

Ivan Jacob Agaloos Pesigan

August 29, 2025

References

Aalen et al.: Can we believe the DAGs? A comment on the relationship between causal DAGs and mechanisms **Aalen-Roysland-Gran-et-al-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](https://doi.org/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](https://doi.org/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.

**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](https://doi.org/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. Alarming, 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-et-al-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](https://doi.org/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 in-

cluded. 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).

Asparouhov et al.: Dynamic structural equation models

Asparouhov-Hamaker-Muthen-2017

Tihomir Asparouhov, Ellen L. Hamaker, and Bengt Muthén. “Dynamic structural equation models”. In: *Structural Equation Modeling: A Multidisciplinary Journal* 25.3 (Dec. 2017), pp. 359–388. DOI: [10.1080/10705511.2017.1406803](https://doi.org/10.1080/10705511.2017.1406803).

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.

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](https://doi.org/10.1080/10705511.2018.1511375).

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

Jacqueline M Barker and Jane R Taylor. “Habitual alcohol seeking: Modeling the transition from casual drinking to addiction”. In: *Neuroscience & Biobehavioral Reviews* 47 (Nov. 2014), pp. 281–294. ISSN: 0149-7634. DOI: [10.1016/j.neubiorev.2014.08.012](https://doi.org/10.1016/j.neubiorev.2014.08.012).

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.

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](https://doi.org/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 improve 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.

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](https://doi.org/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 choosing 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.

**Bernardo et al.: Pathways from collectivist coping to life satisfaction among Chinese:
The roles of locus-of-hope** **Bernardo-Wang-Pesigan-et-al-2017**

Allan B.I. Bernardo et al. “Pathways from collectivist coping to life satisfaction among Chinese: The roles of locus-of-hope”. In: *Personality and Individual Differences* 106 (Feb. 2017), pp. 253–256. ISSN: 0191-8869. DOI: [10.1016/j.paid.2016.10.059](https://doi.org/10.1016/j.paid.2016.10.059).

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.

Biesanz et al.: Assessing mediational models: Testing and interval estimation for indirect effects**Biesanz-Falk-Savalei-2010**

Jeremy C. Biesanz, Carl F. Falk, and Victoria Savalei. “Assessing mediational models: Testing and interval estimation for indirect effects”. In: *Multivariate Behavioral Research* 45.4 (Aug. 2010), pp. 661–701. DOI: [10.1080/00273171.2010.498292](https://doi.org/10.1080/00273171.2010.498292).

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

Blanca-Arnau-LopezMontiel-et-al-2013

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](https://doi.org/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

Boettiger-Eddelbuettel-2017

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](https://doi.org/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.

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

Kenneth A. Bollen and Jennie E. Brand. “A general panel model with random and fixed effects: A structural equations approach”. In: *Social Forces* 89.1 (Sept. 2010), pp. 1–34. ISSN: 1534-7605. DOI: [10.1353/sof.2010.0072](https://doi.org/10.1353/sof.2010.0072).

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.

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

Bond-Greenfield-Patterson-et-al-2014

Jason C. Bond et al. “Adjustments for drink size and ethanol content: New results from a self-report

diary and transdermal sensor validation study”. In: *Alcoholism: Clinical and Experimental Research* 38.12 (Dec. 2014), pp. 3060–3067. ISSN: 1530-0277. DOI: [10.1111/acer.12589](https://doi.org/10.1111/acer.12589).

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 ≥ 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.

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

Juan Carlos Bou and Albert Satorra. “Univariate versus multivariate modeling of panel data: Model specification and goodness-of-fit testing”. In: *Organizational Research Methods* 21.1 (June 2017), pp. 150–196. ISSN: 1552-7425. DOI: [10.1177/1094428117715509](https://doi.org/10.1177/1094428117715509).

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.

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

Laura F. Bringmann et al. “A network approach to psychopathology: New insights into clinical longitudinal data”. In: *PLoS ONE* 8.4 (Apr. 2013). ISSN: 1932-6203. DOI: [10.1371/journal.pone.0060188](https://doi.org/10.1371/journal.pone.0060188).

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.

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

CastroSchilo-Ferrer-2013

Laura Castro-Schilo and Emilio Ferrer. “Comparison of nomothetic versus idiographic-oriented methods for making predictions about distal outcomes from time series data”. In: *Multivariate Behavioral Research* 48.2 (Mar. 2013), pp. 175–207. ISSN: 1532-7906. DOI: [10.1080/00273171.2012.736042](https://doi.org/10.1080/00273171.2012.736042).

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.

Chen et al.: Vector autoregression, structural equation modeling, and their synthesis in neuroimaging data analysis **Chen-Daniel-Ziad-et-al-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](https://doi.org/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.

