Uncovering General, Shared, and Unique Temporal Patterns in Ambulatory Assessment Data

Stephanie T. Lane, Kathleen M. Gates,
and Hallie K. Pike
University of North Carolina at Chapel Hill

Adriene M. Beltz
University of Michigan

Aidan G. C. Wright
University of Pittsburgh

Abstract

Intensive longitudinal data provide psychological researchers with the potential to better understand individual-level temporal processes. While the collection of such data has become increasingly common, there are a comparatively small number of methods well-suited for analyzing these data, and many methods assume homogeneity across individuals. A recent development rooted in structural equation and vector autoregressive modeling, Subgrouping Group Iterative Multiple Model Estimation (S-GIMME), provides one method for arriving at individual-level models composed of processes shared by the sample, a subset of the sample, and a given individual. As this algorithm was motivated and validated for use with neuroimaging data, its performance is less understood in the context of ambulatory assessment data. Here, we evaluate the performance of the S-GIMME algorithm across various conditions frequently encountered with daily diary (compared to neuroimaging) data; namely, a smaller number of variables, a lower number of time points, and smaller autoregressive effects. We demonstrate, for the first time, the importance of the autoregressive effects in recovering data-generating connections and directions, and the ability to use S-GIMME with lengths of data commonly seen in daily diary studies. We demonstrate the use of S-GIMME with an empirical example evaluating the general, shared, and unique temporal processes associated with a sample of individuals with borderline personality disorder (BPD). Finally, we underscore the need for methods such as S-GIMME moving forward given the increasing use of intensive longitudinal data in psychological research, and the potential for these data to provide novel insights into human behavior and mental health.

Translational Abstract

Psychological researchers increasingly collect intensive longitudinal data. This kind of research involves a shift away from a more traditional approach of collecting a few occasions of data from many persons in order to inform sample-level conclusions, and instead entails collecting many occasions of data from a smaller number of persons in order to understand individual-level processes. With this wealth of data, researchers face a challenge when selecting an analysis method to meaningfully parse heterogeneity across individuals. In our study, we demonstrate the use of a method rooted in the structural equation modeling framework, the Subgrouping Group Iterative Multiple Model Estimation (S-GIMME) algorithm. Here, we evaluate its performance using a Monte Carlo simulation study, in which we generate data consistent with characteristics often encountered by applied researchers. Our goal is to understand the algorithm's ability to identify data-generating subgroups and recover processes at the level of the entire sample, the level of the subgroups, and the level of the individual. We follow this with an empirical data analysis, in which we uncover relationships between symptom data over time in a sample of individuals with borderline personality disorder. We demonstrate, for the first time, the adequacy of the S-GIMME algorithm for use with daily diary data. More specifically, we show that it provides robust results with as few as 25 individuals for whom 60 time points of data have been collected. Our demonstration underscores the need for methods such as S-GIMME moving forward in order to uncover novel insights into human behavior and mental health.

Keywords: time series analysis, idiographic methods, structural equation modeling, daily diary data

It is now well recognized that individuals comprising a priori categories, such as gender, age, or psychiatric diagnosis, are remarkably heterogeneous in their characteristic manners of mood, behavior, or cognition. Presumably, this rich variability within a given category reveals individual differences (perhaps unrelated to category) that undermine the validity of the categorization. Indeed, this heterogeneity poses challenges to psychological and behavioral classification systems, including those used in high-stakes settings (e.g., psychiatric diagnosis and treatment). Additionally, these categories have problematically nebulous boundaries, and supposedly discrete groups may overlap substantially. A prime example of these issues can be found in psychiatric diagnoses, which are paradoxically both too specific (i.e., high rates of covariation or "co-morbidity"; Lilienfeld, Waldman, & Israel, 1994) and too general (i.e., high heterogeneity within diagnosis; Hyman, 2010). These challenges have generated a broad and growing literature on quantitative modeling approaches to uncover the symptomatic patterns of psychopathology (for review, see, e.g., Kotov et al., 2017). The present article evaluates one analytic approach for ambulatory data analysis, Subgrouping Group Iterative Multiple Model Estimation (S-GIMME; Gates, Lane, Varangis, Giovanello, & Guiskewicz, 2017), for the modeling of emotional, behavioral, and cognitive processes. S-GIMME arrives at general patterns of relations that exist for the entire sample, patterns that are shared by only a subset of individuals, and patterns that are unique to a given individual. Originally developed for use on functional brain imaging data, we evaluate the utility of this approach for data qualities that are sometimes seen in analysis on ecological momentary assessments.

To date, efforts to generate quantitatively derived taxonomies have largely relied on assessment approaches that are temporally distal from the putative mechanisms of interest; that is, indicators are assessed in cross-section or at a temporal scale (e.g., years) that miss the target underlying dynamic processes in real time that serve to define and maintain the constructs of interest (e.g., psychopathology). However, over the past two decades, there has been a surge in method development and implementation to study mood, behavior, and cognition as they unfold in the naturalistic settings of daily life. Referred to variously as experience sampling, ecological momentary assessment, or ambulatory assessment (Shiffman, Stone, & Hufford, 2008; Trull & Ebner-Priemer, 2013), these approaches generally use intensive longitudinal data collection designs to sample behavior on a time scale closer to the dynamic processes they intend to measure. Research in psychopathology increasingly relies on these methods to understand the nuanced and complex dynamic processes of mental illness (e.g., Ebner-Priemer et al., 2007; Hamaker & Wichers, 2017; Muehlenkamp et al., 2009; Myin-Germeys, van Os, Schwartz, Stone, & Delespaul, 2001; Pe et al., 2015; Sadikaj, Moskowitz, Russell, Zuroff, & Paris, 2013; Shiffman et al., 2002; Silk, Steinberg, & Morris, 2003; Smyth et al., 2007; Trull et al., 2008; Wright & Simms, 2016; see also Myin-Germeys et al., 2009, and Trull & Ebner-Priemer, 2013, for reviews). This work has almost exclusively relied on the aforementioned diagnostic classification schemes by placing individuals in predefined groups and conducting analysis under the assumption of within-group homogeneity. However, if the assumption of within-group homogeneity does not hold for a given sample, then results do not reflect the symptomatology present for individuals seeking treatment. To understand individual-level symptoms and experiences, temporal processes must be examined. Further, it has now been demonstrated that heterogeneity in temporal processes also exists in normative samples, such as the structure of personality (e.g., Borkenau & Ostendorf, 1998; Molenaar & Campbell, 2009; Ram & Grimm, 2009; Ram, Rabbitt, Stollery, & Nesselroade, 2005).

Even though diagnostic categories may not accurately reflect the symptoms or experiences of individuals, there may exist groups of individuals who share similar behavioral or cognitive profiles. Thus, research on the temporal processes of individuals would ideally be extended to parsing individuals into subgroups based on similarity in the dynamic processes underlying their behavior. Identifying these subgroups first requires quantitative methods for arriving at individual-level models that accurately and reliably capture the unique patterns of effects among variables across time (Molenaar, 2004), a goal that has gained much traction in human brain neuroimaging literature (Smith et al., 2011). From this work, Bayes net approaches—a class of algorithms for model selection—have been identified as the most promising analytic methods for arriving at data-driven models of temporal processes (Mumford & Ramsey, 2014; Ramsey, Hanson, & Glymour, 2011; Ramsey, Sanchez-Romero, & Glymour, 2014). Among those, Group Iterative Multiple Model Estimation (GIMME; Gates & Molenaar, 2012) has surfaced as a highly reliable model building procedure for arriving at individual-level patterns of effects underlying dynamic processes. One recent extension to the original GIMME algorithm is the ability to cluster individuals into subgroups based on similarities in their dynamic processes. Termed "Subgrouping GIMME" (S-GIMME; Gates et al., 2017), this method provides an unsupervised classification approach to arrive at data-driven subgroups. S-GIMME thus identifies dynamic relations that exist at the group (i.e., sample), subgroup, and individual levels, providing estimates for these relations separately for each individual.

Previous work involving GIMME and S-GIMME has been rooted in modeling functional MRI (fMRI) brain data; that is, Monte Carlo simulations have focused on data generated to mimic characteristics of fMRI data, and empirical applications have focused on modeling relationships among brain regions of interest (ROIs) from fMRI data (e.g., Beltz et al., 2013; Karunanayaka et al., 2014; Nichols, Gates, Molenaar, & Wilson, 2014; Price et al., 2017; Yang, Gates, Molenaar, & Li, 2015). The present article provides evidence that, with some consideration, S-GIMME can be applied to ambulatory assessment data, of which daily diary data are a specific type. Here, daily diaries were specifically chosen because (a) they are an oft-used variant of ambulatory assessment both historically (e.g., Borkenau & Ostendorf, 1998; Lebo & Nesselroade, 1978; Searles, Perrine, Mundt, & Helzer, 1995; Zevon & Tellegen, 1982) and at present (e.g., Bachrach & Read, 2017; Castro-Schilo & Ferrer, 2013; Gadassi et al., 2016; Hamaker, Grasman, & Kamphuis, 2016; Lee et al., 2017; Zimmermann et al., in press), and (b) the sampling rate is sometimes conducted at relatively equal spacing, an advantageous feature for

time-series data in the current modeling context.[1] Nevertheless, application to daily diary data presents novel analytic concerns that may be absent in neuroimaging data.

Motivation for Simulation Study

One important difference in the qualities of the data results from the psychophysiological nature of fMRI data. In fMRI data, the extent to which a variable predicts itself at subsequent time points, also known as the autoregressive (AR) effect (Chatfield, 2003), is sizable. The magnitude of these AR effects stems from the slowness of the biological process under study relative to the temporal resolution of data collection, such that values at a given time are highly predictive of values at the same brain region at the next time point. In contrast to fMRI data, for which an AR effect concerns an ROI predicting itself at the next time point (often two seconds later), there may be little reason to expect a construct examined in a daily diary study, such as affect, to *strongly* predict itself at the next time point because changes may occur on the order of hours (or minutes) rather than days. Even when data are obtained multiple times within a day, the AR effects have been shown to vary across individuals and constructs (see, e.g., Fernandez, Fisher, & Chi, 2017; A. J. Fisher & Boswell, 2016), suggesting that the assumption of strong AR effects may not hold in these types of data.

When the rate of measurement is slow relative to the process being studied, time-lagged effects surface as contemporaneous (Granger, 1969). As described in detail below, strong contemporaneous relations can prevent the estimation of AR effects in the GIMME process, thus changing the search space from what is seen in brain imaging situations in which the AR effects are often (but not always) high. In these cases, multiple solutions may surface from within a GIMME framework that describe the data approximately equally well (Beltz & Molenaar, 2016). Accordingly, one issue that the present article seeks to investigate is whether the inclusion of AR relations in the base model circumvents this problem and provides reliable results even in the presence of relatively low AR estimates. By estimating AR effects (even if small), one can directly examine whether a given variable statistically predicts another variable—at a lag or contemporaneously—more so than the reverse (Granger, 1969). It has previously been shown in simulations emulating fMRI data that failure to include both contemporaneous and lagged effects leads to spurious results when both types of effects exist in the true generative model (Gates, Molenaar, Hillary, Ram, & Rovine, 2010). However, it remains unknown whether the inclusion of small, potentially nonsignificant, AR effects perhaps seen in daily diary data will allow for the reliable recovery of directed effects among variables.

