**Focusing Personality Assessment on the Person: Modeling General, Shared, and Person
Specific Processes in Personality and Psychopathology**

Aidan G.C. Wright¹

Kathleen M. Gates²

Cara Arizmendi²

Stephanie Lane²

William C. Woods¹

Elizabeth A. Edershile¹

¹University of Pittsburgh; ²University of North Carolina, Chapel Hill

We would like to thank Barbara De Clercq, Filip De Fruyt, Lize Verbeke, Bart Wille, Joeri Hofmans, Karla Van Leeuwen, & Leonard Simms for organizing and inviting the first author to the Expert Meeting on Personality Assessment in Oostduinkirke, Belgium, and the European Association for Personality Psychology for providing funding for the meeting.

This research was supported by grants from the National Institute of Mental Health (F32 MH097325, L30 MH101760, Wright; R01 MH080086, Simms). The opinions expressed are solely those of the authors and not those of the funding source.

Correspondence concerning this article should be addressed to Aidan Wright, Department of Psychology, University of Pittsburgh, 4119 Sennott Square, 210 S. Bouquet St., Pittsburgh, PA, 15260. E-mail: aidan@pitt.edu

Abstract

Personality and psychopathology are composed of dynamic and interactive processes among diverse psychological systems, manifesting over time and in response to an individual's natural environment. Ambulatory assessment techniques promise to revolutionize assessment practices by allowing access to the dynamic data necessary to study these processes directly. Assessing manifestations of personality and psychopathology naturalistically in an individual's own ecology allows for dynamic modeling of key behavioral processes. However, advances in dynamic data collection have highlighted the challenges of both fully understanding an individual (via idiographic models) and how s/he compares to others (as seen in nomothetic models). Methods are needed that can simultaneously model idiographic (i.e., person-specific) processes and nomothetic (i.e., general) structure from intensive longitudinal personality assessments. Here we present a method, Group Iterative Multiple Model Estimation (GIMME) for simultaneously studying general, shared (i.e., in subgroups), and person-specific processes in intensive longitudinal behavioral data. We first provide introduction to the GIMME method, followed by a demonstration of its use in a sample of individuals diagnosed with personality disorder who completed daily diaries over 100 consecutive days.

Keywords: Ambulatory Assessment; Personality Dynamics; Group Iterative Multiple Model Estimation; Idiographic; Nomothetic.

Public Significance Statement: Ambulatory assessment techniques (e.g., ecological momentary assessment) generate information that can be used to develop personalized models of an individual's behavior. Techniques are presented that allow for the simultaneous development of many personalized models and searches for shared features across models.

Focusing Personality Assessment on the Person: Modeling General, Shared, and Person Specific Processes in Personality and Psychopathology

Contemporary psychological assessment relies heavily on models of personality and psychopathology that were developed using nomothetic principles. In this top-down approach, individuals are compared to normative distributions on one or more dimensions of functioning (e.g., extraversion, anxiety) to ascertain their position relative to others. This approach is rooted in well validated models of individual differences (e.g., the Big Five). Yet it also yields an incomplete picture of an individual's relevant processes, because the same model is applied to all individuals. As currently implemented these models are static (i.e., based on behavioral tendencies assessed in cross-section) and human behavior¹ is dynamic (i.e., plays out over time and different situations). Recently there has been a push towards collecting data in ways that might reveal important dynamic processes using techniques, such as ambulatory assessment, that involve sampling individuals intensively and repeatedly over time and situations (Hamaker & Wichers, 2017). This trend has been accompanied by calls to use idiographic statistical techniques to develop “personalized models” of individuals' processes from the many repeated assessments collected (Fisher, 2015; Molenaar, 2004; Roche et al., 2014; van Os et al., 2013; Wichers et al., 2011; Wright et al., 2016). These echo Allport's (1937) much earlier call that personality psychology should focus on the study of the whole person as the unit of analysis (i.e., personology). Here we present a statistical approach, group iterative multiple model estimation (GIMME; Gates & Molenaar, 2012), that is designed to bridge the idiographic and nomothetic levels of analysis by estimating person-specific models from intensive longitudinal data, and then searching for commonalities across models.

¹ Here we use the term “behavior” very generally, including thoughts, feelings, and overt behavior.

How personality and psychopathology are conceptualized and studied has implications for how they are assessed. Models of both the basic structure of personality (e.g., the Big Five/Five-Factor Model; Costa & McCrae, 2008; Goldberg, 1990) and the emerging structure of psychopathology that dovetails with the structure of basic personality (Kotov et al., 2017; Wright & Simms, 2015) are grounded in between-person statistical analyses. As such, they readily translate to developing inventories to assess an individual's personality profile or elevated features of psychopathology, especially when considered relative to their peers. These assessment instruments can be administered via self- or informant-report, or as an interview by a trained clinician. Using the resulting data, predictions can be made about how the assessed individual might behave across a large range of situations, or summarize the experiences someone has had over circumscribed period of time (e.g., the past six months).

