this paper. SLURM, initially developed for large Linux clusters at the Lawrence Livermore National Laboratory (LLNL), is a simple cluster manager that can scale to thousands of processors. SLURM is designed to be flexible and fault-tolerant and can be ported to other clusters of different size and architecture with minimal effort. We are certain that SLURM

will benefit both users and system architects by providing them with a simple, robust, and highly scalable parallel job execution environment for their cluster system.

Yu et al.: Psychometric validation of the Internet Gaming Disorder-20 Test among Chinese middle school and university students   Yu-Pesigan-Zhang-etal-2019

Abstract: Background and aims: Internet gaming disorder (IGD) was proposed in Diagnostic and Statistical Manual of Mental Disorders of American Psychiatric Association as an area warranting more research attention. High prevalence of excessive Internet game use and related addictions has been reported in China, especially among youth; however, there is a lack of psychometrically and theoretically sound instruments for assessing IGD in the Chinese language. Methods: This study aimed to examine the psychometric properties of a Chinese version of the Internet Gaming Disorder Test (IGD-20 Test) among Chinese middle school ($n = 569$; $M_{\mathrm{age}} = 13.34$; 46.2% females) and university students ($n = 523$; $M_{\mathrm{age}} = 20.12$; 48.4% females) samples in Beijing, China. All participants voluntarily completed an anonymous questionnaire. Results: Confirmatory factor analysis results showed that the Chinese version of the IGD-20 Test had five factors (i.e., salience-tolerance, mood modification, withdrawal, conflict, and relapse). Measurement invariance was confirmed across the two samples. The test score was positively associated with the modified Young's Internet Addiction Test for gaming addiction. Concurrent validation was further demonstrated by the IGD-20 Test's positive correlation with weekly gameplay and depression symptoms. The latent profile analysis showed four different gamer classes (i.e., regular gamers, low-risk engaged gamers, high-risk engaged gamers, and probable disordered gamers), with the estimated prevalence of 2.1% of the last group. Discussion and conclusion: The IGD-20 Test was applicable to Chinese youth and its Chinese version generally demonstrated good psychometric properties.

Abstract: The aim of this study was to examine the effects of the two cognitive factors proposed by social cognitive theory to be highly influential on behavior (i.e., outcome expectancies and self-efficacy), in addition to optimism and loneliness, on social networking addiction among university students. In the study, 395 Chinese students (145 males, $M_{\text{age}} = 19.05$, age range = 17-27 years) voluntarily completed an online, anonymous questionnaire regarding their Internet use. Almost all of the participants (99%) were found to be using online social networking sites, and findings showed that social networking addiction was strongly correlated with Internet addiction. As hypothesized, more negative outcome expectancies and lower self-efficacy with regard to reducing Internet use were associated with higher social networking addictive tendencies. The results of the path analysis showed that low optimism was an indirect risk factor of social networking addiction through outcome expectancies and self-efficacy, while loneliness was a direct risk factor. The findings provide practical implications to preventive intervention for social networking addiction among youth.

K.-H. Yuan et al.: Three likelihood-based methods for mean and covariance structure analysis with nonnormal missing data Yuan-Bentler-2000

Abstract: Survey and longitudinal studies in the social and behavioral sciences generally contain missing data. Mean and covariance structure models play an important role in analyzing such