A second difference is that daily diary data may possess fewer group-level relationships compared with fMRI data. This notion again stems directly from fMRI being a biological process, and as such, having known properties that exist for all individuals. For example, a strong connection is typically found between left and right hemispheres of the same region; further, this connection is often found across all individuals (e.g., Zelle, Gates, Fiez, Sayette, & Wilson, 2017). Thus, for data containing multiple bilateral ROIs, there will be a larger number of group-level relationships. Although daily diary data may also contain relationships common to all individuals (e.g., stress and negative affect), there may be

fewer in general because there are no biological constraints underlying group-level relations.

A third difference is that daily diary studies likely will have far fewer observations per person than fMRI studies. As with any time series analysis, the power to detect effects increases with greater numbers of observations across the time. For instance, it is rare for an fMRI study to have fewer than 150 observations, but daily dairy studies rarely collect that number of observations.

A related and final difference is the number of variables (or ROIs). Given the decrease in sample size, researchers may opt for fewer variables to model in order to ensure model convergence and to facilitate interpretation. For these reasons, the present article investigates the performance of S-GIMME on simulated behavioral data, which are expected to differ from fMRI data in the following ways: (a) different parameters governing temporal effects, (b) shorter time series, (c) fewer group-level relationships, and (d) a smaller number of variables.

Empirical Study

Following examination of simulated data, we provide evidence of heterogeneity in the temporal processes of symptoms across days in a sample of individuals diagnosed with borderline personality disorder (BPD). That is, we investigate whether patterns of relations among symptoms vary across individuals within a sample of individuals diagnosed with BPD. BPD, in particular, is a psychiatric diagnosis plagued by heterogeneity in clinical presentation (Hallquist & Pilkonis, 2012; Wright et al., 2013). As defined by the *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.; *DSM–5*; American Psychiatric Association, 2013), BPD is a syndrome composed of interpersonal (e.g., unstable and intense relationships), affective (e.g., marked reactivity of mood; difficulty controlling anger), behavioral (e.g., impulsivity; recurrent suicidal behavior), and cognitive (e.g., identity disturbance; transient psychosis) features. Thus, there is diversity in the component features of the syndrome. Further, like many *DSM–5* diagnoses, BPD is a polythetic category, such that only five of the nine criteria are necessary for diagnosis. As such, there are 256 possible combinations of features that would satisfy the threshold for BPD diagnosis. This fact, along with seminal clinical theory (Kernberg, 1975), has motivated the search for meaningful subgroupings to parse this observed heterogeneity. In turn, parsing this heterogeneity may inform more personalized treatment.

Previous efforts have enlisted cluster analysis or finite mixture modeling of various concurrently assessed diagnostic, personality, and behavioral features to empirically establish more homogeneous groups (Bradley, Zittel Conklin, & Westen, 2005; Critchfield, Clarkin, Levy, & Kernberg, 2008; Hallquist & Pilkonis, 2012; Lenzenweger, Clarkin, Yeomans, Kernberg, & Levy, 2008; Wright et al., 2013). These studies have largely used other dispositional scales to derive more homogenous subgroups but not direct assessments of *processes*, as they might unfold over time and across contexts. However, in both the *DSM–5* (American Psychiatric Association, 2013) and the broader clinical literature, the cardinal impairments of BPD are dynamic and contextualized.

---

[1] Note that equally spaced time intervals is currently a requirement of the GIMME approach. For more a more detailed discussion regarding the requirements for GIMME's use, see Beltz and Gates (2017).

For instance, the symptoms of BPD involve not only merely elevated negative affect but also "reactivity of mood," as well as frantic efforts to avoid abandonment, unstable relationships, transient psychosis, difficulty controlling anger, and impulsivity. These features are generally understood to have important dynamic relationships to each other (e.g., self-harm may serve to regulate negative emotions; Miskewicz et al., 2015); further, the individual variability present in these relations likely gives rise to the observed heterogeneity in the presentation of symptoms. To date, there has been scant investigation into the potential between-person heterogeneity in within-person couplings of BPD features over time.

Method

Definition of Unified Structural Equation Modeling (uSEM) Implemented in S-GIMME

The current article uses the S-GIMME algorithm (Gates et al., 2017; Gates & Molenaar, 2012), which reliably recovers data-generating group-, subgroup-, and individual-level relations at the level of the person. This algorithm is rooted in the uSEM framework (Kim, Zhu, Chang, Bentler, & Ernst, 2007), a variant of structural vector AR modeling that simultaneously estimates both contemporaneous and lagged relations. The uSEM capitalizes on the known ability to conduct time-series analysis from within an SEM framework, which allows one to draw on the many strengths of SEM to arrive at individual-level models with directed (not correlational) relationships among variables of interest. The inclusion of the lagged effects, including the AR effects, allows specification of the direction of a relationship from within a Granger causality framework. That is, a Variable X is said to "Granger-cause" Variable Y if it explains variance in Variable Y above and beyond the variance explained by the AR term of Y (Granger, 1969). This is typically evaluated by looking at the error variance; as a proxy, we examine whether the candidate variable (X in our example) would significantly improve model fit if it were added as a predictor to Variable Y after considering the AR effect for Variable Y. Note that this definition centers on the improvement of prediction in Variable Y when using information from Variable X; it does not represent causality in an absolute sense. Importantly, making stronger claims of causality rests in part on the assumption that all relevant variables are in the model, a known limitation of drawing claims of causality in SEM (see omitted variable problem: Bollen, 1989). Results must be interpreted as causal, with the caveat that the only putative third-variable influences were those contained in the set used and not the universe of variables.

Including the AR relations during the path search procedure in S-GIMME automatically tests for Granger causality in each step.[2] Each additional path added provides a significant improvement upon the reduced model (which, in this case, includes the AR effect for the target variable) as well as any other predictors that have been added. The interpretation of lagged relations that are added is quite straightforward: deviation from the mean at the prior time point relates to deviation from the mean at the next time point. Directed contemporaneous relations can emerge because (a) the time delay for cause and effect is small (or unmeasurable) relative to the time interval of data collection, or (b) both variables are

caused by a different unmeasured variable at a prior time point (Granger, 1988). The latter reason provides further motivation for researchers to carefully consider variables to include in the model to avoid omitting necessary third-variable influences. In many cases, it is reasonable to expect that relationships that are truly lagged in nature may emerge as contemporaneous because of the temporal resolution of the data collection. In fMRI, for instance, the biological process is slow (on the order of seconds) relative to the underlying neuronal activity it aims to capture (which occurs on the order of milliseconds). Thus, it is expected that effects are larger for contemporaneous relations. The same may occur when measurements are taken daily but the process changes hourly, or any other such situation in which the measurement rate is slower than the rate under which the process evolves.

The general uSEM may be formally defined as

$$\eta_t = A\eta_t + \phi\eta_{t-1} + \zeta_t,$$

where A is a $p \times p$ matrix containing the contemporaneous relations among p variables (with a zero diagonal to prevent contemporaneous self-prediction), ϕ is a $p \times p$ matrix containing the lagged relations among p variables with AR effects on the diagonal, η is the observed time series, and ζ contains residuals with a mean of zero and diagonal covariance matrix, assumed to be white-noise processes and therefore contain no temporal dependencies.

These relations can be further decomposed for each individual into group-, subgroup-, and individual-level relations. That is, certain relations exist for the entire sample, certain relations exist within a given subgroup k, and certain relations exist for a given individual i. This decomposition can be expressed as

$$\eta_{i,t} = (A_i + A_{i,k}^s + A_i^g)\eta_{i,t} + (\phi_i + \phi_{i,k}^s + \phi_i^g)\eta_{i,t-1} + \zeta_{i,t}.$$

Here, A, ϕ, and ζ are defined as before, where the superscripts s and g indicate that these paths exist at the subgroup and group levels, respectively. Importantly, should no subgroup division exist (i.e., all individuals are in one "subgroup"), these relations will be contained in the group-level matrix. Parameter matrices A and ϕ that lack a superscript denote matrices containing only the individual-level relations. Finally, the subscript i on all matrices indicates that each relation is estimated at the individual level, even in cases in which there are group- and subgroup-level patterns of relations.

Assumptions of S-GIMME

S-GIMME requires that the assumption of stationarity be met but does not require the assumption of ergodicity be met. We first define stationarity because it is part of an ergodicity assumption. Stationary (formally, "weakly stationary") multivariate time series data has statistical properties—such as mean, variance, and covariance—that are constant over time and only depend on the lag. Data that has a trend violates this assumptions because the mean level changes across time. This issue can be accommodated prior

---

[2] In brain imaging data, AR relations are typically the first relations opened in the model search, as they are strong. Therefore, without placing them in the null model at the start of estimation, they are usually selected shortly thereafter.

to analysis using preprocessing steps described in Beltz and Gates (2017). In a similar vein, data for which the autocovariance value (at a given lag) changes across time violates the stationarity assumption, as do data for which the covariance among variables changes across time. For these instances, the researcher may want to consider approaches that allow for these estimates to vary across time (e.g., Molenaar, Beltz, Gates, & Wilson, 2016). To be ergodic, (a) a process must be stationary, and (b) each unit (here, person) must have the same dynamic processes (Molenaar, 2008). Although S-GIMME requires that each unit of data (however defined) be stationary, it does not require that all individuals have the same process: Individuals can vary in their processes. The rate of recovery of data-generating relations is improved when some relations are consistently found across individuals, as this helps to detect signal from noise, but the entire process does not have to be the same either in terms of the pattern of relations or the estimates of those relations. Two other noteworthy assumptions related to any SEM analysis also exist for the S-GIMME framework: that the data are normally distributed and have equal intervals between them.

Model Selection Procedure

Group-level search. The ultimate goal of GIMME (and S-GIMME) is to allow for individual-level estimates and relations in order to attend to heterogeneity in dynamic processes. However, GIMME begins with a group-level search for two related reasons. First, when using model selection procedures for SEM, the recovery of individual-level relationships is greatly improved when the starting model is closer to the data-generating model (MacCallum, 1986). This follows from suggestions that model search procedures work best when starting from a point in the model formation that is as close as possible to the final model (MacCallum, Roznowski, & Necowitz, 1992). Second, there likely is noise in any process being studied, and looking for similarities in relations across individuals aids in detecting signal from noise. Taken together, by looking for consistencies across individuals, GIMME arrives at group-level paths that greatly aid in accurately recovering individual-level relations. Full details of the model selection procedure can be found in Gates et al. (2017) and Gates and Molenaar (2012). Here, we briefly describe the relevant steps. The model selection procedure is implemented in the freely distributed R package, gimme (Lane, Gates, & Molenaar, 2014; Lane & Gates, 2017), which allows for both traditional GIMME and S-GIMME to be implemented. For each individual, the model search procedure begins with a null model; that is, all values in the A matrix and ϕ matrix are zero. Optionally, a researcher may choose to have this null model estimate the diagonal of the ϕ matrix (i.e., the AR effects) rather than initially constrain them to be zero. Although it is standard (and the default in gimme) to have the AR elements freely estimated in the null model, it is ultimately the decision of the researcher. For this reason, investigation into best practices is warranted here.