However, structural models of personality and psychopathology are descriptive accounts of individual differences, providing little explanation for how these differences manifest. In contrast, contemporary personality theories elaborate within-person dynamic processes that putatively give rise to the major individual difference domains (e.g., Baumert et al., 2017; DeYoung, 2015; Fleeson & Jayawickreme, 2015; Pincus, Hopwood, & Wright, in press). Thus, what we recognize as an individual's personality is actually their particular patterning of behavior as it fluctuates across time and in response to situations (Wright, 2011). A corollary of these theories, is that the structure of an individual's personality is not wholly reducible to their profile of behavioral means, but rather should include the associations (concurrent and lagged) among behavioral states and relevant environmental features. Likewise, although often assessed using cross-sectional methods, psychopathology is usually defined explicitly in terms of dynamic processes among several mutually influencing mental and physiological systems, and failures or maladaptive efforts to regulate these systems. For instance, compulsive behaviors emerge as one

tries to gain control over severe anxiety and obsessive thoughts. In borderline personality pathology, perceptions of abandonment lead to intense negative affect and behaviors to keep a partner present and engaged. Although at the between-person level psychopathology can be organized into similar domains to those in basic personality, at the within-person level, cardinal features often involve interactions among different domains. Behavioral treatments seek to disrupt these within-person processes, making them desirable targets for direct assessment.

Personality and psychopathology are increasingly being measured using ambulatory assessment techniques, sampling individuals intensively and repeatedly in their naturalistic environment (e.g., Moskowitz, Russell, Sadikaj, & Sutton, 2009). By virtue of the intensive repeated measurement, individuals are assessed at different points along a temporally unfolding process which allows for modeling of dynamic processes (Hamaker & Wichers, 2017). Modern mobile computing allows for nearly ubiquitous access to an individual's behavior, which can be sampled via self-report survey, multimedia capture (e.g., brief videos, photos), or passive sensing devices (e.g., GPS, ambient noise from the microphone). For our purposes, the key contribution of intensively repeated assessments is the ability to capture fluctuations of personality and psychopathology states over brief periods of time. As the nuances of data capture are outside the scope of the present paper we refer readers to several good reviews (e.g., Carpenter, Wycoff, & Trull, 2016; Shiffman, Stone, & Hufford, 2008; Trull & Ebner-Primer, 2013).

Dynamic assessment data, as gathered via some ambulatory assessment technique, often generates large amounts of observations for each individual. When this is the case, the resultant multivariate time-series can be used to develop person-specific statistical models. In basic personality, there is longstanding precedent for using such data to develop idiographic models, that provide an idiosyncratic personality structure (e.g., Borkenau & Ostendorf, 1998; Cattell, Cattell, & Rhymer, 1947; Zevon & Tellegen, 1982). More recently, there have been numerous

calls for using ambulatory assessment data coupled with idiographic analyses to develop “personalized” models of psychopathology (e.g., Fisher, 2015; Litten et al., 2015; Roche et al., 2014; van Os et al., 2013; Wright et al., 2016). This is consistent with a broader push towards “personalized medicine” across clinical fields (Hamburg & Collins, 2010). Several compelling examples have been published in recent years (Bak et al., 2016; Wichers & Groot, 2016).

As enticing as the idiographic modeling approach is, it suffers from two main impediments, scalability and generalizability, that have traditionally limited its wide use in research (Connor et al., 2009). First, because a model needs to be developed for each individual, it is cumbersome to implement in any large-scale sample. Second, by maintaining such a crisp focus on the individual, it struggles to contribute to generalizable principles of personality and psychopathology (Beltz, Wright, Sprague, & Molenaar, 2016; Spencer & Schöner, 2003). Thus, in many respects this is an incomplete approach. On the one hand, person-specific models might be informative in clinical settings, when the individual’s processes are the focus. On the other hand, not being able to compare the results of that personalized model to any other in any systematic way is a practical limitation (e.g., are the processes in the normative range, or are they similar to an identified clinical group?). These are most likely why person-specific models have not caught on in the broader literature.

What is needed is a method that can bridge the idiographic and nomothetic divide, providing efficient personalized analysis of an individual’s dynamic associations among relevant behavioral domains, while also nesting these person-specific analyses within a broader nomothetic structure. Such an approach should allow for the presumably high-heterogeneity in processes across individuals, while searching for commonalities when present. This could contribute to bottom-up development of behavioral and psychopathological taxonomies, seeking out patterns of processes that are shared among individuals. Here we introduce a method that

satisfies these needs, *group iterative multiple model estimation* (GIMME; Gates & Molenaar, 2012). Originally developed for use with functional MRI brain data, we demonstrate the utility of this approach on EMA data gathered on a daily basis in the present paper.