Cheung: Multivariate meta-analysis as structural equation models **Cheung-2013**

Mike W.-L. Cheung. "Multivariate meta-analysis as structural equation models". In: *Structural*

Equation Modeling: A Multidisciplinary Journal 20.3 (July 2013), pp. 429–454. ISSN: 1532-8007.
DOI: [10.1080/10705511.2013.797827](https://doi.org/10.1080/10705511.2013.797827).

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.

Chow et al.: Equivalence and differences between structural equation modeling and state-space modeling techniques **Chow-Ho-Hamaker-et-al-2010**

Sy-Miin Chow, Moon-ho R. Ho, et al. “Equivalence and differences between structural equation modeling and state-space modeling techniques”. In: *Structural Equation Modeling: A Multidisciplinary Journal* 17.2 (Apr. 2010), pp. 303–332. DOI: [10.1080/10705511003661553](https://doi.org/10.1080/10705511003661553).

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.: The cusp catastrophe model as cross-sectional and longitudinal mixture structural equation models **Chow-Witkiewitz-Grasman-et-al-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](https://doi.org/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 equation 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.

Clapp et al.: Examining the social ecology of a bar-crawl: An exploratory pilot study **Clapp-Madden-Mooney-et-al-2017**

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](https://doi.org/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.

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

Crocetti-Hale-Dimitrova-et-al-2014

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 (Aug. 2014), pp. 159–174. ISSN: 1573-3319. DOI: [10.1007/s10566-014-9275-9](https://doi.org/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.

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

Patrick J. Curran and Daniel J. Bauer. "The disaggregation of within-person and between-person effects in longitudinal models of change". In: *Annual Review of Psychology* 62.1 (Jan. 2011), pp. 583–619. DOI: [10.1146/annurev.psych.093008.100356](https://doi.org/10.1146/annurev.psych.093008.100356).

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.

Curran-Howard-Bainter-et-al-2014

Patrick J. Curran, Andrea L. Howard, et al. "The separation of between-person and within-person components of individual change over time: A latent curve model with structured residuals." In: *Journal of Consulting and Clinical Psychology* 82.5 (2014), pp. 879–894. ISSN: 0022-006X. DOI: [10.1037/a0035297](https://doi.org/10.1037/a0035297).

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.

Silvia de Haan-Rietdijk et al. “Discrete- vs. continuous-time modeling of unequally spaced experience sampling method data”. In: *Frontiers in Psychology* 8 (Oct. 2017). ISSN: 1664-1078. DOI: [10.3389/fpsyg.2017.01849](https://doi.org/10.3389/fpsyg.2017.01849).

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**

Pascal R. Deboeck and Aaron J. Boulton. “Integration of stochastic differential equations using structural equation modeling: A method to facilitate model fitting and pedagogy”. In: *Structural Equation Modeling: A Multidisciplinary Journal* 23.6 (Oct. 2016), pp. 888–903. ISSN: 1532-8007. DOI: [10.1080/10705511.2016.1218763](https://doi.org/10.1080/10705511.2016.1218763).

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**

Pascal R. Deboeck and Kristopher J. Preacher. “No need to be discrete: A method for continuous time mediation analysis”. In: *Structural Equation Modeling: A Multidisciplinary Journal* 23.1 (June 2015), pp. 61–75. DOI: [10.1080/10705511.2014.973960](https://doi.org/10.1080/10705511.2014.973960).

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.

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

Marina Demeshko et al. “A novel continuous and structural VAR modeling approach and its application to reactor noise analysis”. In: *ACM Transactions on Intelligent Systems and Technology* 7.2 (Nov. 2015), pp. 1–22. DOI: [10.1145/2710025](https://doi.org/10.1145/2710025).

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.

Doñamayor et al.: The involuntary nature of binge drinking: Goal directedness and awareness of intention
Donamayor-Strelchuk-Baek-et-al-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](https://doi.org/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.

Charles C. Driver, Johan H. L. Oud, and Manuel C. Voelkle. “Continuous time structural equation modeling with R package ctsem”. In: *Journal of Statistical Software* 77.5 (2017). ISSN: 1548-7660. DOI: [10.18637/jss.v077.i05](https://doi.org/10.18637/jss.v077.i05).

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.

Charles C. Driver and Manuel C. Voelkle. “Hierarchical Bayesian continuous time dynamic mod-

eling.” In: *Psychological Methods* 23.4 (Dec. 2018), pp. 774–799. ISSN: 1082-989X. DOI: [10.1037/met0000168](https://doi.org/10.1037/met0000168).

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**

Paul Dudgeon. “Some improvements in confidence intervals for standardized regression coefficients”. In: *Psychometrika* 82.4 (Mar. 2017), pp. 928–951. DOI: [10.1007/s11336-017-9563-z](https://doi.org/10.1007/s11336-017-9563-z).

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.

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](https://doi.org/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](https://doi.org/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.: 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](https://doi.org/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.