data. Two promising methods for dealing with missing data are a direct maximum-likelihood and a two-stage approach based on the unstructured mean and covariance estimates obtained by the EM-algorithm. Typical assumptions under these two methods are ignorable nonresponse and normality of data. However, data sets in social and behavioral sciences are seldom normal, and experience with these procedures indicates that normal theory based methods for nonnormal data very often lead to incorrect model evaluations. By dropping the normal distribution assumption, we develop more accurate procedures for model inference. Based on the theory of generalized estimating equations, a way to obtain consistent standard errors of the two-stage estimates is given. The asymptotic efficiencies of different estimators are compared under various assumptions. We also propose a minimum chi-square approach and show that the estimator obtained by this approach is asymptotically at least as efficient as the two likelihood-based estimators for either normal or nonnormal data. The major contribution of this paper is that for each estimator, we give a test statistic whose asymptotic distribution is chisquare as long as the underlying sampling distribution enjoys finite fourth-order moments. We also give a characterization for each of the two likelihood ratio test statistics when the underlying distribution is nonnormal. Modifications to the likelihood ratio statistics are also given. Our working assumption is that the missing data mechanism is missing completely at random. Examples and Monte Carlo studies indicate that, for commonly encountered nonnormal distributions, the procedures developed in this paper are quite reliable even for samples with missing data that are missing at random.

K.-H. Yuan et al.: Biases and standard errors of standardized regression coefficients
Yuan-Chan-2011

Abstract: The paper obtains consistent standard errors (SE) and biases of order $O(1/n)$ for the sample standardized regression coefficients with both random and given predictors. Analytical results indicate that the formulas for SEs given in popular text books are consistent only when the

297

population value of the regression coefficient is zero. The sample standardized regression coefficients are also biased in general, although it should not be a concern in practice when the sample size is not too small. Monte Carlo results imply that, for both standardized and unstandardized sample regression coefficients, SE estimates based on asymptotics tend to under-predict the empirical ones at smaller sample sizes.

Y. Yuan et al.: Bayesian mediation analysis.   Yuan-MacKinnon-2009

Abstract: In this article, we propose Bayesian analysis of mediation effects. Compared with conventional frequentist mediation analysis, the Bayesian approach has several advantages. First, it allows researchers to incorporate prior information into the mediation analysis, thus potentially improving the efficiency of estimates. Second, under the Bayesian mediation analysis, inference is straightforward and exact, which makes it appealing for studies with small samples. Third, the Bayesian approach is conceptually simpler for multilevel mediation analysis. Simulation studies and analysis of 2 data sets are used to illustrate the proposed methods.

Yzerbyt et al.: New recommendations for testing indirect effects in mediational models: The need to report and test component paths   Yzerbyt-Muller-Batailler-etal-2018

Abstract: In light of current concerns with replicability and reporting false-positive effects in psychology, we examine Type I errors and power associated with 2 distinct approaches for the assessment of mediation, namely the component approach (testing individual parameter estimates in the model) and the index approach (testing a single mediational index). We conduct simulations that

examine both approaches and show that the most commonly used tests under the index approach risk inflated Type I errors compared with the joint-significance test inspired by the component approach. We argue that the tendency to report only a single mediational index is worrisome for this reason and also because it is often accompanied by a failure to critically examine the individual causal paths underlying the mediational model. We recommend testing individual components of the indirect effect to argue for the presence of an indirect effect and then using other recommended procedures to calculate the size of that effect. Beyond simple mediation, we show that our conclusions also apply in cases of within-participant mediation and moderated mediation. We also provide a new R-package that allows for an easy implementation of our recommendations.

Zamboanga et al.: A "player's" game? Masculinity and drinking games participation among White and Asian American male college freshmen

Abstract: Research with college students indicates that conformity to distinct masculine norms is associated with heavy drinking and alcohol-related problems. Drinking games (DGs) involve heavy alcohol consumption, and although such games have been characterized as a male-dominated activity, studies have not examined how gender-relevant factors such as conformity to masculine norms are associated with DGs participation. Moreover, the extent to which these associations vary as a function of race/ethnicity warrants further exploration because the exact role that these factors play in increasing college students' risk for DGs participation is unclear. Thus, the primary aim of this study was to examine the associations between distinct masculine norms and frequency of DGs participation (while controlling for typical alcohol consumption) in a sample of White ($n = 328$) and Asian American ($n = 136$) college men ($M_{age} = 18.11$ years, SD $= 0.35$). A secondary aim was to test the degree to which such relationships are similar between these groups. Male

college freshmen from a public university completed self-report questionnaires. Results indicated that increased levels of conformity to the masculine norms of being a playboy and heterosexual presentation were significantly associated with more frequent DGs participation for White but not Asian American college men. Implications for intervention and future research directions are discussed.