GIMME

Few options exist for analyzing intensively gathered longitudinal data for multiple individuals in a manner that allows for individual-level heterogeneity as well as general patterning of the sample. The most common approach is multilevel modeling, where each time point (i.e., observation) is considered to be nested within individuals. Indeed, Connor et al. (2009) argued that multilevel models (MLM) effectively bridge nomothetic and idiographic approaches, because they simultaneously estimate within-person associations among variables (e.g., daily stress→negative affect) and between-person differences in those associations. The main benefit of this approach is that it provides an estimate for each individual based on an assumed distribution derived from the sample as a whole. When the data conform to this assumption, the estimates may be better than when conducting individual-level analysis (Liu, 2017). A drawback is that everyone must have the same *pattern of relations* which must be specified *a priori*. Large models with more than a few outcomes can be difficult to estimate within a single model using MLM. Moreover, they make strong assumptions about the distribution of such effects (e.g., continuously and normally distributed), which when violated can lead to biased estimates and erroneous inferences. These can be altered, but must be done so prior to modeling and still places restrictions on the individual-level estimates.

Person-specific, data-driven methods for model selection circumvent the need for specific hypotheses and allows for unrestricted individual-level variation in model structures. In this way, they are well-suited for assessing individual-level processes. These purely individual-level

approaches can detect the pattern of relations among variables with no need for prior information and no assumption of equivalence in that patterning across individuals. While a number of approaches for arriving at data-driven relations for individual data have been proposed (see Henry & Gates, 2017; Smith et al., 2011), relatively few individual-level data-driven approaches and methodology have been developed and extensively tested. A major criticism of these approaches is that they can overfit the data, meaning that spurious effects might be added that simply model noise rather than a true aspect of the person's process.

GIMME has emerged as a promising approach for conducting a multivariate search for the pattern of associations among intensively sampled data. It has five main qualities that set it apart from other methods. For one, GIMME has been extensively evaluated in simulation studies using benchmark individual-level data (Gates & Molenaar, 2012; see Smith et al., 2011). Results indicate that GIMME consistently and reliably can detect the true patterns of effects even in heterogeneous samples at higher rates than competing methods. Two, enabling individual-level paths provides a marked benefit over MLM, which requires that all individuals have the same pattern of relations. Three, the approach for identifying the pattern of relations for individuals has a strict stopping criterion that greatly attenuates the risk of overfitting. Much like other successful data-driven approaches (Mumford & Ramsey, 2014), GIMME does not add all effects that are significant but rather uses a set of stopping criteria that favors parsimony to guard against overfitting. Four, each effect is estimated separately for individuals with no assumptions placed on the distribution of estimates. This is true even for effects that exist and are estimated for all individuals. So, much like multilevel modeling there will be a unique estimate for each individual, but the estimates obtained with GIMME will be unrestrained. Five, a recent innovation clusters individuals based on their patterns and estimates of effects (Gates et al., 2017). This greatly aids in understanding the varied patterns of results that may occur across

individuals by finding sets of paths that co-occur in subsets of individuals. It also may identify subgroupings that meaningfully differ in their psychological processes.

GIMME is based on a unified structural equation modeling (uSEM; Kim et al., 2007) framework. uSEM provides estimates for both lagged and contemporaneous effects. In this way, it bridges the two most common methods used for analyzing time-series data: vector autoregression (VAR) and structural equation modeling (SEM). VAR models estimate lagged influences among variables. For instance, one can see how disinhibition and recklessness on one day predicts negative affect on the next day. Additionally, autoregressive effects are modeled, which indicates how each variable predicts itself at the next time point. These have recently been highlighted as being very useful in psychological contexts (Bulteel, Mestdagh, Tuerlinckx, & Ceulemans, 2017), and provide a solid foundation for the model search to begin. Importantly, no additional effects will be added if these effects sufficiently describe the data. Structural equation models provide complementary estimates of contemporary, directed (i.e., not correlational) relations. As an example, perhaps levels of stress on a specific day can be used to statistically predict levels of negative affect on that same day, indicating that on days where one is stressed they are also likely to have negative affect after controlling for other influences on negative affect.

Such contemporaneous effects can occur when the mechanism being studied changes faster than the rate of data collection (Granger, 1969). This is likely to occur in the assessment of behaviors in a clinical population as they may shift quickly (e.g., on the order of minutes) whereas the frequency of data collection may be larger (e.g., on the order of hours or days). Here, the underlying behaviors might shift in the time between data collection time points, preventing the assessment of lead-lag relations since the causal mechanism occurred in the time between assessments. In these cases, one can only assess if having a given emotion at a given

time (e.g., negative affect) is likely to occur when another attribute is present (e.g., stress). The direction would indicate that one attribute tends to statistically predict another attribute after taking into account other influences on the target attribute. For instance, if we find that stress tends to statistically predict reports of negative affect after controlling for the autoregressive influence (e.g., the prior day's negative affect predicts current negative affect), we can ascertain that knowledge of the person's stress level reveals insight into that person's current level of negative affect. The reverse may not be true; the person's negative affect may not contemporaneously predict stress after controlling for the autoregressive influence of stress. When both lagged and contemporaneous effects exist in the underlying process they both must be modeled to prevent spurious findings and ensure reliable inferences (Gates et al., 2010).