Efron: Bayesian inference and the parametric bootstrap

Efron-2012

Bradley Efron. “Bayesian inference and the parametric bootstrap”. In: *The Annals of Applied Statistics* 6.4 (Dec. 2012). DOI: [10.1214/12-aos571](https://doi.org/10.1214/12-aos571).

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 exponential 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.

Enders et al.: A Bayesian approach for estimating mediation effects with missing data

Enders-Fairchild-MacKinnon-2013

Craig K. Enders, Amanda J. Fairchild, and David P. MacKinnon. “A Bayesian approach for estimating mediation effects with missing data”. In: *Multivariate Behavioral Research* 48.3 (May 2013), pp. 340–369. ISSN: 1532-7906. DOI: [10.1080/00273171.2013.784862](https://doi.org/10.1080/00273171.2013.784862).

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.

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](https://doi.org/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-et-al-2012

Sacha Epskamp, Angélique O. J. Cramer, et al. “qgraph: Network visualizations of relationships in psychometric data”. In: *Journal of Statistical Software* 48.4 (2012). ISSN: 1548-7660. DOI: [10.18637/jss.v048.i04](https://doi.org/10.18637/jss.v048.i04).

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-et-al-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](https://doi.org/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., $n \geq 1$ 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.

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

Andrea K. Finlay et al. "Leisure activities, the social weekend, and alcohol use: Evidence from a daily study of first-year college students". In: *Journal of Studies on Alcohol and Drugs* 73.2 (Mar. 2012), pp. 250–259. ISSN: 1938-4114. DOI: [10.15288/jsad.2012.73.250](https://doi.org/10.15288/jsad.2012.73.250).

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.

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](https://doi.org/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.

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-et-al-2010

Kathleen M. 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”. In: *NeuroImage* 50.3 (Apr. 2010), pp. 1118–1125. DOI: [10.1016/j.neuroimage.2009.12.117](https://doi.org/10.1016/j.neuroimage.2009.12.117).

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.

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

Thomas K Greenfield et al. "Risks of alcohol use disorders related to drinking patterns in the U.S. general population". In: *Journal of Studies on Alcohol and Drugs* 75.2 (Mar. 2014), pp. 319–327. ISSN: 1938-4114. DOI: [10.15288/jsad.2014.75.319](https://doi.org/10.15288/jsad.2014.75.319).

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 ($n_s = 5,922$ 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.

Fei Gu, Kristopher J. Preacher, and Emilio Ferrer. “A state space modeling approach to mediation analysis”. In: *Journal of Educational and Behavioral Statistics* 39.2 (Apr. 2014), pp. 117–143. ISSN: 1935-1054. DOI: [10.3102/1076998614524823](https://doi.org/10.3102/1076998614524823).

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.

E. L. Hamaker et al. “Modeling affect dynamics: State of the art and future challenges”. In: *Emotion Review* 7.4 (July 2015), pp. 316–322. ISSN: 1754-0747. DOI: [10.1177/1754073915590619](https://doi.org/10.1177/1754073915590619).

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.

Ellen L. Hamaker et al.: A critique of the cross-lagged panel model

Hamaker-Kuiper-Grasman-2015

Ellen L. Hamaker, Rebecca M. Kuiper, and Raoul P. P. P. Grasman. “A critique of the cross-lagged panel model”. In: *Psychological Methods* 20.1 (2015), pp. 102–116. DOI: [10.1037/a0038889](https://doi.org/10.1037/a0038889).

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.

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

Hamaker-Schuurman-Zijlmans-2016

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](https://doi.org/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.

**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](https://doi.org/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

Hecht-Voelkle-2019

Martin Hecht and Manuel C. Voelke. “Continuous-time modeling in prevention research: An illustration”. In: *International Journal of Behavioral Development* 45.1 (Nov. 2019), pp. 19–27. ISSN: 1464-0651. DOI: [10.1177/0165025419885026](https://doi.org/10.1177/0165025419885026).

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.

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

Tracy E. Herring et al. “Utility of the Athlete Drinking Scale for assessing drinking motives among high school athletes”. In: *Addictive Behaviors* 60 (Sept. 2016), pp. 18–23. ISSN: 0306-4603. DOI: [10.1016/j.addbeh.2016.03.026](https://doi.org/10.1016/j.addbeh.2016.03.026).

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.

Hesterberg: What teachers should know about the bootstrap: Resampling in the undergraduate statistics curriculum **Hesterberg-2015**

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](https://doi.org/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.

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

Ralph Hingson, Wenxing Zha, and Daniel Smyth. “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”. In: *Journal of Studies on Alcohol and Drugs* 78.4 (July 2017), pp. 540–548. ISSN: 1938-4114. DOI: [10.15288/jsad.2017.78.540](https://doi.org/10.15288/jsad.2017.78.540).

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.