Upon estimation of the null model for each individual, S-GIMME examines the modification index (MI) associated with each fixed parameter. MIs are Lagrange Multiplier equivalents tests that indicate the anticipated improvement in model fit should that parameter be freely estimated (Jöreskog & Sörbom, 1986). There are unique MI estimates for all of the elements in the ϕ

matrix (when AR effects are not included initially) and for the estimates in the A matrix (excluding the diagonal). For each of these estimates, GIMME counts the number of individuals for whom it is significant after a Bonferroni correction of $\alpha = .05/N$. From these counts, the MI that is significant for the largest proportion of individuals is selected for inclusion in each individual's model (i.e., considered a "group-level" relation). The selection of group-level paths terminates when no path is significant for a prespecified proportion of individuals. Informed by prior research and simulations, the default cutoff is 75% (Gates & Molenaar, 2012; Smith et al., 2011). Although 75% is a strict cutoff for what constitutes a majority, this cutoff aids in the subsequent search for paths that may exist at a subgroup level. The researcher can adjust this value if a more or less strict criterion is desired. A group-level pruning procedure is conducted following the group-level search, in which paths are removed sequentially that are no longer significant for 75% of individuals at a Bonferroni-corrected $\alpha = .05/N$.

Subgroup-level search. Following the identification of the group-level model, S-GIMME searches for the existence of subgroups. These subgroups are defined based on shared characteristics of individuals' temporal processes. Specifically, an adjacency matrix is created that counts the number of similar effects shared by each dyad of two individuals, i and j. Here, similarity is defined using the correspondence between individuals' already-estimated values contained in the A and ϕ matrices (collectively referred to here as β) as well as the correspondence of the expected parameter change (EPC) associated with each potential path that could be added to the model. The EPC for a given parameter is preferred to the MI because the EPC also provides an indication of the direction of the effect (should it be freely estimated). If both the significance and sign of individuals i and j match for a given β or EPC value, the similarity count increases by 1. In this way, the adjacency matrix that forms the basis for subgrouping is constructed. This matrix is adjusted by the lowest similarity count to increase the sparseness of the matrix. Generation of an appropriate similarity matrix is highly critical for accurate clustering of time series data (Liao, 2005), and prior work indicated that the count matrix constructed here is optimal when compared with other feature-selection techniques, such as label propagation, Infomap, and fast modularity (Gates et al., 2017).

The community detection procedure known as Walktrap is then applied to this adjacency matrix (Pons & Latapy, 2006). Briefly, as a random walk approach, Walktrap performs short random walks, generally with three to five steps, and merges communities in a bottom-up fashion using traditional Ward's clustering (Ward, 1963). Then, a quality function called *modularity* (Newman, 2004) is used to identify the optimal cut point in the dendrogram. In this way, the researcher does not need to specify a priori the number of subgroups or subjectively decide given the options for cluster solutions. Full details of this procedure can be found in Pons and Latapy (2006). A benefit of Walktrap is that it is one of the few approaches that reliably recovers cluster (i.e., subgroup or community) assignments on data that are in the form of a dense correlation matrix or a count matrix as used in S-GIMME (Gates, Henry, Steinley, & Fair, 2016). Once subgroup membership is established, a search is conducted within each subgroup using the same procedure as the group-level search. Here, the starting model is now the previously established group-level model. This search terminates when no path is significant for the majority of individ-

uals (here, 51%) in that subgroup. This cutoff can also be adjusted by the researcher. As before, a subgroup-level pruning procedure is conducted, in which paths are removed sequentially that are no longer significant for 51% of individuals at a Bonferroni-corrected $\alpha = .05/N$, where N represents the number of individuals in the sample.

Individual-level search. By this point in the search procedure, all individuals enter the individual-level search with a model composed of (potentially) group- and (potentially) subgroup-level paths. The individual-level search adds paths that are significant for an individual at $\alpha = .01$ until an "excellent" model is obtained. Here, an excellent fit is obtained when two of four fit indices are excellent (RMSEA <.05, SRMR <.05, CFI >.95, NNFI >.95; Brown, 2006). An individual's final model is composed of (potentially) group-, (potentially) subgroup-, and individual-level paths, all of which are estimated at the individual level. Importantly, if a sample is so heterogeneous that no paths exist at the group or subgroup level, then the final model for an individual will be entirely composed of paths unique to that individual.

Estimation. In order to represent the lagged and contemporaneous relationships, a block Toeplitz structure is used, enabling estimates, which can be considered quasi-maximum likelihood estimates. The block Toeplitz covariance matrix is structured to represent the covariance among lag-0 variables, the covariance among lag-1 variables, and the covariance among lag-0 and lag-1 variables. Importantly, assuming normality, these parameters approximate true ML estimates (Hamaker, Dolan, & Molenaar, 2002). Finally, the default estimation method is full information maximum likelihood (Enders & Bandalos, 2001), which allows for the presence of missing data under the assumption that they are missing at random.

Stimulation Study: Monte Carlo Simulation and Evaluation Criteria

Data generation. We designed a Monte Carlo simulation consistent with empirical daily diary data. The intentions of this simulation were to evaluate the reliability of the recovery of paths with and without the AR effects across common conditions. Here, in the data-generating process, we varied (a) the number of time points, T, to be 30, 60, 90, and 120; (b) the number of variables, p, to be 5 and 10; and (c) the number of individuals, N, to be 25, 75, and 150. When applying S-GIMME to these data, two conditions were tested – AR fixed and AR freed. This yielded a fully factorial design, with $4 \times 2 \times 3 \times 2$ conditions. These conditions were evaluated across $R = 100$ replications. The data were generated using an algebraic manipulation of the formula depicted in the equation below.

$$\eta_{i,t,k} = (I_p - (A_i + A_{i,k}^s + A_i^g))^{-1}(\phi_i + \phi_{i,k}^s + \phi_i^g)\eta_{i,t-1}$$
$$+ (I_p - (A_i + A_{i,k}^s + A_i^g))^{-1}\zeta_{i,t}.$$

The values for the AR paths were set to 0.2 ($SD = .1$); this value reflects that AR processes may not be as strong in daily diary data as in fMRI data, with prior simulation work placing them at 0.6 (Gates & Molenaar, 2012). Therefore, for ambulatory data that has higher AR estimates (perhaps due to shorter time intervals, as seen in A. J. Fisher & Boswell, 2016), prior work supports that accurate model recovery should be high. The path weights for the off-

diagonal elements in ϕ were set to -0.4 ($SD = 0.1$), and the path weights for the off-diagonal elements in A were set to 0.4 ($SD = 0.1$). The errors were generated to be Gaussian white-noise errors, that is, normally distributed for all variables and all individuals. All individual time series were checked for stability over time, such that the mean and variance of the time series remained relatively constant.

To prevent the results from being driven by a specific network structure (e.g., one that is particularly easy or difficult to recover), we generated each set of paths for each replication randomly, holding constant the overall density of the network structure at approximately 20%. Of the paths generated for each individual, 25% percent were generated to exist at the group level (not including AR effects), 50% were generated to exist at the subgroup level, and 25% were generated to exist at the individual level. Group- and subgroup-level paths were varied across replications but not within replications; individual-level paths were varied randomly both across *and* within replication. For each simulated time series, a series of $T + 50$ observations were generated, where the first 50 observations were discarded to remove deviations due to initialization of the time series.

Hubert-Arabie adjusted Rand index. In order to evaluate the accuracy of subgroup recovery, we computed the Hubert-Arabie adjusted Rand index (ARI_{HA}; Hubert & Arabie, 1985). It is considered an adjusted index because it corrects for the grouping of elements by chance. This index compares the vector of recovered subgroup assignment with the vector of true subgroup assignment, in which values closer to 1 indicate better recovery of the true subgroup structure. The ARI_{HA} can be formally expressed as

$$ARI_{HA} = \frac{\binom{N}{2}(a + d) - [(a + b)(a + c) + (c + d)(b + d)]}{\binom{N}{2}^2 - [(a + b)(a + c) + (c + d)(b + d)]},$$

where N indicates the number of individuals, a indicates the number of pairs assigned to the same community who truly belonged in the same community, b indicates the number of pairs assigned to the different communities but truly belonged to the same community, c indicates the number of pairs assigned to the same community but truly belonged in different communities, and d indicates the number of pairs assigned to different communities that truly belonged in different communities. Perfect recovery results in an ARI_{HA} of 1, in which excellent recovery is >.90, good recovery is >.8, moderate recovery is >.65, and poor recovery is <.65 (Steinley, 2004).

Recovery indices. To evaluate the accuracy of the path recovery, we computed four indices popularized by Ramsey et al. (2011), which evaluate both the *presence* and the *direction* of the paths in the true and recovered individual-level models: path recall, path precision, direction recall, and direction precision. Here, recall pertains to the proportion of true paths (or directions) in the recovered model relative to the paths (or directions) in the data-generating model. Thus, recall evaluates S-GIMME's power to find relations that existed in the data-generating model. Direction recall takes into account both the detection of the presence of a path as well as the directionality, and as such, is always lower than path recall. Neither the path recall nor the direction recall takes into account recovered paths that did not exist in the data-generating model (false positives). Therefore, we also evaluate

precision, which is a proportion of the true paths (or directions) in the recovered model relative to the total paths (or directions) in the recovered model. Using these pieces of information, we are able to evaluate the algorithm's ability to accurately recover an individual's data-generating model (i.e., sensitivity), while balancing both false positives and false negatives (specificity).

Metamodels. We estimated a series of four-factor general linear models to examine mean differences in the path recovery using the aforementioned measures of recall and precision. Each GLM was estimated using all main effects (i.e., AR open or closed, number of variables, sample size, and length of observations) and two-, three-, and four-way interactions. Any effects with η_p^2 exceeding .01 were examined further.

Empirical Study

As a final step, we applied S-GIMME to an available clinical data set to demonstrate its performance in real-world empirical daily diary data. Our goal was to examine potential group, subgroup, and individual-specific paths among cardinal BPD manifestations at the daily level. BPD was selected as a target for these investigations because it is a highly heterogeneous diagnosis, by definition, and this heterogeneity has been studied using traditional dispositional assessments in cross-section (e.g., Hallquist & Pilkonis, 2012; Wright et al., 2013). For the current example, we drew the data from a study of the processes involved in the daily manifestation of personality disorder (see, e.g., Wright, Beltz, Gates, Molenaar, & Simms, 2015; Wright, Hopwood, & Simms, 2015; Wright & Simms, 2016). Wright and Simms (2016) provide a detailed description of the method and procedure involved in data collection. In brief, a large group of participants ($N = 628$) was recruited from outpatient psychiatric treatment settings and received diagnostic clinical interviews. Subsequently, a subset of those participants ($n = 116$) who met the criterion for any personality disorder were enrolled in the daily diary study, provided they also had daily access to the Internet. Written informed consent was obtained prior to participation. The relevant institutional review board approved all study procedures. Participants attended an in-person training and assessment session during which study procedures were explained and self-report measures were completed via computer. Starting in the evening of the in-person assessment, participants began completing daily diaries via secure website every evening for 100 consecutive days. Surveys were to be completed at (roughly) the same time each day, between 8:00 p.m. and 12:00 a.m. However, participants were allowed to deviate from this schedule if necessary (e.g., working nightshift), provided (a) they completed diaries at the end of their day, and (b) the diaries were completed at roughly the same time each day. Participants received daily e-mail reminders and were also provided several paper diaries they could use in the event of technological difficulties. Compensation was provided for daily participation at the rate of $100 for ≥80% participation, and prorated at $1/day for <80%. Participation was also incentivized though recurring raffles ($10 drawing every 5 days for those providing at least four diaries) and drawings for additional money and tablet computers at the end of the study, with the odds of winning proportionally tied to participation. Of the total number of participants in the daily diary study, we selected 36 who met the threshold for a BPD diagnosis and had ≥60 days' worth of daily diary responses.