Zamboanga et al.: Frequency of drinking games participation and alcohol-related problems in a multiethnic sample of college students: Do gender and ethnicity matter?

Abstract: Introduction: A drinking game (DG) is a high-risk, social drinking activity that consists of certain rules (i.e., when to drink and how much to consume) designed to promote inebriation and that requires each player to perform a cognitive and/or motor task (Zamboanga et al., 2013). Research suggests that non-White or female students who play DGs are at an increased risk of experiencing alcohol-related problems. Thus, this study examined whether the associations between DG participation and alcohol-related problems were similar for men and women and across ethnic groups. Method: College students (N = 7409; 73% women; 64% White, 8% Black, 14% Hispanic, 14% Asian) from 30 U.S. colleges/universities completed self-report questionnaires. Results: Controlling for age, site, Greek membership (i.e., membership in a fraternity or sorority), and typical alcohol consumption, results indicated that the association between DG participation and alcohol-related problems was stronger for men compared to women. With respect to ethnicity, the association between these variables was stronger among Black women than Black men. Conclusions: Findings from this large-scale study highlight the need to closely investigate how gender and ethnicity moderate the associations between DG participation and alcohol-related problems. College

intervention efforts designed to address high-risk drinking behaviors such as DG participation might consider paying close attention to ethnic minority populations, perhaps particularly Black women.

Zeileis: Econometric computing with HC and HAC covariance matrix estimators

<div align="right">Zeileis-2004</div>

Achim Zeileis. "Econometric computing with HC and HAC covariance matrix estimators". In: *Journal of Statistical Software* 11.10 (2004). DOI: 10.18637/jss.v011.i10.

Abstract: Data described by econometric models typically contains autocorrelation and/or heteroskedasticity of unknown form and for inference in such models it is essential to use covariance matrix estimators that can consistently estimate the covariance of the model parameters. Hence, suitable heteroskedasticity consistent (HC) and heteroskedasticity and autocorrelation consistent (HAC) estimators have been receiving attention in the econometric literature over the last 20 years. To apply these estimators in practice, an implementation is needed that preferably translates the conceptual properties of the underlying theoretical frameworks into computational tools. In this paper, such an implementation in the package sandwich in the R system for statistical computing is described and it is shown how the suggested functions provide reusable components that build on readily existing functionality and how they can be integrated easily into new inferential procedures or applications. The toolbox contained in sandwich is extremely flexible and comprehensive, including specific functions for the most important HC and HAC estimators from the econometric literature. Several real-world data sets are used to illustrate how the functionality can be integrated into applications.

Zeileis: Object-oriented computation of sandwich estimators   Zeileis-2006

Achim Zeileis. "Object-oriented computation of sandwich estimators". In: *Journal of Statistical Software* 16.9 (Aug. 2006). DOI: 10.18637/jss.v016.i09.

Abstract: Sandwich covariance matrix estimators are a popular tool in applied regression modeling for performing inference that is robust to certain types of model misspecification. Suitable implementations are available in the R system for statistical computing for certain model fitting functions only (in particular lm()), but not for other standard regression functions, such as glm(), nls(), or survreg(). Therefore, conceptual tools and their translation to computational tools in the package sandwich are discussed, enabling the computation of sandwich estimators in general parametric models. Object orientation can be achieved by providing a few extractor functions' most importantly for the empirical estimating functions' from which various types of sandwich estimators can be computed.