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

Hunter-2017

Michael D. Hunter. “State space modeling in an open source, modular, structural equation modeling environment”. In: *Structural Equation Modeling: A Multidisciplinary Journal* 25.2 (Oct. 2017), pp. 307–324. DOI: [10.1080/10705511.2017.1369354](https://doi.org/10.1080/10705511.2017.1369354).

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.

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**

Mark P. Jensen and Dennis C. Turk. “Contributions of psychology to the understanding and treatment of people with chronic pain: Why it matters to ALL psychologists”. In: *American Psychologist* 69.2 (2014), pp. 105–118. ISSN: 0003-066X. DOI: [10.1037/a0035641](https://doi.org/10.1037/a0035641).

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.

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](https://doi.org/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. With random predictors, this method also fails to account for the sampling variability of the pre-

dictor 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-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](https://doi.org/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.

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

Anne E. Kazak. “Editorial: Journal article reporting standards”. In: *American Psychologist* 73.1 (Jan. 2018), pp. 1–2. ISSN: 0003-066X. DOI: [10.1037/amp0000263](https://doi.org/10.1037/amp0000263).

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

Ken Kelley and Kristopher J. Preacher. “On effect size”. In: *Psychological Methods* 17.2 (2012), pp. 137–152. ISSN: 1082-989X. DOI: [10.1037/a0028086](https://doi.org/10.1037/a0028086).

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.

David A. Kenny and Charles M. Judd. “Power anomalies in testing mediation”. In: *Psychological Science* 25.2 (Dec. 2013), pp. 334–339. ISSN: 1467-9280. DOI: [10.1177/0956797613502676](https://doi.org/10.1177/0956797613502676).

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.

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](https://doi.org/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.

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

Joel Koopman et al. “Small sample mediation testing: Misplaced confidence in bootstrapped confidence intervals”. In: *Journal of Applied Psychology* 100.1 (2015), pp. 194–202. DOI: [10.1037/a0036635](https://doi.org/10.1037/a0036635).

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

Jolanda J. Kossakowski et al. "Data from 'Critical slowing down as a personalized early warning signal for depression'". In: *Journal of Open Psychology Data* 5 (Feb. 2017). ISSN: 2050-9863. DOI: [10.5334/jopd.29](https://doi.org/10.5334/jopd.29).

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.

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

Rebecca M. Kuiper and Oisin Ryan. "Drawing conclusions from cross-lagged relationships: Re-considering the role of the time-interval". In: *Structural Equation Modeling: A Multidisciplinary Journal* 25.5 (Mar. 2018), pp. 809–823. DOI: [10.1080/10705511.2018.1431046](https://doi.org/10.1080/10705511.2018.1431046).

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 con-

flicts 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.

Kuppens: It's about time: A special section on affect dynamics**Kuppens-2015**

Peter Kuppens. "It's about time: A special section on affect dynamics". In: *Emotion Review* 7.4 (July 2015), pp. 297–300. ISSN: 1754-0747. DOI: [10.1177/1754073915590947](https://doi.org/10.1177/1754073915590947).

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.

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). Ed. by Attila Gursoy, e0177459. DOI: [10.1371/journal.pone.0177459](https://doi.org/10.1371/journal.pone.0177459).

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

Joyce L. Y. Kwan and Wai Chan. "Comparing standardized coefficients in structural equation

modeling: A model reparameterization approach”. In: *Behavior Research Methods* 43.3 (Apr. 2011), pp. 730–745. DOI: [10.3758/s13428-011-0088-6](https://doi.org/10.3758/s13428-011-0088-6).

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.

Kwan et al.: Comparing squared multiple correlation coefficients using structural equation modeling **Kwan-Chan-2014**

Joyce L. Y. Kwan and Wai Chan. “Comparing squared multiple correlation coefficients using structural equation modeling”. In: *Structural Equation Modeling: A Multidisciplinary Journal* 21.2 (Apr. 2014), pp. 225–238. DOI: [10.1080/10705511.2014.882673](https://doi.org/10.1080/10705511.2014.882673).

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

Lachowicz-Preacher-Kelley-2018

Mark J. Lachowicz, Kristopher J. Preacher, and Ken Kelley. “A novel measure of effect size for mediation analysis”. In: *Psychological Methods* 23.2 (June 2018), pp. 244–261. ISSN: 1082-989X. DOI: [10.1037/met0000165](https://doi.org/10.1037/met0000165).

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.

Leffingwell et al.: Continuous Objective Monitoring of Alcohol Use: Twenty-First Century Measurement Using Transdermal Sensors Leffingwell-Cooney-Murphy-et-al-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](https://doi.org/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.

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-et-al-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](https://doi.org/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 appropri-

ate 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.

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

Scott E. Maxwell, David A. Cole, and Melissa A. Mitchell. “Bias in cross-sectional analyses of longitudinal mediation: Partial and complete mediation under an autoregressive model”. In: *Multivariate Behavioral Research* 46.5 (Sept. 2011), pp. 816–841. DOI: [10.1080/00273171.2011.606716](https://doi.org/10.1080/00273171.2011.606716).