Daily personality disorder manifestations. Daily expression of BPD was measured using 30 items created for this project. Details related to full item set and scale development can be found in Wright and Simms's (2016) supplemental materials. Daily items included the stem, "Over the past 24 hours . . ." and were rated on an 8-point response scale for each item anchored with *not at all* (0) and *very much so* (7). Of the 30 items, seven were selected for being (a) core features of the BPD construct, and (b) exhibiting enough variability across the majority of participants to be suitable for S-GIMME analyses. These items included measures of mood lability ("My mood was up and down"), anxiousness ("I felt anxious"), depression ("I felt depressed"), anger ("I lost my temper"), impulsivity ("I did something on impulse"), emptiness ("My relationships felt empty"), and urgency ("I acted on my emotions"). One participant was removed due to no variability on the urgency item. Thus, the final analyses used a total of 35 participants.

MANOVA. We conducted a one-way MANOVA to evaluate whether the average level of symptoms across the days used in the present study differed across the subgroups obtained from S-GIMME. We wished to examine whether the subgroups based on temporal processes differed in terms of symptom severity. As such, we conducted a MANOVA to examine potential between-subgroup differences in symptom severity using the same variables used in the analysis of temporal processes (described below): participant's mood lability, impulsivity, anger, anxiousness, depression, emptiness, and urgency.

Results

Convergence

Across all simulation conditions, there was noticeable variability in the percentage of converged cases. The lowest convergence was seen in the five-variable condition with 30 time points and AR effects closed at the beginning of estimation, with approximately 48% of individual-level models terminating normally. The models that did not terminate normally either (a) halted at a model one path prior to nonconvergence, or (b) failed to converge at all. However, if AR effects were open at the beginning of estimation, more than 99% of cases converged normally, meaning that the final model experienced no convergence issues and at least two of four fit indices met the criteria to be considered "excellent."

Meta-Model Results

As seen previously (Gates & Molenaar, 2012), the performance of S-GIMME did not fluctuate across varying levels of sample size in the present study. For this reason, Figure 1 collapses across sample size to display the path recovery of S-GIMME for the conditions when the AR effects are fixed (closed) and freed (open) across varying lengths of time and number of variables. As can clearly be seen, S-GIMME conducted with AR effects freely estimated greatly improves the accurate and reliable recovery of paths. Further, a larger number of time points greatly improves the recovery of data-generating relationships. Finally, though we find promising results in both the five- and 10-variable conditions, performance is better in the five-variable condition. Specifics
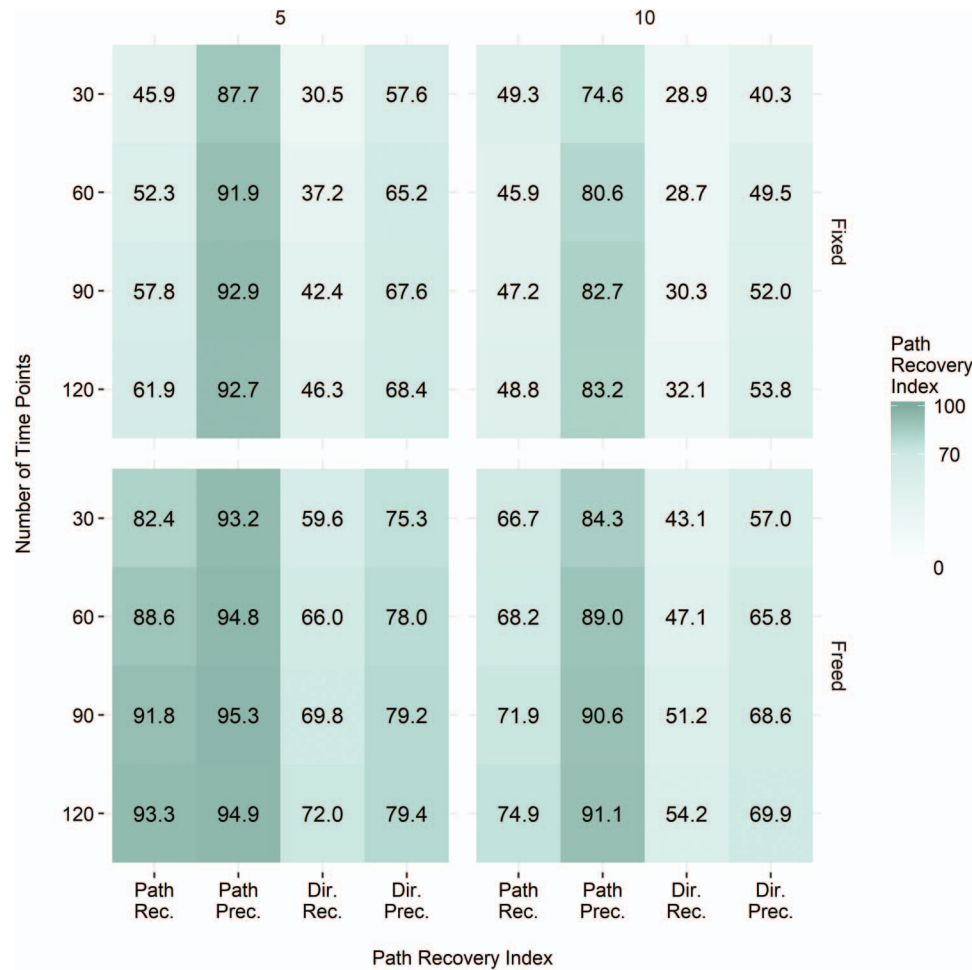
Figure 1. All outcomes by number of variables, time, and autoregressive estimation. See the online article for the color version of this figure.

regarding precision and direction as well as performance across the specific conditions are discussed below.

Path and direction recall. We statistically evaluated the recovery of data-generating paths across four simulation factors corresponding to conditions frequently encountered in daily diary data: the estimation of AR effects, sample size, number of time points, and number of variables. The outcome measures relating to path recovery were evaluated using an ANOVA that included main effects and interactive effects among the simulation factors. For the recall of true paths, main effects were present for opening the AR effects in estimation ($\eta_p^2 = .34$), the number of time points ($\eta_p^2 = .19$), and the number of variables ($\eta_p^2 = .02$). Thus, the recovery of true paths was improved in the presence of freely estimated AR effects, a larger number of time points, and a smaller number of variables.

A three-way interaction between AR, number of time points, and number of variables clarified these main effects; this interaction explained 6% of the variance ($\eta_p^2 = .06$) in true path recovery. The same interaction effect was present for the recall of true directions ($\eta_p^2 = .04$). Increased rates of path and direction recovery was observed in cells in which the AR paths were open, the

number of time points was greater (see Figure 2 for path recall), and the number of variables was lower. The total number of individuals, or sample size, did not meaningfully impact the path recovery. There was a large amount of variability present depending on simulation condition. For example, collapsing over sample size, the average path recovery in the condition with T = 120, AR effects open, and five variables was 93.3% compared with 45.9% in the condition with T = 30, AR effects closed, and 10 variables.

Path and direction precision. Whereas our two measures of recall provide information regarding our ability to recover paths that existed in the true model, our two measures of precision provide information regarding the quality of the paths we recovered in terms of false positives. That is, precision addresses the number of true paths in the recovered model relative to the number of total paths in the recovered model. Here, again, the number of individuals was not influential. The interaction of AR estimation with the number of variables was the most important determination of precision, in which precision was highest when AR effects were freed for estimation and the number of variables was smaller. There was also an interaction between the number of time points and the number of variables ($\eta_p^2 = .02$), such that the greatest
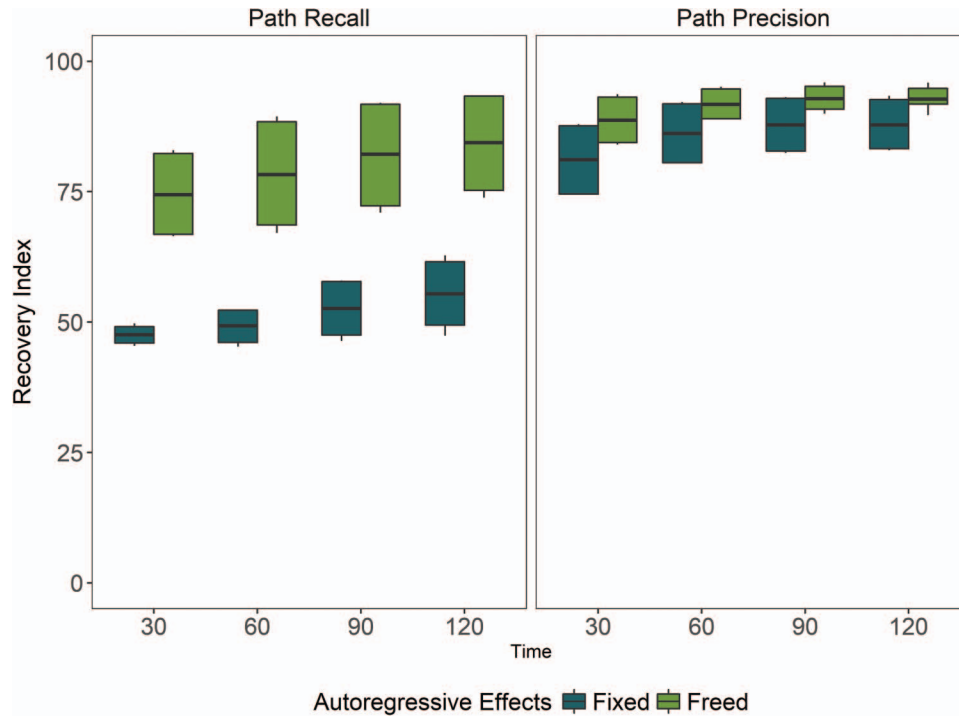
Figure 2.   Path recall and precision by time and autoregressive estimation. See the online article for the color version of this figure.

precision of paths is obtained when the number of variables is low and the number of time points is high. If the number of variables is large relative to the number of time points (e.g., in the condition with 10 variables and 30 time points), then the precision of paths is low. Figure 1 summarizes the findings across all four measures of path recovery. Table 1 contains comprehensive results for path recall by simulation condition.