Empirical Example

To demonstrate the GIMME approach applied to ambulatory assessment data, we present an empirical example using daily diary data collected over 100-days in a sample of individuals diagnosed with personality disorders. Our goal is to illustrate the types of findings GIMME can generate when used with dynamic personality data. Therefore, we chose a set of daily behavioral variables that correspond to basic domains of personality, as well as a measure of daily stress, and a measure of difficulties with daily functioning. The behavioral domains we included were interpersonal dominance vs. submissiveness, affiliation vs. separation, negative affect, and positive affect. These analyses were largely exploratory; that is, we did not hypothesize the emergence of specific group-level or subgroup-level effects. However, it seemed plausible that we might observe a group-level association between stress and negative affect, and positive affect and dominance. Moreover, we had a descriptive aim of determining just how heterogeneous the individual models might be.

Method

Participants

The sample used in this study was collected as part of a project designed to investigate daily processes of behavior in individuals with personality disorders. More detailed information about recruitment and sample characteristics can be found in Wright, Hopwood, and Simms (2015), Wright, Beltz, et al. (2015), and Wright and Simms (2016), which also used this same sample. One hundred and sixteen participants attended the baseline assessment for the daily diary study and were enrolled in a 100-day daily diary protocol. To ensure accurate estimation of each individual's model, participants providing fewer than 60 daily observations were excluded (Lane et al., in press). Only 19 individuals were excluded for providing less than 60 diaries. Three additional participants reported no variability in daily functioning, and were excluded because models cannot be estimated for a participant when a variable is a constant. Thus, the effective sample size for the GIMME analyses was $n = 94$. All of the individuals had at least one personality disorder diagnoses, with many having a Borderline Personality Disorder diagnoses (39%) and/or Narcissist Personality Disorder (20%).

Procedure

A complete description of the study was provided and written informed consent was obtained prior to participation. The University at Buffalo institutional review board approved all study procedures. Participants were asked to complete daily diaries via secure website every evening for 100 consecutive days. Surveys were to be completed at (roughly) the same time each day, between 8pm and 12am. However, participants were allowed to deviate from this schedule if necessary (e.g., working nightshift) so long as (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 email reminders and also were 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 $\geq 80\%$ participation, and prorated at \$1/day for $< 80\%$. Participation also was incentivized through recurring raffles (\$10 drawing every 5 days for those providing at least 4 diaries) and drawings for additional money and tablet computers at the end of the study, with the odds of winning proportionally tied to participation.

Measures

Daily Affect. Daily affect was measured using a subset of Positive and Negative Affect Schedule (PANAS; Thompson, 2007) items. The PANAS uses a 5-point scale (*very slightly or not at all, a little, moderately, quite a bit, and very much*) for participants to rate mood states. Participants were asked to report on their mood “over the last 24 hours.” Daily positive affect was measured as the mean of the following five items: *Active, Alert, Attentive, Determined, and Inspired*. Daily negative affect was measured as the mean of: *Afraid, Ashamed, Hostile, Nervous, and Upset*. In the full sample, the resulting affective domains were uncorrelated ($r = .04$) when pooled across participants.

Daily Interpersonal Behavior. Daily interpersonal behavior was measured using a subset of the Interpersonal Adjective Scales (Wiggins, 1995) items. One adjective from each octant-scale was provided (e.g., *Assertive, Critical, Indifferent, Introverted, etc.*), and participants were asked to rate how well each term described their social behavior over the past 24 hours using an 8-point scale ranging from *Extremely Inaccurate* to *Extremely Accurate*. Daily dominance and affiliation scores were calculated after first subtracting daily mean endorsement to control for overall endorsement, followed by combining the scores based on circumplex weights. In the full sample, this resulted in two essentially orthogonal dimensions ($r = .07$) of interpersonal behavior when pooled across participants.

Daily Stress. Daily stress was measured using a self-report version of the Daily Inventory of Stressful Events (Almeida et al., 2002), which consists of seven questions that ask

whether specific stressful events have occurred within the last 24 hours. The events include: 1. *Having had an argument or disagreement with someone*, 2. *Something occurring that could have led to an argument or disagreement but it was allowed to pass*, 3. *A stressful event at work or school*, 4. *A stressful event at home*, 5. *Experiencing discrimination on the basis of age, sex, or race*, 6. *Something stressful happening to a close friend or relative*, 7. *Anything else that most people would consider stressful*. Endorsed events then were rated for severity on a 4-point scale with the anchors of *Not at all*, *Not Very*, *Somewhat*, and *Very*. We used the sum of the rated severity, across items, as an index of daily perceived stress.

Daily Functioning. Difficulties with daily functioning were assessed using a single item, referenced to the last 24-hours: *How much difficulty did you have in taking care of important tasks or responsibilities?* This item was rated on a 5-point scale with the anchors of *None*, *Mild*, *Moderate*, *Severe*, *Extreme (Could not do)*.