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.

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>.

Miocevic et al.: A tutorial in Bayesian potential outcomes mediation analysis
Miocevic-Gonzalez-Valente-et-al-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](https://doi.org/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.

Molenaar: Equivalent Dynamic Models**Molenaar-2017**

Peter C. M. Molenaar. “Equivalent Dynamic Models”. In: *Multivariate Behavioral Research* 52.2 (Feb. 2017), pp. 242–258. ISSN: 1532-7906. DOI: [10.1080/00273171.2016.1277681](https://doi.org/10.1080/00273171.2016.1277681).

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.

Moneta et al.: Causal search in structural vector autoregressive models**Moneta-Chlas-Entner-et-al-2011**

Alessio Moneta et al. “Causal search in structural vector autoregressive models”. In: *Journal of Machine Learning Research - Proceedings Track* 12 (Jan. 2011), pp. 95–114.

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.

Neale et al.: OpenMx 2.0: Extended structural equation and statistical modeling

Neale-Hunter-Pritikin-et-al-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](https://doi.org/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.

Northcote et al.: Accuracy of self-reported drinking: Observational verification of ‘last occasion’ drink estimates of young adults

Northcote-Livingston-2011

Jeremy Northcote and Michael Livingston. “Accuracy of self-reported drinking: Observational ver-

ification of ‘last occasion’ drink estimates of young adults”. In: *Alcohol and Alcoholism* 46.6 (Sept. 2011), pp. 709–713. ISSN: 0735-0414. DOI: [10.1093/alcalc/agr138](https://doi.org/10.1093/alcalc/agr138).

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.

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

OLaughlin-Martin-Ferrer-2018

Kristine D. O’Laughlin, Monica J. Martin, and Emilio Ferrer. “Cross-sectional analysis of longitudinal mediation processes”. In: *Multivariate Behavioral Research* 53.3 (Apr. 2018), pp. 375–402. ISSN: 1532-7906. DOI: [10.1080/00273171.2018.1454822](https://doi.org/10.1080/00273171.2018.1454822).

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**

ORourke-MacKinnon-2018

Holly P. O’Rourke and David P. MacKinnon. “Reasons for testing mediation in the absence of an intervention effect: A research imperative in prevention and intervention research”. In: *Journal of Studies on Alcohol and Drugs* 79.2 (Mar. 2018), pp. 171–181. DOI: [10.15288/jsad.2018.79.171](https://doi.org/10.15288/jsad.2018.79.171).

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.

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

Zita Oravecz, Francis Tuerlinckx, and Joachim Vandekerckhove. “A hierarchical latent stochastic differential equation model for affective dynamics”. In: *Psychological Methods* 16.4 (2011), pp. 468–490. DOI: [10.1037/a0024375](https://doi.org/10.1037/a0024375).

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 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).

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](https://doi.org/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 discrete and 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.

Pastor et al.: On the multilevel nature of meta-analysis: A tutorial, comparison of software programs, and discussion of analytic choices **Pastor-Lazowski-2017**

Dena A. Pastor and Rory A. Lazowski. “On the multilevel nature of meta-analysis: A tutorial, comparison of software programs, and discussion of analytic choices”. In: *Multivariate Behavioral Research* 53.1 (Sept. 2017), pp. 74–89. ISSN: 1532-7906. DOI: [10.1080/00273171.2017.1365684](https://doi.org/10.1080/00273171.2017.1365684).

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.

**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](https://doi.org/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 ($P \leq 0.001$) in these behaviors from ages 18 to 19/20 was observed. Significant decreases ($P \leq 0.05$) 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 ($P \leq 0.001$). 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.

Pesigan et al.: Brief report: Identity processes in Filipino late adolescents and young adults: Parental influences and mental health outcomes

Pesigan-Luyckx-Alampay-2014

Ivan Jacob Agaloos Pesigan, Koen Luyckx, and Liane Peña Alampay. “Brief report: Identity processes in Filipino late adolescents and young adults: Parental influences and mental health outcomes”. In: *Journal of Adolescence* 37.5 (May 2014), pp. 599–604. ISSN: 1095-9254. DOI: [10.1016/j.adolescence.2014.04.012](https://doi.org/10.1016/j.adolescence.2014.04.012).

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.

Piasecki: Assessment of alcohol use in the natural environment**Piasecki-2019**

Thomas M. Piasecki. “Assessment of alcohol use in the natural environment”. In: *Alcoholism: Clinical and Experimental Research* 43.4 (Mar. 2019), pp. 564–577. ISSN: 1530-0277. DOI: [10.1111/acer.13975](https://doi.org/10.1111/acer.13975).

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.