Performance by level.   All results discussed thus far focus on recall and precision of all paths regardless of the data-generating level (e.g., group, subgroup, individual level). Figure 3 displays the path recovery of group-, subgroup-, and individual-level paths across the number of time points with estimated AR paths. We observe that the recovery of group-level relationships is most robust within S-GIMME. The recovery of subgroup-level relationships is next, followed by the recovery of individual-level relationships. Importantly, the recovery of subgroup-level paths will be in part determined by the quality of the data-driven subgrouping. If a path is not added in the subgroup-level search, it may be recovered at the individual level. The high recovery of subgroup-level paths points to the benefits of searching for consistencies across individuals in order to arrive at reliable models. The recovery of individual-level paths was lowest, averaging around 60% recall, suggesting that S-GIMME may achieve a model characterized by "excellent" fit before all individual-level paths have been added. The trade-off of the use of a fit index as a stopping criterion is that although not all paths may be recovered, favoring parsimony prevents false positives.

Relative bias.   Finally, we assessed the relative bias of the recovered path estimates by condition. There existed variability across conditions, though no individual effect exceeded a η_p^2 value

of .01. A graphic depiction (see Figure 4) shows the amount of bias present across the number of time points, split by AR condition. For the conditions in which the AR effects are freed at the start of estimation, the relative bias stays within ±10% across the number of time points, in which bias is lowest when T = 120. These findings are consistent with prior work evaluating bias when using the block Toeplitz method to implement SEM-based time series models.

Subgroup recovery.   Finally, we examined the ARI_{HA} as a measure of subgroup recovery. Subgroup results generally follow those seen for recall and precision because the algorithm uses path information to cluster individuals. Therefore, the most accurate recovery of subgroups was present in conditions with AR effects freed and 120 time points. For the 10-variable condition, the ARI_{HA} exceeded .8, provided that T ≥90 and AR effects were freely estimated. The subgroup recovery of the five-variable condition was poorer, though it, too, increased with an increasing number of time points (range = .09–.47 between T = 30 and T = 120, respectively). A graphical examination of the results revealed that freeing the estimation of the AR effects was the most important determination of subgroup recovery; however, no simulation factor exceeded the cutoff for η_p^2. Table 2 shows comprehensive results for subgroup recovery by simulation condition.

Empirical Study

S-GIMME discovered three subgroups from the 35 individuals. These subgroups are labeled "A," "B," and "C," in order of size, for descriptive purposes. Subgroup A comprised the largest proportion of the sample, with *n* = 24; subgroup B was the second

Table 1
Path Recall Across All Simulation Conditions

AR effects	Number of time points	Sample size	Path recall: 5 variables	Path recall: 10 variables
Fixed	30	25	46.30	48.86
		75	45.90	49.19
		150	45.45	49.75
	60	25	52.28	45.27
		75	52.21	46.04
		150	52.38	46.40
	90	25	57.64	46.37
		75	57.97	47.49
		150	57.81	47.60
	120	25	62.78	47.40
		75	61.72	49.77
		150	61.14	49.33
Freed	30	25	82.93	67.06
		75	82.47	66.73
		150	81.82	66.46
	60	25	89.43	68.87
		75	88.62	68.53
		150	87.73	67.11
	90	25	92.05	72.70
		75	91.75	72.12
		150	91.72	70.98
	120	25	93.43	75.77
		75	93.38	75.07
		150	93.08	73.86

Note. AR = autoregressive relationships.

largest ($n = 10$). Subgroup C contained only one individual; this singleton group demonstrates the strength of the Walktrap approach for clustering individuals. Rather than combining individuals who do not share temporal processes into a "super" subgroup,

it allows the existence of subgroups with very low numbers. This individual can thus be considered an outlier with respect to the multivariate pattern of temporal relations. Thus, it is not included in subsequent analysis.

Figure 5 depicts the patterns of effects found for the group, Subgroups A and B, and individuals. Only one group-level path was found—after controlling for other individual-level lagged and contemporaneous effects, feelings of depression within the last 24 hr can be statistically predicted by reports of mood lability at that same time period. The small number of group-level paths suggests a very high degree of heterogeneity in this sample, supporting the decision to conduct subgroup- and individual-level searches for temporal patterns rather than treating individuals as the same (as is frequently done within diagnostic categories). In contrast to the relatively sparse group-level structure, each subgroup was typified by five to six subgroup-level paths. The network of paths that defined Subgroup A included contemporaneous prediction of Anxiety by Depression and Mood Lability, Depression also predicted Emptiness, Anxiety predicted Anger, Anger predicted Urgency, and Urgency predicted Impulsivity. Interestingly, Subgroup B included a path that linked two of the same daily BPD manifestations as in Subgroup A but in the opposite direction: Anxiety predicted Mood Lability, Anxiety also predicted Impulsivity, Mood Lability predicted Urgency and Anger, and Urgency predicted Emptiness. Each of these groups is suggestive of potentially important clinical differences that are given greater treatment in the Discussion section.

Although the sample is small and any group comparison is likely to be low powered to detect anything but sizable effects, we conducted a one-way MANOVA with the seven indicators of daily BPD symptomology as independent variables and subgroup as-
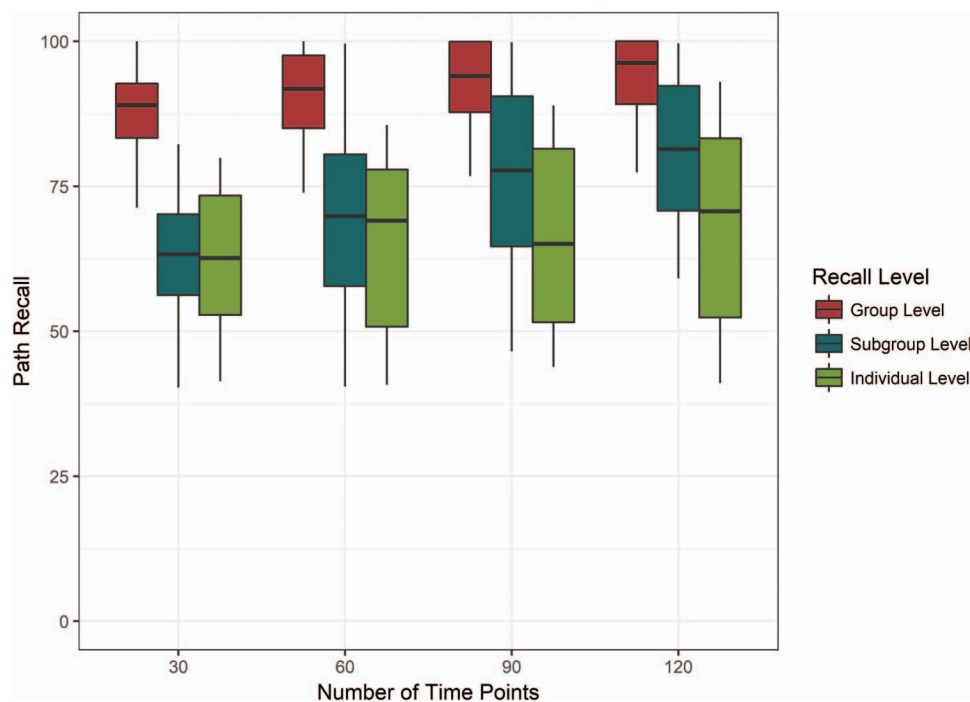

Figure 3. Recall by level. See the online article for the color version of this figure.
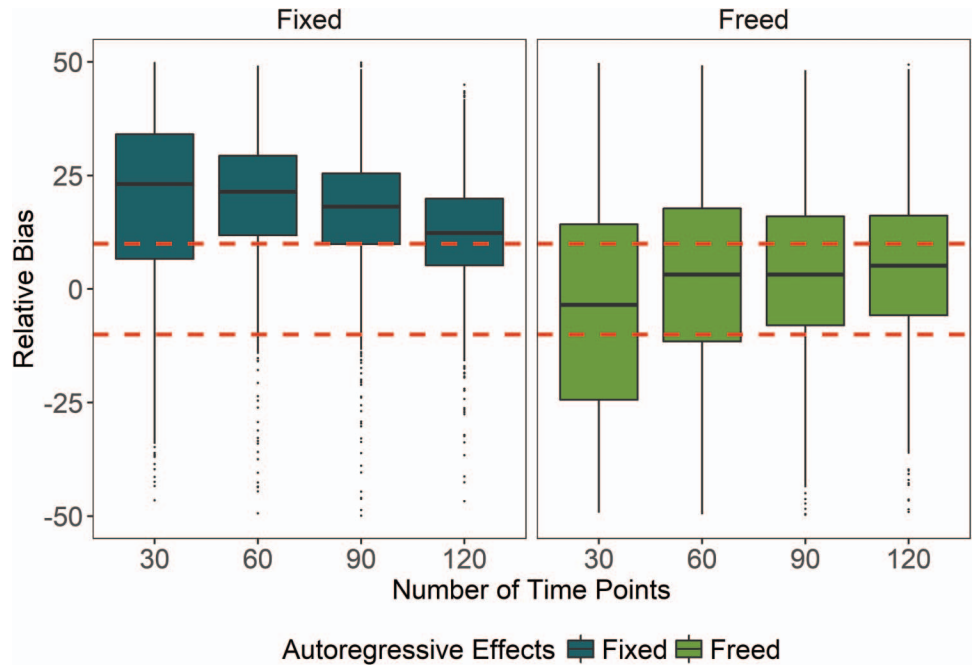
Figure 4. Relative bias by time and autoregressive estimation. See the online article for the color version of this figure.

signment as dependent variables, similar to a discriminant analysis. Results yielded a nonsignificant main effect for BPD symptomology, Pillai's Trace = 0.35, $F(7, 26) = 1.98$, $p > .05$, suggesting no differences in the average daily symptom endorse-

ment based on subgroup allocation. Thus, these subgroups are not typified by symptom level but rather by similarities in their symptom dynamics. This demonstration illustrates that the relations of symptoms across time reveals information above and beyond simply averaging the data.