Analytic Plan

GIMME

All GIMME analyses described here were conducted using the ‘gimmeSEM’ function provided by the *gimme* R package (Lane, Gates, & Molenaar, 2017). Details of the algorithm can be found elsewhere (Gates & Molenaar, 2012; Gates et al., 2017) and tutorials exist for the interested user (Beltz & Gates, 2017; Lane & Gates, 2017). We provide a brief description here and suggest the reader refer to these documents for further details. The GIMME algorithm begins by first searching for lagged and contemporaneous relations (i.e., regression paths) that would significantly improve the uSEM fits for the majority of individuals (or dyads or other unit of analysis). The default for what constitutes the majority is 75%, but this can be altered. Significance is assessed by looking at the modification indices for each putative path and correcting for multiple comparisons. Importantly, GIMME selects paths that are significant for

individuals when the models are estimated separately for each individual (i.e., no aggregation takes place). This follows from the perspective that each individual is a unique sample and the goal is to identify effects that consistently occur across these samples. Whereas outliers with particularly strong effects may drive results from other aggregation approaches, GIMME ensures that all individuals provide equal contribution to the selection of group-level effects such that these effects are descriptive of most of the group. Additionally, by only looking at significance this allows for paths to be added that might be positive for some individuals and negative for others. The only criterion is that it be significant for the majority.

After obtaining the pattern of group level effects there is the option to cluster individuals into subgroups based on the estimates of the group-level paths and the patterns of effects that are likely to emerge at the individual level. Since identifying group-level paths has been shown to improve upon the reliability in recovering of individual-level paths (Gates & Molenaar, 2012), it follows that identifying paths that replicate across a subset of individuals may further aid in model building by getting the model closer to the final pattern (MacCallum, 1993). Hence subgrouping serves two purposes: to further refine the data-search algorithm and to identify patterns of paths that consistently occur for a subset of individuals. That is not to say that it will identify firm subtypes or classes, but rather group individuals who share similar patterns of coefficients. The clustering option first generates a similarity matrix based on the path estimates and pattern of relations and then utilizes a community detection algorithm called Walktrap (Pons & Latapy, 2005) to place individuals into subgroups with other individuals that have similar estimates and patterns of effects. Walktrap was chosen since it has a number of benefits over competing approaches. For one, it does not require that the researcher indicate the number of subgroups. Two, it has been shown to reliably recover both the true number of subgroups and place individuals into the appropriate group for data similar to the similarity matrix used here

(Gates, Henry, Stienly, & Fair, 2016). Three, Walktrap performs well even when the sample size is small and there are small subgroups. Note that the Walktrap algorithm is not based on a mixture distribution approach (Hallquist & Wright, 2014), and therefore does not provide class membership probabilities or similar metrics of an individual's fit to a group pattern.

Once subgroups are obtained the same steps used to arrive at group-level paths is conducted again but this time within each subgroup. Here, 50% is used as the majority cut-off (this can be altered if desired). This culminates in paths that are specific to subgroups in addition to the previously found group-level paths. Using the group- and subgroup-level paths as priors, GIMME then searches for individual-level effects. This process continues until an excellent fit is found (i.e., the model appropriately describes the variance and covariance found in the data). At the end of the algorithm, all individuals have unique estimates for the group- and subgroup-level paths, in addition to unique path structures (Gates, Lane, Varangis, Giovanello, & Guskiewicz, 2017). These effects enable general (group), shared (subgroup), and person-specific insights. It must be noted that individuals within subgroups may also share similarities with individuals in other subgroups. The classification is mainly done to improve the model search and help to parse similarities in subsets of individuals. These similarities in patterns of relations may or may not indicate a subset of individuals that is distinct from other individuals in terms of their psychological processes.

GIMME has been tested on data with between 5 and 15 variables. Currently, GIMME is similar to network approaches (e.g., Fried & Cramer, 2017) by modeling only observed variables. The daily diary variables included in each individual's model included daily ratings of dominant behavior (*Dominance*), affiliative behavior (*Affiliation*), positive affect (*PA*), negative affect (*NA*), experienced stress (*Stress*), and difficulties with performing daily tasks (*Functioning*). These six variables, in the 94 participants previously diagnosed with PD, were

submitted to GIMME as implemented in the freely distributed R package *gimme* (Lane, Gates, & Molenaar, 2017). For the current model, we used the *gimme* package defaults as these have been shown to reliably recover true effects (Gates & Molenaar, 2012; Gates et al., 2017; Lane et al., in press). These defaults are: freely estimating all autoregressive paths at the group level (i.e., for all individuals in the sample), a cutoff of .75 for group-level paths (i.e., indicating that a path must be significant in 75% or more of the individual models in order for it to be included as a group-level path), and a cutoff of .50 for subgroup-level paths (i.e., indicating that a path must be significant in 50% or more of the individual models in a subgroup in order for it to be included as a subgroup-level path). We then explored subgroup differences in estimated paths and external variables to identify if meaningful inferences could be made. All *gimme* input syntax, individual data files, output files, and diagrams are available at the following website: <https://osf.io/95hyr/>