Preacher et al.: Effect size measures for mediation models: Quantitative strategies for communicating indirect effects **Preacher-Kelley-2011**

Kristopher J. Preacher and Ken Kelley. "Effect size measures for mediation models: Quantitative strategies for communicating indirect effects". In: *Psychological Methods* 16.2 (2011), pp. 93–115. ISSN: 1082-989X. DOI: [10.1037/a0022658](https://doi.org/10.1037/a0022658).

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](https://doi.org/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](https://doi.org/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.

Reichardt: Commentary: Are three waves of data sufficient for assessing mediation?

Reichardt-2011

Charles S. Reichardt. "Commentary: Are three waves of data sufficient for assessing mediation?" In: *Multivariate Behavioral Research* 46.5 (Sept. 2011), pp. 842–851. ISSN: 1532-7906. DOI: [10.1080/00273171.2011.606740](https://doi.org/10.1080/00273171.2011.606740).

Abstract: Maxwell, Cole, and Mitchell (2011) demonstrated that simple structural equation models, when used with cross-sectional data, generally produce biased estimates of mediated effects. I extend those results by showing how simple structural equation models can produce biased estimates of mediated 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.

John D. Roache et al. "Processing transdermal alcohol concentration (TAC) data to detect low-level drinking". In: *Alcohol* 81 (Dec. 2019), pp. 101–110. ISSN: 0741-8329. DOI: [10.1016/j.alcohol.2018.08.014](https://doi.org/10.1016/j.alcohol.2018.08.014).

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.

Yves Rosseel. “lavaan: An R package for structural equation modeling”. In: *Journal of Statistical Software* 48.2 (2012). DOI: [10.18637/jss.v048.i02](https://doi.org/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.

Sacks et al.: 2010 national and state costs of excessive alcohol consumption

Sacks-Gonzales-Bouchery-et-al-2015

Jeffrey J. Sacks et al. “2010 national and state costs of excessive alcohol consumption”. In: *American Journal of Preventive Medicine* 49.5 (Nov. 2015), e73–e79. ISSN: 0749-3797. DOI: [10.1016/j.amepre.2015.05.031](https://doi.org/10.1016/j.amepre.2015.05.031).

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 2013, we estimated the total cost of excessive drinking (e.g., inflation rate in cost of medical care). The total cost, cost to government, and costs for binge 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; underaged 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 five 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.

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](https://doi.org/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

Rianne Margaretha Schouten, Peter Lugtig, and Gerko Vink. “Generating missing values for simulation purposes: A multivariate amputation procedure”. In: *Journal of Statistical Computation and Simulation* 88.15 (July 2018), pp. 2909–2930. DOI: [10.1080/00949655.2018.1491577](https://doi.org/10.1080/00949655.2018.1491577).

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.

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](https://doi.org/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-et-al-2016**

Noémi K. Schuurman, Emilio Ferrer, et al. “How to compare cross-lagged associations in a multilevel autoregressive model”. In: *Psychological Methods* 21.2 (2016), pp. 206–221. ISSN: 1082-989X. DOI: [10.1037/met0000062](https://doi.org/10.1037/met0000062).

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.

Schuurman et al.: Measurement error and person-specific reliability in multilevel autoregressive modeling **Schuurman-Hamaker-2019**

Noémi K. Schuurman and Ellen L. Hamaker. “Measurement error and person-specific reliability in multilevel autoregressive modeling”. In: *Psychological Methods* 24.1 (Feb. 2019), pp. 70–91. ISSN: 1082-989X. DOI: [10.1037/met0000188](https://doi.org/10.1037/met0000188).

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.

Noémi K. Schuurman, Jan H. Houtveen, and Ellen L. Hamaker. “Incorporating measurement error in $n = 1$ psychological autoregressive modeling”. In: *Frontiers in Psychology* 6 (July 2015). ISSN: 1664-1078. DOI: [10.3389/fpsyg.2015.01038](https://doi.org/10.3389/fpsyg.2015.01038).

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.

Shrout: Commentary: Mediation analysis, causal process, and cross-sectional data

Shrout-2011

Patrick E. Shrout. “Commentary: Mediation analysis, causal process, and cross-sectional data”. In: *Multivariate Behavioral Research* 46.5 (Sept. 2011), pp. 852–860. DOI: [10.1080/00273171.2011.606718](https://doi.org/10.1080/00273171.2011.606718).

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.

Simons et al.: Quantifying alcohol consumption: Self-report, transdermal assessment, and prediction of dependence symptoms **Simons-Wills-Emery-et-al-2015**

Jeffrey S. Simons et al. “Quantifying alcohol consumption: Self-report, transdermal assessment, and prediction of dependence symptoms”. In: *Addictive Behaviors* 50 (Nov. 2015), pp. 205–212. ISSN: 0306-4603. DOI: [10.1016/j.addbeh.2015.06.042](https://doi.org/10.1016/j.addbeh.2015.06.042).