Discussion

The analytic method evaluated here enables researchers to reliably parse individuals into clusters based on similarities in their emotional, behavioral, and cognitive processes. From the S-GIMME results, researchers can make inferences regarding dynamic processes that are generalizable to the population from which individuals are drawn. Additionally, the results reveal patterns of effects that exist for a subset of individuals and, finally, individual-level effects. Originally developed for use on brain imaging data, we found that with minor adaptation (specifically, forcing AR effects to be freed at the start of model building), S-GIMME can be used for data with the same qualities as those seen in daily diary studies. Specifically, these results provide evidence that including the AR effects greatly improves recovery of the contemporaneous and lagged temporal effects among variables when conducting S-GIMME for model selection. This information may be particularly useful for ambulatory assessment, such as daily self-report, that may be collected at a low temporal resolution relative to the process of interest. In this case, as seen here in the empirical example, the majority of effects may surface contemporaneously, particularly if the relations among variables occur faster than the data collection (as described in Granger, 1988). Further, we find our results to be robust to the formation of the data-generating network structure, as a strength of the current study is our use of randomly generated network structures. The
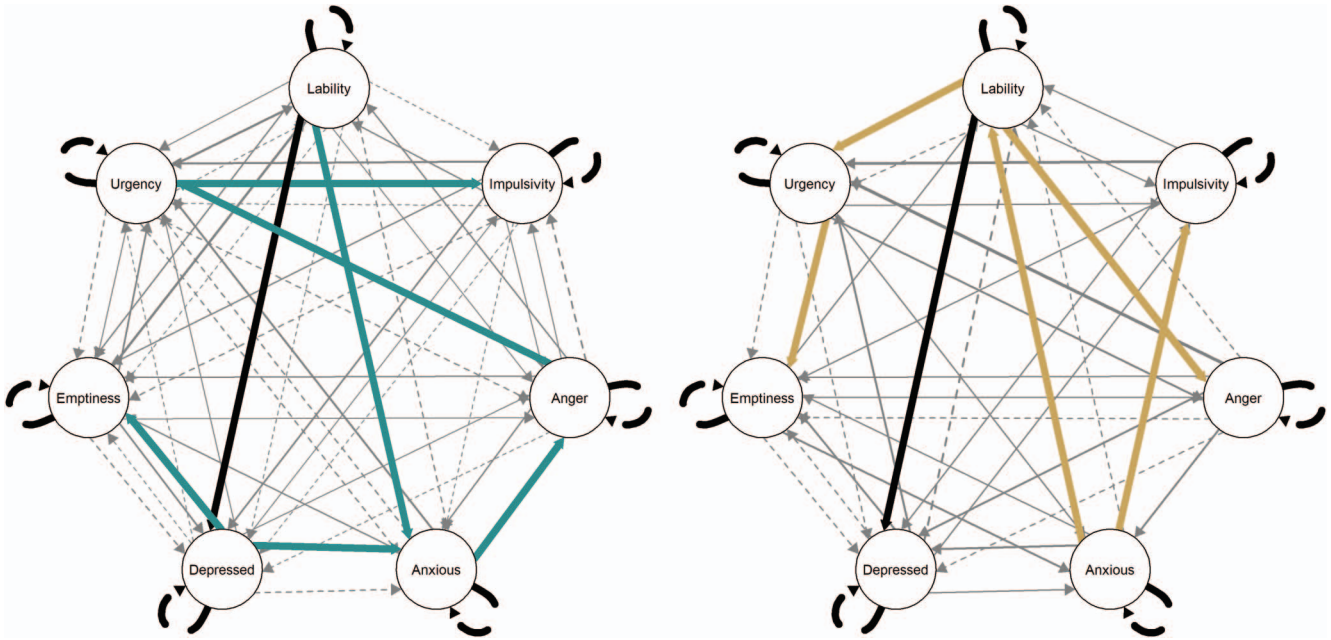
Table 2

Subgroup Recovery Across All Simulation Conditions

AR effects	Number of time points	Sample size	ARI_{HA}: 5 variables	ARI_{HA}: 10 variables
Fixed	30	25	.00	.00
		75	.00	.00
		150	.00	.00
	60	25	.01	.00
		75	.01	.00
		150	.00	.00
	90	25	.08	.01
		75	.05	.01
		150	.04	.01
	120	25	.15	.04
		75	.11	.04
		150	.09	.03
Freed	30	25	.12	.37
		75	.10	.42
		150	.06	.37
	60	25	.35	.71
		75	.31	.68
		150	.30	.66
	90	25	.39	.81
		75	.45	.85
		150	.44	.82
	120	25	.42	.85
		75	.50	.91
		150	.48	.89

Note. ARI_{HA} = Hubert-Arabie Adjusted Rand Index.

Figure 5. Depiction of relationships for Subgroups A and B. See the online article for the color version of this figure.

results from the simulation study suggest that S-GIMME is a viable option, assuming certain data features, for studying heterogeneity in multivariate processes assessed using daily diary studies—namely, data characterized by moderate to strong AR effects, a longer time series (e.g., T > 60), a relatively small number of variables (e.g., five to 10), and equally spaced time intervals are best suited for use with S-GIMME.

The present project found lower rates of recovery than what has been seen previously in data generated to emulate brain imaging data. A few factors decrease the utility of S-GIMME in the context of daily diary studies. First, many effects may be missed for studies that only have 30 data points (e.g., daily measures for 1 month). Here, we saw that having a higher number of variables ($n = 10$) when there were under 60 time points decreased accurate model recovery. However, there were also relatively few false positives, indicating that the effects that were recovered, although potentially more sparse, were accurate and reliable. Second, the AR effect sizes were much smaller than had been previously tested. The present article provides evidence that S-GIMME has satisfactory rates of recovery of the data-generating relations even when AR effects are small. Still, the rate of recovery was worse than seen in data generated to have higher AR effect sizes. S-GIMME favors parsimony by using fit indices as a stopping criteria rather than adding all paths that might be significant. Although this keeps false positives at a minimum, we see that it might have an unintended effect of stopping too early in the search and not adding paths that were in the data-generating model. More work needs to be done to identify the best fit indices for this context.

The model search procedure that began with AR effects estimated greatly improved recovery of the presence and direction of effects from within an S-GIMME framework. This finding aligns with statistical theory and prior research that demonstrate the utility of examining how variables relate after controlling for the influence each variable has on itself at a later time (Gates et al., 2010; Granger, 1969). However, other noteworthy options exist for accurately detecting the direction of prediction between variables. One option, GIMME for Multiple Solutions (GIMME-MS; Beltz & Molenaar, 2016) generates a set of possible connectivity maps from which the optimal solution can be selected using decision criteria (e.g., based on information criteria or residuals). Other Bayesian network approaches perform well in terms of accurate model recovery in the absence of AR effects. One such approach, independent multisample greedy equivalence search (iMaGES; Ramsey et al., 2011) uses information across all individuals in a manner that does not produce false paths. A second approach in the Bayesian network class of algorithms originally designed for cross-sectional studies, linear non-Gaussian acyclical model (LiNGAM; Hyvärinen, Zhang, Shimizu, & Hoyer, 2010; Shimizu, Hoyer, Hyvärinen, & Kerminen, 2006), performs well for non-Gaussian data but likely requires a greater number of observations than typical in daily diary studies. These latter two approaches, like the original GIMME, performed well on a canonical set of benchmark fMRI data (Gates & Molenaar, 2012; Mumford & Ramsey, 2014; Smith et al., 2011). Much like including ARs for all individuals with S-GIMME, GIMME-MS, iMaGES, and LiNGAM appear to be useful when contemporaneous effects are large compared with lagged effects and when AR effects are low or inconsistent (i.e., present for some variables but not for others).

Other areas of growth for model selection pertain specifically to qualities of daily diary data. As noted above, multiple items were removed from the empirical demonstration due to there being no variability for some individuals. Analysis from within a general linear modeling framework (including time series) cannot imme-

diately utilize variables that are constant across the measurement period. Additionally, it is highly likely that these processes may change across time, reflecting nonstationarity (and a violation of the assumptions of S-GIMME). This is particularly true if the participants are measured during an intervention. In these cases, trends may be addressed through preprocessing or other approaches (e.g., splitting of time series) prior to S-GIMME analysis. Another option is to increase the lag order to account for periodic effects (e.g., a lag of 7 for weekly effects), but this may only be tenable when a small number of variables is used. Adding an exogenous basis vector that contains the shape of the trend is another option that will be shortly available. Non-normality in the distribution of the data presents another issue. At present, it is unknown the influence that this will have on S-GIMME's recovery rate. A more pressing issue is that of measurement error. Within the SEM framework, it is rather straightforward to arrive at latent variables in an effort to obtain more precise measurements of underlying constructs. Extension of GIMME to manage latent variables is currently underway by integrating the gimme package with MIIVsem (Z. Fisher, Bollen, Gates, & Rönkkö, 2017).

Further, given the data-driven nature of the GIMME approach, the role of measurement error should be investigated in future work, particularly when considering how different modalities of data collection may affect the precision of our instrument and, therefore, our error in measurement. Finally, some designs will encounter unequal spacing between observations. This unequal spacing may occur in ecological momentary assessments for which the measurement is taken only when the participant indicates a life event occurred (as described in the review by Laurenceau & Bolger, 2005) or when there are no measurements obtained while the participants are at sleep (e.g., the publicly available data described in A. J. Fisher, 2015) or in burst designs. These sorts of irregular time series pose methodological and conceptual problems for contemporary time-series methods, not just GIMME. A number of approaches have been proposed and implemented to generate series with equal spacing from these types of data, such as continuous time modeling or rescaling continuous time to integer time (see, e.g., Asparouhov, Hamaker, & Muthén, 2018). Users interested in implementing GIMME with irregularly spaced data may wish to create equally spaced time series in advance by using interpolation and resampling methods (Beltz & Gates, 2017; Fisher & Boswell, 2016). However, the effect of these approaches on parameter estimates for ambulatory assessment data and recovery of the correct model remains unstudied in GIMME and poorly understood across all analytic approaches. These issues warrant much more attention for time series analysis for intensively gathered longitudinal data collected for psychological studies.

To illustrate the utility of using S-GIMME method, we examined group-, subgroup-, and individual-level temporal models in a sample of 35 patients diagnosed with BPD who completed daily diaries of various clinical features over 60 days or more. We found that only a single group-level path existed tying together all of these participants with the same diagnosis. This path suggests that the common feature in this group is the within-day prediction of depression from mood lability. At the same time, we found rich individual heterogeneity in contemporaneous and lagged paths. Moreover, we found that two subgroups emerged that were markedly distinct in the daily associations among several variables. We also found that the subgroups could not be differentiated using the average daily endorsement of the symptom variables examined in their temporal processes, though we note that this was a low-powered comparison for all but large effects. Nevertheless, this demonstrates that the information regarding how the symptoms relate across time differed from what can be reaped by looking at average scores alone. Although necessarily speculative given the limitations of the data (e.g., modest sample size), interpreting the subgroup specific models offers compelling results and beckon further analyses in larger samples. The processes that differentiate the Subgroup A from Subgroup B may best be understood as internalizing versus externalizing processes. That is, the majority of relationships within Subgroup A were between the predominantly affective variables (e.g., Depression, Anxiety, Mood Lability, Anger, and Emptiness); the disinhibitory behaviors (Urgency, Impulsivity) are linked but are largely separate from these processes.