Zeileis et al.: Various versatile variances: An object-oriented implementation of clustered covariances in R   Zeileis-Koll-Graham-2020

Abstract: Clustered covariances or clustered standard errors are very widely used to account for correlated or clustered data, especially in economics, political sciences, and other social sciences. They are employed to adjust the inference following estimation of a standard least-squares regression or generalized linear model estimated by maximum likelihood. Although many publications just refer to "the" clustered standard errors, there is a surprisingly wide variety of clustered covariances, particularly due to different flavors of bias corrections. Furthermore, while the linear regression model is certainly the most important application case, the same strategies can be employed in more general models (e.g., for zero-inflated, censored, or limited responses). In R, functions for covariances in clustered or panel models have been somewhat scattered or available only for certain modeling functions, notably the (generalized) linear regression model. In contrast, an object-oriented approach to "robust"' covariance matrix estimation - applicable beyond lm() and glm() - is available in the sandwich package but has been limited to the case of cross-section or time series data. Starting

with sandwich 2.4.0, this shortcoming has been corrected: Based on methods for two generic functions (estfun() and bread()), clustered and panel covariances are provided in vcovCL(), vcovPL(), and vcovPC(). Moreover, clustered bootstrap covariances are provided in vcovBS(), using model update() on bootstrap samples. These are directly applicable to models from packages including MASS, pscl, countreg, and betareg, among many others. Some empirical illustrations are provided as well as an assessment of the methods' performance in a simulation study.

G. Zhang: Testing process factor analysis models using the parametric bootstrap

Guangjian Zhang. "Testing process factor analysis models using the parametric bootstrap". In: *Multivariate Behavioral Research* 53.2 (Jan. 2018), pp. 219–230. ISSN: 1532-7906. DOI: 10.1080/00273171.2017.1415123.

Abstract: Process factor analysis (PFA) is a latent variable model for intensive longitudinal data. It combines P-technique factor analysis and time series analysis. The goodness-of-fit test in PFA is currently unavailable. In the paper, we propose a parametric bootstrap method for assessing model fit in PFA. We illustrate the test with an empirical data set in which 22 participants rated their effects everyday over a period of 90 days. We also explore Type I error and power of the parametric bootstrap test with simulated data.

G. Zhang et al.: Bootstrap standard error estimates in dynamic factor analysis

Guangjian Zhang and Michael W. Browne. "Bootstrap standard error estimates in dynamic factor analysis". In: *Multivariate Behavioral Research* 45.3 (May 2010), pp. 453–482. ISSN: 1532-7906. DOI: 10.1080/00273171.2010.483375.

Abstract: Dynamic factor analysis summarizes changes in scores on a battery of manifest variables over repeated measurements in terms of a time series in a substantially smaller number of latent fac-

tors. Algebraic formulae for standard errors of parameter estimates are more difficult to obtain than in the usual intersubject factor analysis because of the interdependence of successive observations. Bootstrap methods can fill this need, however. The standard bootstrap of individual timepoints is not appropriate because it destroys their order in time and consequently gives incorrect standard error estimates. Two bootstrap procedures that are appropriate for dynamic factor analysis are described. The moving block bootstrap breaks down the original time series into blocks and draws samples of blocks instead of individual timepoints. A parametric bootstrap is essentially a Monte Carlo study in which the population parameters are taken to be estimates obtained from the available sample. These bootstrap procedures are demonstrated using 103 days of affective mood self-ratings from a pregnant woman, 90 days of personality self-ratings from a psychology freshman, and a simulation study.

G. Zhang et al.: Dynamic factor analysis with multivariate time series of multiple individuals: An error-corrected estimation method. Zhang-Lee-Li-etal-2025

Abstract: Intensive longitudinal data, increasingly common in social and behavioral sciences, often consist of multivariate time series from multiple individuals. Dynamic factor analysis, combining factor analysis and time series analysis, has been used to uncover individual-specific processes from single-individual time series. However, integrating these processes across individuals is challenging due to estimation errors in individual-specific parameter estimates. We propose a method that integrates individual-specific processes while accommodating the corresponding estimation error. This method is computationally efficient and robust against model specification errors and nonnormal data. We compare our method with a Naive approach that ignores estimation error using both empirical and simulated data. The two methods produced similar estimates for fixed effect parameters, but the proposed method produced more satisfactory estimates for random effects than the

Naive method. The relative advantage of the proposed method was more substantial for short to moderately long time series (T = 56-200).