Results

As noted above, although each model is ultimately estimated individually, there are (potentially) three sets of paths contributing to each model. First, the group-level paths are freely estimated for all individuals in the sample. In our model, these included both the autoregressive paths for each variable that were included in the model from the start, in addition to any concurrent or lagged associations that the model “discovered” in its iterative searches. Second, there are the subgroup paths, which are paths shared (yet freely estimated for all members) among a subgroup of the sample. Finally, individual-level paths are those paths that are freely estimated for each individual, but do not rise to the thresholds for subgroup (50%) or group (75%) level paths. We describe results at each level. Figure 1 depicts all estimated paths, including group (black), subgroup (green), and individual paths (grey). To elaborate on the visual presentation of the results, Table 1 contains two matrices that summarize (a) the number of individual models each path was included in, (b) the mean coefficient (i.e., regression weight)

across models in which each path was estimated, and (c) the standard deviation of estimated paths. Recall that the total $N=94$, therefore any path with 94 as the number of models it was estimated in was a group level path.

Group Model (General Paths)

Group level paths are depicted in black in Figures 1 and 2. All six autoregressive paths were freely estimated for each individual. For all six variables, estimates ranged from small, non-significant, negative values to large, significant, positive values when looking across individuals (see Table 1). Large significant values for the autoregressive effects indicate relative consistency across days for a given variable (holding constant the influence of any other variables that predict the target variable). Conceptually, the AR weight has been described as a measure of inertia (Kuppens, Allen, & Sheeber, 2010) or stability in the variable of interest. This does not imply absolute stability, but rather that a given day's measure can be predicted well by the prior day's measure on that construct. *Dominance* and *Functioning* were the only variables to have significant negative autoregressive estimates for any participant, and these were for two and one individuals, respectively. Negative autoregressive paths have been thought to reflect a feedback system whereby the system cycles between high and low values. Only one additional path surpassed the group-level cutoff: contemporaneous *NA* regressed on *Stress*. The average effect was strong, all individuals exhibited positive estimates for this relation, and all but seven were significant (92.5%). This suggests that in general, experiencing stress predicts experiencing negative affect during the same day after taking into account the lagged influence of negative affect on the prior day as well as other covariates included as needed for each individual.

Sub-Group Models (Shared Paths)

Three subgroups were found, and participants were unevenly divided across subgroups. The largest, Subgroup 1, included 60 individuals, next largest, Subgroup 3, included 23

individuals, followed by Subgroup 2 with 11 individuals. Importantly, the subgroup order and numbering is arbitrary. One must examine and interpret the results for each subgroup to arrive at any inferences. Subgroup plots can be found in Figure 2. Additionally, summary tables similar to Table 1, but specific to each subgroup, can be found in the online supplemental materials.

Starting with Subgroup 1, we found one shared path, Affiliation regressed on NA. Estimates for this path in the sample ranged from small, non-significant, positive effects, to large, significant, negative effects. Thus, for a subset of the sample, affect and interpersonal behavior were associated contemporaneously, such that experiencing negative affect predicted withdrawing from others. Descriptively this is important information – that a relatively large subset of individuals may have this relation. Hence when working with individuals who have personality disorder diagnoses this may be an aspect of their emotional and behavioral processes to attend to and assess.

Subgroups 2 and 3 provide interesting examples of groupings of individuals that share no paths (with the exception of the group-level paths all in the sample share). Recall that the subgrouping algorithm within GIMME uses the total similarity of individuals' model patterns, not solely based on the presence of shared paths (Gates, Lane, et al., 2017). As a result, it is possible for the algorithm to form groups based on overall similarity of their group-level estimates, but lacking in any shared paths that cross the 50% threshold. Subgroups 2 and 3 are examples of such groups.

We recognize that a feature that is intended to provide interpretive facility of the resulting models does not guarantee easily observable similarities. When this occurs, researchers may be interested to examine differences in paths across groups statistically in order to clarify shared features within a subgroup. Here we explored some differences on baseline personality pathology features and differences in group-level paths to illustrate how one might go about this.

We found that the largest subgroup (number 1) had lower autoregressive estimates than Subgroup 2 on stress and PA, as well as a lower average AR for affection than seen in Subgroup 3 (See Table 2). Perhaps most interesting were the differences on the *Stress*→*NA* path ($F_{2,91} = 17.89, p < .001$). Post hoc tests revealed that Subgroup 2 had a lower average estimate for the path from Stress to NA when compared to Subgroup 1 (average difference = $-0.43, p < .001$) and Subgroup 3 (average difference = $-0.40, p < .001$). This suggests that Subgroup 2 has a particularly lower propensity to have negative affect on days in which they experience stress (controlling for other variables in their model). Hence, we can see that the estimates contributed to the separation of individuals into subgroups.