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**

Hermann Singer. “SEM modeling with singular moment matrices part II: ML-estimation of sampled stochastic differential equations”. In: *The Journal of Mathematical Sociology* 36.1 (Jan. 2012), pp. 22–43. ISSN: 1545-5874. DOI: [10.1080/0022250x.2010.532259](https://doi.org/10.1080/0022250x.2010.532259).

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.

Zsuzsika Sjoerds et al. "Behavioral and neuroimaging evidence for overreliance on habit learning in alcohol-dependent patients". In: *Translational Psychiatry* 3.12 (Dec. 2013), e337–e337. ISSN: 2158-3188. DOI: [10.1038/tp.2013.107](https://doi.org/10.1038/tp.2013.107).

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.

Kathryn E. Smith and Adrienne Juarascio. "From ecological momentary assessment (EMA) to ecological momentary intervention (EMI): Past and future directions for ambulatory assessment and interventions in eating disorders". In: *Current Psychiatry Reports* 21.7 (June 2019). DOI: [10.1007/s11920-019-1046-8](https://doi.org/10.1007/s11920-019-1046-8).

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.

Hairong Song and Emilio Ferrer. “Bayesian estimation of random coefficient dynamic factor models”. In: *Multivariate Behavioral Research* 47.1 (Feb. 2012), pp. 26–60. ISSN: 1532-7906. DOI: [10.1080/00273171.2012.640593](https://doi.org/10.1080/00273171.2012.640593).

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.

A. B. Taylor et al.: Four applications of permutation methods to testing a single-mediator model

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](https://doi.org/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.

Tibshirani: Regression shrinkage and selection via the lasso: A retrospective

Tibshirani-2011

Robert Tibshirani. "Regression shrinkage and selection via the lasso: A retrospective". In: *Journal of the Royal Statistical Society Series B: Statistical Methodology* 73.3 (Apr. 2011), pp. 273–282. ISSN: 1467-9868. DOI: [10.1111/j.1467-9868.2011.00771.x](https://doi.org/10.1111/j.1467-9868.2011.00771.x).

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

Davood Tofighi and Ken Kelley. "Indirect effects in sequential mediation models: Evaluating meth-

ods for hypothesis testing and confidence interval formation”. In: *Multivariate Behavioral Research* 55.2 (June 2019), pp. 188–210. DOI: [10.1080/00273171.2019.1618545](https://doi.org/10.1080/00273171.2019.1618545).

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.: Monte Carlo confidence intervals for complex functions of indirect effects

Tofighi-MacKinnon-2015

Davood Tofighi and David P. MacKinnon. “Monte Carlo confidence intervals for complex functions of indirect effects”. In: *Structural Equation Modeling: A Multidisciplinary Journal* 23.2 (Aug. 2015), pp. 194–205. DOI: [10.1080/10705511.2015.1057284](https://doi.org/10.1080/10705511.2015.1057284).

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.

Usami et al.: A unified framework of longitudinal models to examine reciprocal relations **Usami-Murayama-Hamaker-2019**

Satoshi Usami, Kou Murayama, and Ellen L. Hamaker. “A unified framework of longitudinal models to examine reciprocal relations”. In: *Psychological Methods* 24.5 (Oct. 2019), pp. 637–657. ISSN: 1082-989X. DOI: [10.1037/met0000210](https://doi.org/10.1037/met0000210).

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.

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](https://doi.org/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.

Jay M. Ver Hoef. “Who invented the delta method?” In: *The American Statistician* 66.2 (May 2012), pp. 124–127. ISSN: 1537-2731. DOI: [10.1080/00031305.2012.687494](https://doi.org/10.1080/00031305.2012.687494).

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.

Voelkle et al.: Continuous time modelling with individually varying time intervals for oscillating and non-oscillating processes **Voelkle-Oud-2012**

Manuel C. Voelkle and Johan H. L. Oud. “Continuous time modelling with individually varying time intervals for oscillating and non-oscillating processes”. In: *British Journal of Mathematical and Statistical Psychology* 66.1 (Mar. 2012), pp. 103–126. DOI: [10.1111/j.2044-8317.2012.02043.x](https://doi.org/10.1111/j.2044-8317.2012.02043.x).

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.

Manuel C. Voelkle, Johan H. L. Oud, et al. “An SEM approach to continuous time modeling of panel data: Relating authoritarianism and anomia”. In: *Psychological Methods* 17.2 (2012), pp. 176–192. DOI: [10.1037/a0027543](https://doi.org/10.1037/a0027543).

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.

Matti Vuorre and Niall Bolger. “Within-subject mediation analysis for experimental data in cognitive psychology and neuroscience”. In: *Behavior Research Methods* 50.5 (Dec. 2017), pp. 2125–2143. ISSN: 1554-3528. DOI: [10.3758/s13428-017-0980-9](https://doi.org/10.3758/s13428-017-0980-9).