Conclusion

The present article introduces one approach for analyzing daily diary data for heterogeneous populations that satisfies a few issues seen in the analysis of temporal processes across individuals—namely, S-GIMME arrives at individual-level models that are accurate and reliable even when parameters mimic those seen in daily diary studies. Additionally, S-GIMME provides robust results for as few as 60 time points, which is far fewer than the number that has been tested in neuroimaging research. Finally, S-GIMME performs well even with a relatively low number of participants. Ultimately, our demonstration underscores the need for methods such as S-GIMME moving forward given the increasing use of intensive longitudinal data in psychological research and the potential for these data to provide novel insights into human behavior and mental health.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Asparouhov, T., Hamaker, E. L., & Muthén, B. (2018). Dynamic structural equation models. *Structural Equation Modeling: An Multidisciplinary Journal, 3,* 359–388. http://dx.doi.org/10.1080/10705511.2017.1406803

Bachrach, R. L., & Read, J. P. (2017). Peer alcohol behavior moderates within-level associations between posttraumatic stress disorder symptoms and alcohol use in college students. *Psychology of Addictive Behaviors, 31,* 576–588. http://dx.doi.org/10.1037/adb0000285

Beltz, A. M., & Gates, K. M. (2017). Network Mapping with GIMME. *Multivariate Behavioral Research, 52,* 789–804. http://dx.doi.org/10.1080/00273171.2017.1373014

Beltz, A. M., Gates, K. M., Engels, A. S., Molenaar, P. C., Pulido, C., Turrisi, R., . . . Wilson, S. J. (2013). Changes in alcohol-related brain networks across the first year of college: A prospective pilot study using fMRI effective connectivity mapping. *Addictive Behaviors, 38,* 2052–2059. http://dx.doi.org/10.1016/j.addbeh.2012.12.023

Beltz, A. M., & Molenaar, P. C. (2016). Dealing with multiple solutions in structural vector autoregressive models. *Multivariate behavioral research, 51*(2–3), 357–373.

Bollen, K. A. (1989). *Structural equations with latent variables.* New York, NY: Wiley. http://dx.doi.org/10.1002/9781118619179

Borkenau, P., & Ostendorf, F. (1998). The Big Five as states: How useful is the five-factor model to describe intraindividual variations over time?

Journal of Research in Personality, 32, 202–221. http://dx.doi.org/10.1006/jrpe.1997.2206

Bradley, R., Zittel Conklin, C., & Westen, D. (2005). The borderline personality diagnosis in adolescents: Gender differences and subtypes. *Journal of Child Psychology and Psychiatry, 46,* 1006–1019. http://dx.doi.org/10.1111/j.1469-7610.2004.00401.x

Brown, T. A. (2006). *Confirmatory Factor Analysis for Applied Research.* New York, NY: Guilford.

Castro-Schilo, L., & Ferrer, E. (2013). Comparison of nomothetic versus idiographic-oriented methods for making predictions about distal outcomes from time series data. *Multivariate Behavioral Research, 48,* 175–207. http://dx.doi.org/10.1080/00273171.2012.736042

Chatfield, C. (2003). *The analysis of time series: An introduction.* Boca Raton, FL: Chapman and Hall.

Critchfield, K. L., Clarkin, J. F., Levy, K. N., & Kernberg, O. F. (2008). Organization of co-occurring Axis II features in borderline personality disorder. *British Journal of Clinical Psychology, 47,* 185–200. http://dx.doi.org/10.1348/014466507X240731

Ebner-Priemer, U. W., Kuo, J., Kleindienst, N., Welch, S. S., Reisch, T., Reinhard, I., . . . Bohus, M. (2007). State affective instability in borderline personality disorder assessed by ambulatory monitoring. *Psychological Medicine, 37,* 961–970. http://dx.doi.org/10.1017/S0033291706009706

Enders, C. K., & Bandalos, D. L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural Equation Modeling, 8,* 430–457. http://dx.doi.org/10.1207/S15328007SEM0803_5

Fernandez, K. C., Fisher, A. J., & Chi, C. (2017). Development and initial implementation of the Dynamic Assessment Treatment Algorithm (DATA). *PLoS ONE, 12,* e0178806. http://dx.doi.org/10.1371/journal.pone.0178806

Fisher, A. J. (2015). Toward a dynamic model of psychological assessment: Implications for personalized care. *Journal of Consulting and Clinical Psychology, 83,* 825–836. http://dx.doi.org/10.1037/ccp0000026

Fisher, A. J., & Boswell, J. F. (2016). Enhancing the personalization of psychotherapy with dynamic assessment and modeling. *Assessment, 23,* 496–506. http://dx.doi.org/10.1177/1073191116638735

Fisher, Z., Bollen, K., Gates, K., & Rönkkö, M. (2017). *MIIVsem: model implied instrumental variable (MIIV) estimation of structural equation models* (R package Version 0.5.2). Retrieved from https://CRAN.R-project.org/package=MIIVsem

Gadassi, R., Bar-Nahum, L. E., Newhouse, S., Anderson, R., Heiman, J. R., Rafaeli, E., & Janssen, E. (2016). Perceived partner responsiveness mediates the association between sexual and marital satisfaction: A daily diary study in newlywed couples. *Archives of Sexual Behavior, 45,* 109–120. http://dx.doi.org/10.1007/s10508-014-0448-2

Gates, K. M., Henry, T., Steinley, D., & Fair, D. A. (2016). A Monte Carlo evaluation of weighted community detection algorithms. *Frontiers in Neuroinformatics, 10,* 45.

Gates, K. M., Lane, S. T., Varangis, E., Giovanello, K., & Guiskewicz, K. (2017). Unsupervised classification during time-series model building. *Multivariate Behavioral Research, 52,* 129–148. http://dx.doi.org/10.1080/00273171.2016.1256187

Gates, K. M., & Molenaar, P. C. (2012). Group search algorithm recovers effective connectivity maps for individuals in homogeneous and heterogeneous samples. *NeuroImage, 63,* 310–319. http://dx.doi.org/10.1016/j.neuroimage.2012.06.026

Gates, K. M., Molenaar, P. C., Hillary, F. G., Ram, N., & Rovine, M. J. (2010). Automatic search for fMRI connectivity mapping: An alternative to Granger causality testing using formal equivalences among SEM path modeling, VAR, and unified SEM. *NeuroImage, 50,* 1118–1125. http://dx.doi.org/10.1016/j.neuroimage.2009.12.117

Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica, 37,* 424–438. http://dx.doi.org/10.2307/1912791

Granger, C. W. (1988). Some recent development in a concept of causality. *Journal of Econometrics, 39,* 199–211. http://dx.doi.org/10.1016/0304-4076(88)90045-0

Hallquist, M. N., & Pilkonis, P. A. (2012). Refining the phenotype of borderline personality disorder: Diagnostic criteria and beyond. *Personality Disorders: Theory, Research, and Treatment, 3,* 228–246. http://dx.doi.org/10.1037/a0027953

Hamaker, E. L., Dolan, C. V., & Molenaar, P. C. (2002). On the nature of SEM estimates of ARMA parameters. *Structural Equation Modeling, 9,* 347–368. http://dx.doi.org/10.1207/S15328007SEM0903_3

Hamaker, E. L., Grasman, R. P., & Kamphuis, J. H. (2016). Modeling BAS dysregulation in bipolar disorder: Illustrating the potential of time series analysis. *Assessment, 23,* 436–446. http://dx.doi.org/10.1177/1073191116632339

Hamaker, E. L., & Wichers, M. (2017). No time like the present: Discovering the hidden dynamics in intensive longitudinal data. *Current Directions in Psychological Science, 26,* 10–15. http://dx.doi.org/10.1177/0963721416666518

Hubert, L., & Arabie, P. (1985). Comparing partitions. *Journal of Classification, 2,* 193–218. http://dx.doi.org/10.1007/BF01908075

Hyman, S. E. (2010). The diagnosis of mental disorders: The problem of reification. *Annual review of clinical psychology, 6,* 155–179.

Hyvärinen, A., Zhang, K., Shimizu, S., & Hoyer, P. O. (2010). Estimation of a structural vector autoregression model using non-Gaussianity. *Journal of Machine Learning Research, 11,* 1709–1731.

Jöreskog, K. G., & Sörbom, D. (1986). *LISREL VI: Analysis of linear structural relationships by maximum likelihood, instrumental variables, and least squares methods.* Mooresville, IN: Scientific Software.

Karunanayaka, P., Eslinger, P. J., Wang, J. L., Weitekamp, C. W., Molitoris, S., Gates, K. M., . . . Yang, Q. X. (2014). Networks involved in olfaction and their dynamics using independent component analysis and unified structural equation modeling. *Human Brain Mapping, 35,* 2055–2072. http://dx.doi.org/10.1002/hbm.22312

Kernberg, O. F. (1975). A systems approach to priority setting of interventions in groups. *International Journal of Group Psychotherapy, 25,* 251–275. http://dx.doi.org/10.1080/00207284.1975.11491899

Kim, J., Zhu, W., Chang, L., Bentler, P. M., & Ernst, T. (2007). Unified structural equation modeling approach for the analysis of multisubject, multivariate functional MRI data. *Human Brain Mapping, 28,* 85–93. http://dx.doi.org/10.1002/hbm.20259

Kotov, R., Krueger, R. F., Watson, D., Achenbach, T. M., Althoff, R. R., Bagby, R. M., . . . Zimmerman, M. (2017). The Hierarchical Taxonomy of Psychopathology (HiTOP): A dimensional alternative to traditional nosologies. *Journal of Abnormal Psychology, 126,* 454–477. http://dx.doi.org/10.1037/abn0000258

Lane, S. T., & Gates, K. M. (2017). Automated selection of robust individual-level structural equation models for time series data. *Structural Equation Modeling: A Multidisciplinary Journal, 24,* 768–782.

Lane, S. T., Gates, K. M., & Molenaar, P. (2014). *gimme: Group Iterative Multiple Model Estimation* (R package version 0.1–7). https://CRAN.R-project.org/package=gimme

Laurenceau, J. P., & Bolger, N. (2005). Using diary methods to study marital and family processes. *Journal of Family Psychology, 19,* 86–97. http://dx.doi.org/10.1037/0893-3200.19.1.86

Lebo, M. A., & Nesselroade, J. R. (1978). Intraindividual differences dimensions of mood change during pregnancy identified in five P-technique factor analyses. *Journal of Research in Personality, 12,* 205–224. http://dx.doi.org/10.1016/0092-6566(78)90098-3

Lee, C. M., Cronce, J. M., Baldwin, S. A., Fairlie, A. M., Atkins, D. C., Patrick, M. E., . . . Leigh, B. C. (2017). Psychometric analysis and validity of the daily alcohol-related consequences and evaluations mea-

sure for young adults. *Psychological Assessment, 29,* 253–263. http://dx .doi.org/10.1037/pas0000320

Lenzenweger, M. F., Clarkin, J. F., Yeomans, F. E., Kernberg, O. F., & Levy, K. N. (2008). Refining the borderline personality disorder phenotype through finite mixture modeling: Implications for classification. *Journal of Personality Disorders, 22,* 313–331. http://dx.doi.org/10 .1521/pedi.2008.22.4.313

Liao, T. W. (2005). Clustering of time series data—A survey. *Pattern Recognition, 38,* 1857–1874. http://dx.doi.org/10.1016/j.patcog.2005.01 .025

Lilienfeld, S. O., Waldman, I. D., & Israel, A. C. (1994). A critical examination of the use of the term and concept of comorbidity in psychopathology research. *Clinical Psychology: Science and Practice, 1,* 71–83. http://dx.doi.org/10.1111/j.1468-2850.1994.tb00007.x

MacCallum, R. (1986). Specification searches in covariance structure modeling. *Psychological Bulletin, 100,* 107–120. http://dx.doi.org/10 .1037/0033-2909.100.1.107

MacCallum, R. C., Roznowski, M., & Necowitz, L. B. (1992). Model modifications in covariance structure analysis: The problem of capitalization on chance. *Psychological Bulletin, 111,* 490–504. http://dx.doi .org/10.1037/0033-2909.111.3.490

Miskewicz, K., Fleeson, W., Arnold, E. M., Law, M. K., Mneimne, M., & Furr, R. M. (2015). A contingency-oriented approach to understanding borderline personality disorder: Situational triggers and symptoms. *Journal of Personality Disorders, 29,* 486–502. http://dx.doi.org/10 .1521/pedi.2015.29.4.486

Molenaar, P. C. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement: Interdisciplinary Research and Perspectives, 2,* 201–218.