We also wanted to see if the subgroups related to constructs of interest. Since borderline or narcissistic personality disorder diagnoses were common in the sample, we focused on these variables. Interestingly, Subgroup 1 contained almost all of the individuals who had a narcissistic personality disorder diagnosis ($n = 15$ out of $19; \chi^2 = 5.0, df = 2, p = .082$) and a significantly larger number of individuals with borderline personality disorder than would be expected by chance ($n = 28$ out of $37, \chi^2 = 6.4, df = 2, p = .041$). Taken together, inferences from these results suggest most individuals with a personality disorder may have a relation between experiencing stress and negative affect on the same day. However, this relation varies across individuals. As seen from the subgroup results, those with borderline and/or narcissistic personality disorder diagnoses may be more likely to withdraw from others when experiencing negative affect. In this way, the subgroups help to identify patterns that exist for a subset of individuals that may be meaningful or offer new insights.

Individual Models

Person-specific (i.e., individual-level) models exist for each participant in the sample. A

diagram and path coefficients for each participant can be found in the online materials, and here we present a selection of diagrams in Figure 3 for demonstration purposes. Each participant's map is detailed and invites in-depth interpretation, as might be expected when working with an individual in an applied setting (e.g., assessment before, during, or after psychotherapy, consultation in industry). This sort of detailed interpretation for each individual goes beyond the scope of this presentation. Rather, we wish to draw the reader's attention to several aspects of the obtained person-specific models and use the examples presented in Figure 3 to guide this discussion. In following with color schemes often used in heat map depictions of networks, red here indicates positive ("hot") values with blue reflecting negative ("cool") values. Path width corresponds to the absolute value.

In terms of overarching observations, note that the various presented models differ in their density. That is, some individuals have more significant paths than others, suggesting that daily interpersonal behavior, affect, stress, and functioning are more intertwined, whereas for others these variables function separately or are governed by separate processes. Also, note that lagged effects, which were absent in the general and shared paths, now emerge with greater frequency. This suggests that these variables do, in fact, carry over from day to day in their effects, but that there is high heterogeneity in these effects, with one individual sharing little in common with the next. In line with that, notice the high degree of heterogeneity across individuals in their global structures. A single model would fail to do this diversity justice. It is also useful to compare any given model to the summary matrices in Table 1. Those matrices also highlight the high degree of heterogeneity, and also illustrate that certain paths are much more common than others, even if they do not rise to the level of subgroup or group paths.

Inspecting each person's model reveals interesting hypotheses about their daily processes. For instance, some individuals have lagged links between the same two variables, but the effects

go in opposite directions, suggesting different dynamics. To the extent that Participant 55 experiences NA, the next day s/he engages in dominant behavior, whereas for Participant 51 the opposite is true. Or, take Participant 69, who appears to have an amplification dynamic between stress and NA, such that the more stress s/he experiences, the more s/he experiences NA, which leads to more stress the next day. This sort of vicious cycle is commonly the target of therapeutic interventions. In contrast, Participant 69 exhibits what appears to be a regulatory process between dominance and daily stress, such that increases in dominance lead to more stress, which in turn leads to greater submissiveness, presumably as s/he seeks to interrupt or arrest the process. Participant 86 also has a pattern of paths suggestive of a regulatory process between affiliation and NA. We also see variability in the estimates for $NA \rightarrow Stress$. While all of the individuals shown here had positive estimates for this group-level path, we see that individual 80 has a notably smaller value (as seen in the thinner lines). Other models share several features (e.g., 80 and 94), but differ in others. Some models demonstrate what appear to be the central importance of a specific variable (e.g., NA in participant 55), and other models differ in the overall density and complexity (e.g., compare 86 to 51). Many more specific interpretations could be made about just these six individuals. Our goal is not to interpret every parameter, but instead demonstrate the richness of the emergent person-specific models, and show their relevance to the dynamics of interest in understanding how individuals function in daily life.

Discussion

There is substantial field-wide interest in measuring people intensively and repeatedly to capture dynamic personality processes. The resulting data, which often involve many observations per person, have motivated calls for personalized or person-specific modeling. This

poses a challenge, how to develop a nomothetic science based on idiographic models? GIMME provides a potential solution to this thorny issue, and we have demonstrated it in a sample of patients who completed a long (100-day) daily diary protocol. Our main goal was not to draw firm conclusions about general personality processes. Instead, we sought to provide a proof-of-concept of GIMME with subgrouping capabilities in a large and heterogeneous sample using major domains of personality (affiliation, dominance, PA, NA) and commonly associated variables (stress and dysfunction).