M. X. Zhang et al.: Effects of social and outcome expectancies on hazardous drinking among Chinese university students: The mediating role of drinking motivations

Abstract: Background and Objectives: Based on the theory of reasoned action, the present study investigated the relative effects of drinking outcome expectancies and parental norms, as well as the mediating effect of drinking motivations, on hazardous drinking in Chinese university students. Method: A sample of Chinese university students in Hong Kong and Macao ($N = 973$, $M = 19.82$, $SD = 1.57$, 48.9% males), who reported drinking in the past 3 months, voluntarily completed an anonymous questionnaire. Path analysis was used to test the effects of the variables on hazardous drinking. Results: All the psychosocial variables showed positive correlations with hazardous drinking. In the path model, controlling for sex, parental norms had both direct and indirect effects on hazardous drinking through social and enhancement motivations. Courage had the strongest indirect effect on drinking behavior through social, enhancement, and coping motivations, whereas the relationship between tension reduction and hazardous drinking was mediated by enhancement and coping motivations. Sociality and sexuality only had indirect effect through social and coping motivations respectively. Negative outcome expectancies had no direct nor indirect effects on hazardous drinking. Conclusions: Perceived approval from parents and positive alcohol outcome expectancies may enhance individuals' tendency to engage in hazardous drinking by increasing their motivation to drink to be social, for enjoyment, and to cope with problems. Parents should explicitly show their

disapproval of their children's drinking, and education efforts should focus on decreasing positive outcome expectancies and associated motivations for drinking among Chinese university students.

M. X. Zhang et al.: Psychometric properties of a Chinese version of the Brief Young Adult Alcohol Consequences Questionnaire (B-YAACQ)

Zhang-Pesigan-Kahler-etal-2019

Abstract: Aim: This study evaluated the psychometric properties of the Chinese version of the Brief Young Adult Alcohol Consequences Questionnaire (B-YAACQ). Method: In this study, 1616 Chinese university students (male = 58.66%; M_{age} = 19.88) reporting past-year drinking experience voluntarily completed an anonymous questionnaire. Rasch analysis, reliability analysis, and linear modeling were performed to examine the psychometric properties of the Chinese version of B-YAACQ. Results: Results of Rasch analysis and reliability analysis supported the assumption of uni-dimensionality, local independence, and internal consistency of the 24-item B-YAACQ in our Chinese sample. However, six items had marginal outfit statistics and/or potential gender bias; therefore, a model with 18 items was also tested after removing these items. The 18-item model showed excellent fit to the uni-dimensional model with no gender bias. However, the Pearson correlation between the 24-item and 18-item versions was $r = 0.98$, suggesting highly similar measurement. Both versions demonstrated concurrent validity through positive association with the Alcohol Use Disorder Identification Test (AUDIT) subscales, even after controlling for the effects of age and gender. Conclusion: This study is the first to validate a measurement tool for negative drinking consequences for university students in China. Despite some limitations, the original 24-item B-YAACQ was shown to have satisfactory psychometric properties when applied to Chinese university students. We recommend the shorter 18-item version without significant gender bias for testing gender differences.

Zhiyong Zhang and Lijuan Wang. "Methods for mediation analysis with missing data". In: *Psychometrika* 78.1 (Dec. 2012), pp. 154–184. DOI: 10.1007/s11336-012-9301-5.