Molenaar, P. C. (2008). On the implications of the classical ergodic theorems: Analysis of developmental processes has to focus on intraindividual variation. *Developmental Psychobiology, 50,* 60–69. http:// dx.doi.org/10.1002/dev.20262

Molenaar, P. C. M., Beltz, A. M., Gates, K. M., & Wilson, S. J. (2016). State space modeling of time-varying contemporaneous and lagged relations in connectivity maps. *NeuroImage, 125,* 791–802. http://dx.doi .org/10.1016/j.neuroimage.2015.10.088

Molenaar, P. C. M., & Campbell, C. G. (2009). The new person-specific paradigm in psychology. *Current Directions in Psychological Science, 18,* 112–117. http://dx.doi.org/10.1111/j.1467-8721.2009.01619.x

Muehlenkamp, J. J., Engel, S. G., Wadeson, A., Crosby, R. D., Wonderlich, S. A., Simonich, H., & Mitchell, J. E. (2009). Emotional states preceding and following acts of non-suicidal self-injury in bulimia nervosa patients. *Behaviour Research and Therapy, 47,* 83–87. http:// dx.doi.org/10.1016/j.brat.2008.10.011

Mumford, J. A., & Ramsey, J. D. (2014). Bayesian networks for fMRI: A primer. *NeuroImage, 86,* 573–582. http://dx.doi.org/10.1016/j .neuroimage.2013.10.020

Myin-Germeys, I., Oorschot, M., Collip, D., Lataster, J., Delespaul, P., & van Os, J. (2009). Experience sampling research in psychopathology: Opening the black box of daily life. *Psychological Medicine, 39,* 1533–1547. http://dx.doi.org/10.1017/S0033291708004947

Myin-Germeys, I., van Os, J., Schwartz, J. E., Stone, A. A., & Delespaul, P. A. (2001). Emotional reactivity to daily life stress in psychosis. *Archives of General Psychiatry, 58,* 1137–1144. http://dx.doi.org/10 .1001/archpsyc.58.12.1137

Newman, M. E. (2004). Fast algorithm for detecting community structure in networks. *Physical Review E, 69,* 066133. http://dx.doi.org/10.1103/ PhysRevE.69.066133

Nichols, T. T., Gates, K. M., Molenaar, P. C., & Wilson, S. J. (2014). Greater BOLD activity but more efficient connectivity is associated with better cognitive performance within a sample of nicotine-deprived smokers. *Addiction Biology, 19,* 931–940. http://dx.doi.org/10.1111/adb .12060

Pe, M. L., Kircanski, K., Thompson, R. J., Bringmann, L. F., Tuerlinckx, F., Mestdagh, M., . . . Gotlib, I. H. (2015). Emotion-network density in major depressive disorder. *Clinical Psychological Science, 3,* 292–300. http://dx.doi.org/10.1177/2167702614540645

Pons, P., & Latapy, M. (2006). Computing communities in large networks using random walks. *Journal of Graph Algorithms and Applications, 10,* 191–218. http://dx.doi.org/10.7155/jgaa.00124

Price, R. B., Lane, S., Gates, K., Kraynak, T. E., Horner, M. S., Thase, M. E., & Siegle, G. J. (2017). Parsing heterogeneity in the brain connectivity of depressed and healthy adults during positive mood. *Biological Psychiatry, 81,* 347–357. http://dx.doi.org/10.1016/j .biopsych.2016.06.023

Ram, N., & Grimm, K. J. (2009). Growth mixture modeling: A method for identifying differences in longitudinal change among unobserved groups. *International Journal of Behavioral Development, 33,* 565–576. http://dx.doi.org/10.1177/0165025409343765

Ram, N., Rabbitt, P., Stollery, B., & Nesselroade, J. R. (2005). Cognitive performance inconsistency: Intraindividual change and variability. *Psychology and Aging, 20,* 623–633. http://dx.doi.org/10.1037/0882-7974 .20.4.623

Ramsey, J. D., Hanson, S. J., & Glymour, C. (2011). Multi-subject search correctly identifies causal connections and most causal directions in the DCM models of the Smith et al. simulation study. *NeuroImage, 58,* 838–848. http://dx.doi.org/10.1016/j.neuroimage.2011.06.068

Ramsey, J. D., Sanchez-Romero, R., & Glymour, C. (2014). Non-Gaussian methods and high-pass filters in the estimation of effective connections. *NeuroImage, 84,* 986–1006. http://dx.doi.org/10.1016/j.neuroimage .2013.09.062

Sadikaj, G., Moskowitz, D. S., Russell, J. J., Zuroff, D. C., & Paris, J. (2013). Quarrelsome behavior in borderline personality disorder: Influence of behavioral and affective reactivity to perceptions of others. *Journal of Abnormal Psychology, 122,* 195–207. http://dx.doi.org/10 .1037/a0030871

Searles, J. S., Perrine, M. W., Mundt, J. C., & Helzer, J. E. (1995). Self-report of drinking using touch-tone telephone: Extending the limits of reliable daily contact. *Journal of Studies on Alcohol, 56,* 375–382. http://dx.doi.org/10.15288/jsa.1995.56.375

Shiffman, S., Gwaltney, C. J., Balabanis, M. H., Liu, K. S., Paty, J. A., Kassel, J. D., . . . Gnys, M. (2002). Immediate antecedents of cigarette smoking: An analysis from ecological momentary assessment. *Journal of Abnormal Psychology, 111,* 531–545. http://dx.doi.org/10.1037/0021-843X.111.4.531

Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological momentary assessment. *Annual Review of Clinical Psychology, 4,* 1–32. http:// dx.doi.org/10.1146/annurev.clinpsy.3.022806.091415

Shimizu, S., Hoyer, P. O., Hyvärinen, A., & Kerminen, A. (2006). A linear non-Gaussian acyclic model for causal discovery. *Journal of Machine Learning Research, 7,* 2003–2030.

Silk, J. S., Steinberg, L., & Morris, A. S. (2003). Adolescents' emotion regulation in daily life: Links to depressive symptoms and problem behavior. *Child Development, 74,* 1869–1880. http://dx.doi.org/10 .1046/j.1467-8624.2003.00643.x

Smith, S. M., Miller, K. L., Salimi-Khorshidi, G., Webster, M., Beckmann, C. F., Nichols, T. E., . . . Woolrich, M. W. (2011). Network modelling methods for FMRI. *NeuroImage, 54,* 875–891. http://dx.doi.org/10 .1016/j.neuroimage.2010.08.063

Smyth, J. M., Wonderlich, S. A., Heron, K. E., Sliwinski, M. J., Crosby, R. D., Mitchell, J. E., & Engel, S. G. (2007). Daily and momentary mood and stress are associated with binge eating and vomiting in bulimia nervosa patients in the natural environment. *Journal of Consulting and Clinical Psychology, 75,* 629–638. http://dx.doi.org/10.1037/0022-006X.75.4.629

Steinley, D. (2004). Properties of the Hubert-Arabie adjusted Rand index. *Psychological Methods, 9,* 386–396. http://dx.doi.org/10.1037/1082-989X.9.3.386

Trull, T. J., & Ebner-Priemer, U. (2013). Ambulatory assessment. *Annual Review of Clinical Psychology, 9,* 151–176. http://dx.doi.org/10.1146/annurev-clinpsy-050212-185510

Trull, T. J., Solhan, M. B., Tragesser, S. L., Jahng, S., Wood, P. K., Piasecki, T. M., & Watson, D. (2008). Affective instability: Measuring a core feature of borderline personality disorder with ecological momentary assessment. *Journal of Abnormal Psychology, 117,* 647–661. http://dx.doi.org/10.1037/a0012532

Ward, J. H., Jr. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association, 58,* 236–244. http://dx.doi.org/10.1080/01621459.1963.10500845

Wright, A. G. C., Beltz, A. M., Gates, K. M., Molenaar, P. C. M., & Simms, L. J. (2015). Examining the dynamic structure of daily internalizing and externalizing behavior at multiple levels of analysis. *Frontiers in Psychology, 6,* 1914. http://dx.doi.org/10.3389/fpsyg.2015.01914

Wright, A. G., Hallquist, M. N., Morse, J. Q., Scott, L. N., Stepp, S. D., Nolf, K. A., & Pilkonis, P. A. (2013). Clarifying interpersonal heterogeneity in borderline personality disorder using latent mixture modeling. *Journal of Personality Disorders, 27,* 125–143. http://dx.doi.org/10.1521/pedi.2013.27.2.125

Wright, A. G. C., Hopwood, C. J., & Simms, L. J. (2015). Daily interpersonal and affective dynamics in personality disorder. *Journal of Personality Disorders, 29,* 503–525. http://dx.doi.org/10.1521/pedi.2015.29.4.503

Wright, A. G. C., & Simms, L. J. (2016). Stability and fluctuation of personality disorder features in daily life. *Journal of Abnormal Psychology, 125,* 641–656. http://dx.doi.org/10.1037/abn0000169

Yang, J., Gates, K. M., Molenaar, P., & Li, P. (2015). Neural changes underlying successful second language word learning: An fMRI study. *Journal of Neurolinguistics, 33,* 29–49. http://dx.doi.org/10.1016/j.jneuroling.2014.09.004

Zelle, S. L., Gates, K. M., Fiez, J. A., Sayette, M. A., & Wilson, S. J. (2017). The first day is always the hardest: Functional connectivity during cue exposure and the ability to resist smoking in the initial hours of a quit attempt. *NeuroImage, 151,* 24–32. http://dx.doi.org/10.1016/j.neuroimage.2016.03.015

Zevon, M. A., & Tellegen, A. (1982). The structure of mood change: An idiographic/nomothetic analysis. *Journal of Personality and Social Psychology, 43,* 111–122. http://dx.doi.org/10.1037/0022-3514.43.1.111

Zimmermann, J., Woods, W. C., Ritter, S., Happel, M., Masuhr, O., Jaeger, U., . . . Wright, A. G. C. (in press). Integrating structure and dynamics in personality assessment: First steps toward the development and validation of a Personality Dynamics Diary. *Psychological Assessment.*

E-Mail Notification of Your Latest Issue Online!

Would you like to know when the next issue of your favorite APA journal will be available online? This service is now available to you. Sign up at https://my.apa.org/portal/alerts/ and you will be notified by e-mail when issues of interest to you become available!