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 psy-

chology, 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 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.

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](https://doi.org/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.

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](https://doi.org/10.1159/000441458).

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](https://doi.org/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.

Yu et al.: Psychometric validation of the Internet Gaming Disorder-20 Test among Chinese middle school and university students
Yu-Pesigan-Zhang-et al-2019

Shu M. Yu, Ivan Jacob Agaloos Pesigan, et al. “Psychometric validation of the Internet Gaming

Disorder-20 Test among Chinese middle school and university students”. In: *Journal of Behavioral Addictions* 8.2 (May 2019), pp. 295–305. ISSN: 2063-5303. DOI: [10.1556/2006.8.2019.18](https://doi.org/10.1556/2006.8.2019.18).

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 ($n = 569$; $M_{\text{age}} = 13.34$; 46.2% females) and university students ($n = 523$; $M_{\text{age}} = 20.12$; 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.

Yu et al.: Cognitive and psychosocial health risk factors of social networking addiction

Yu-Wu-Pesigan-2015

Shu M. Yu, Anise Man Sze Wu, and Ivan Jacob Agaloos Pesigan. “Cognitive and psychosocial health risk factors of social networking addiction”. In: *International Journal of Mental Health and Addiction* 14.4 (Dec. 2015), pp. 550–564. ISSN: 1557-1882. DOI: [10.1007/s11469-015-9612-8](https://doi.org/10.1007/s11469-015-9612-8).

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.

Yuan et al.: Biases and standard errors of standardized regression coefficients

Yuan-Chan-2011

Ke-Hai Yuan and Wai Chan. “Biases and standard errors of standardized regression coefficients”. In: *Psychometrika* 76.4 (Aug. 2011), pp. 670–690. DOI: [10.1007/s11336-011-9224-6](https://doi.org/10.1007/s11336-011-9224-6).

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 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.

**Yzerbyt et al.: New recommendations for testing indirect effects in mediational models:
The need to report and test component paths** **Yzerbyt-Muller-Batailler-et-al-2018**

Vincent Yzerbyt et al. “New recommendations for testing indirect effects in mediational models: The need to report and test component paths”. In: *Journal of Personality and Social Psychology* 115.6 (Dec. 2018), pp. 929–943. DOI: [10.1037/pspa0000132](https://doi.org/10.1037/pspa0000132).

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**

Zamboanga-Iwamoto-Pesigan-et-al-2015

Byron L. Zamboanga, Derek K. Iwamoto, et al. “A “player’s” game? Masculinity and drinking games participation among White and Asian American male college freshmen”. In: *Psychology of Men & Masculinity* 16.4 (Oct. 2015), pp. 468–473. ISSN: 1524-9220. DOI: [10.1037/a0039101](https://doi.org/10.1037/a0039101).

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 ($M_{\text{age}} = 18.11$ 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?

Zamboanga-Pesigan-Tomaso-et-al-2015

Byron L. Zamboanga, Ivan Jacob Agaloos Pesigan, et al. "Frequency of drinking games participation and alcohol-related problems in a multiethnic sample of college students: Do gender and ethnicity matter?" In: *Addictive Behaviors* 41 (Feb. 2015), pp. 112–116. ISSN: 0306-4603. DOI: [10.1016/j.addbeh.2014.10.002](https://doi.org/10.1016/j.addbeh.2014.10.002).

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.

G. Zhang: Testing process factor analysis models using the parametric bootstrap

Zhang-2018

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](https://doi.org/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.

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](https://doi.org/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 factors. 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.

Meng Xuan Zhang, Lisbeth Ku, et al. “Effects of social and outcome expectancies on hazardous drinking among Chinese university students: The mediating role of drinking motivations”. In: *Substance Use & Misuse* 55.1 (Sept. 2019), pp. 156–166. ISSN: 1532-2491. DOI: [10.1080/10826084.2019.1658784](https://doi.org/10.1080/10826084.2019.1658784).

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-et-al-2019

Meng Xuan Zhang, Ivan Jacob Agaloos Pesigan, et al. "Psychometric properties of a Chinese version of the Brief Young Adult Alcohol Consequences Questionnaire (B-YAACQ)". In: *Addictive Behaviors* 90 (Mar. 2019), pp. 389–394. ISSN: 0306-4603. DOI: [10.1016/j.addbeh.2018.11.045](https://doi.org/10.1016/j.addbeh.2018.11.045).

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%; $M_{\text{age}} = 19.88$) 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.

Z. Zhang et al.: Methods for mediation analysis with missing data Zhang-Wang-2012

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](https://doi.org/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.

Zyphur et al.: From data to causes I: Building a general cross-lagged panel model (GCLM) **Zyphur-Allison-Tay-et-al-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](https://doi.org/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.

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](https://doi.org/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.