Abstract: Despite wide applications of both mediation models and missing data techniques, formal discussion of mediation analysis with missing data is still rare. We introduce and compare four approaches to dealing with missing data in mediation analysis including listwise deletion, pairwise deletion, multiple imputation (MI), and a two-stage maximum likelihood (TS-ML) method. An R package bmem is developed to implement the four methods for mediation analysis with missing data in the structural equation modeling framework, and two real examples are used to illustrate the application of the four methods. The four methods are evaluated and compared under MCAR, MAR, and MNAR missing data mechanisms through simulation studies. Both MI and TS-ML perform well for MCAR and MAR data regardless of the inclusion of auxiliary variables and for AV-MNAR data with auxiliary variables. Although listwise deletion and pairwise deletion have low power and large parameter estimation bias in many studied conditions, they may provide useful information for exploring missing mechanisms.

Z. Zhang et al.: Mediation analysis with missing data through multiple imputation and bootstrap   Zhang-Wang-Tong-2015

Zhiyong Zhang, Lijuan Wang, and Xin Tong. "Mediation analysis with missing data through multiple imputation and bootstrap". In: *Quantitative Psychology Research*. Springer International Publishing, 2015, pp. 341–355. DOI: 10.1007/978-3-319-19977-1_24.

Zyphur et al.: From data to causes I: Building a general cross-lagged panel model (GCLM)   Zyphur-Allison-Tay-etal-2019

Michael J. Zyphur, Paul D. Allison, et al. "From data to causes I: Building a general cross-lagged

panel model (GCLM)". In: *Organizational Research Methods* 23.4 (May 2019), pp. 651–687. ISSN: 1552-7425. DOI: 10.1177/1094428119847278.

Abstract: This is the first paper in a series of two that synthesizes, compares, and extends methods for causal inference with longitudinal panel data in a structural equation modeling (SEM) framework. Starting with a cross-lagged approach, this paper builds a general cross-lagged panel model (GCLM) with parameters to account for stable factors while increasing the range of dynamic processes that can be modeled. We illustrate the GCLM by examining the relationship between national income and subjective well-being (SWB), showing how to examine hypotheses about short-run (via Granger-Sims tests) versus long-run effects (via impulse responses). When controlling for stable factors, we find no short-run or long-run effects among these variables, showing national SWB to be relatively stable, whereas income is less so. Our second paper addresses the differences between the GCLM and other methods. Online Supplementary Materials offer an Excel file automating GCLM input for Mplus (with an example also for Lavaan in R) and analyses using additional data sets and all program input/output. We also offer an introductory GCLM presentation at https://youtu.be/tHnnaRNPbXs. We conclude with a discussion of issues surrounding causal inference.

Zyphur et al.: From data to causes II: Comparing approaches to panel data analysis
Zyphur-Voelkle-Tay-etal-2019

Michael J. Zyphur, Manuel C. Voelkle, et al. "From data to causes II: Comparing approaches to panel data analysis". In: *Organizational Research Methods* 23.4 (May 2019), pp. 688–716. ISSN: 1552-7425. DOI: 10.1177/1094428119847280.

Abstract: This article compares a general cross-lagged model (GCLM) to other panel data methods based on their coherence with a causal logic and pragmatic concerns regarding modeled dynamics and hypothesis testing. We examine three "static" models that do not incorporate temporal dynamics: random- and fixed-effects models that estimate contemporaneous relationships; and latent

curve models. We then describe "dynamic" models that incorporate temporal dynamics in the form of lagged effects: cross-lagged models estimated in a structural equation model (SEM) or multilevel model (MLM) framework; Arellano-Bond dynamic panel data methods; and autoregressive latent trajectory models. We describe the implications of overlooking temporal dynamics in static models and show how even popular cross-lagged models fail to control for stable factors over time. We also show that Arellano-Bond and autoregressive latent trajectory models have various shortcomings. By contrasting these approaches, we clarify the benefits and drawbacks of common methods for modeling panel data, including the GCLM approach we propose. We conclude with a discussion of issues regarding causal inference, including difficulties in separating different types of time-invariant and time-varying effects over time.