Although our empirical example was not motivated by specific study questions, we nonetheless found interesting results that point to what can be expected in future work in this area. In the results section, we summarized our findings going from the general to the specific, following GIMME's order of operations. Here we reverse the order, beginning with the person-specific level and expanding to the general-group level. Perhaps the most notable finding is the high-degree of heterogeneity across person-specific models. This highlights the motivation for idiographic or personalized models of dynamic personality assessment data: people differ from each other in the structure of their specific processes. Figure 3 illustrates this to some degree, and we encourage the interested reader to peruse the full set of figures (hosted online at <https://osf.io/95hvr/>) to gain a full appreciation of the individual differences in these models. Although the current sample was highly heterogeneous by design, we anticipate based on past experience using GIMME in several data sets of different sorts (Beltz et al., 2016; Price et al., 2017) that massive heterogeneity is likely to be the rule rather than the exception.

The subgrouping feature provides one way to parse through this heterogeneity, identifying individuals that appear to be governed by similar processes. In line with this, we identified a number of subgroups that differed from each other in notable ways. One subgroup was demarcated by shared paths, whereas the others stood out for their lack of shared paths.

Subgroups differed in their size considerably ($ns = 60, 23, 11$). The largest subgroup had only a single shared path, which was a contemporaneous association. As mentioned, two groups had no shared paths, which highlights that the subgrouping is not based solely on the presence of shared paths, but also on the individual estimates of the group-level paths (i.e., direction and strength) and the overall pattern of paths. Shared paths offer a natural first step towards interpreting the groupings, as the overall networks are likely to be too complex to interpret by eye. However, when subgroups lack shared paths to facilitate interpretation, comparisons on external validators may be needed. Here we explored several baseline personality disorder features and path features, and identified differences in narcissistic and borderline features that distinguished group 1 from the others, and several group differences in the strength of lagged effects. We emphasize that these were exploratory and are but a few of the possibly approaches one might take understanding group differences when they are not transparent. We further reiterate that these subgroups are not intended to identify firm subtypes, but rather to provide another tool for parsing through the heterogeneity and facilitating interpretation and comparison of individuals.

A single contemporaneous path emerged as the only group-level effect, daily perceived stress predicted NA. In other words, knowing someone's daily level of experienced stress is informative of their level of NA above other predictors in the model, but the opposite is not true. That there was only a single group-level effect speaks to the heterogeneity in patterns in the person-specific models. Group-level paths are important since, by definition, they replicate across individuals in the sample. Hence these are paths that might be expected to be found across other samples from the same population. In fact, NA predicted by perceived stress (after accounting for its autoregressive effect) is a well-established finding in the daily diary of personality and stress literature (e.g., Bolger et al., 1995; Mroczek et al., 2004). Thus, this finding is consistent with prior literature in other distinct samples using different analytic

methods, thereby lending confidence that our group-level results would replicate in other samples.

An important note is that GIMME likely provides more reliable group-level relations than other approaches for data-driven models. For instance, one could feasibly do model selection search with MLM. However, this is less common and not fully evaluated. The researcher would have to make critical choices, such as does one conduct forward-selection or backward-deletion of potential relations among variables? Because there were so few group-level paths in the present example it is highly possible that spurious relations would result (Molenaar & Campbell, 2009). Since MLM estimates paths for all individuals, none of the individual-level paths would be considered – unless spuriously estimated for the sample, which would likely attenuate estimates since they would be constrained to follow a normal distribution with a mean near zero (since for most individuals, zero would be the estimate). Another popular option is to concatenate individuals' data such that one person's time series is pasted above another's and this continues until the whole sample is in one data set. The problem with this approach is that a few outliers could drive all of the results. GIMME circumvents these issues by requiring that all group-level paths be significant for the majority of individuals.

Potential Applications for GIMME of Dynamic Personality Data

GIMME offers a number of exciting avenues in the domain of dynamic personality assessment. Perhaps the most obvious application is to data collected from putatively homogenous groups, such as diagnostic categories. It is well established that even the best treatments for behavioral disorders only successfully treat a portion of afflicted individuals, and data-driven approaches such as GIMME may help to identify match effective interventions to specific behavioral processes. However, it is relatively unknown how heterogeneous the behavioral processes are of individuals who share a diagnostic category. Our results here would

suggest that they are likely to be highly heterogeneous, although we did not limit our search to a group with a circumscribed diagnostic profile. It is possible, if not likely, that more characteristic group- and subgroup-level paths might emerge in more homogenous samples.

Moreover, subtyping clinical diagnoses based on traditional symptom measures has been largely unsuccessful. Subtyping based on shared processes might be more successful at identifying individuals who share meaningful processes. For instance, when it comes to alcohol use, it is difficult to identify who will go on to develop problematic use over time. In early adulthood, heavy use is normative among many, but only some develop lasting patterns of misuse. Hypotheses related to who is at elevated risk, even among heavy drinkers, are related to dynamic processes of use (e.g., positive reinforcement vs. negative reinforcement pathways). Early identification of cycles of positive and negative reinforcement would allow early identification of at risk individuals, which may not be evident from profiles of use, consequences, or other cross-sectional assessments. Similar examples could be developed for other issues. We wish to be clear that we view GIMME's subgrouping capabilities as a useful tool for identifying individuals with similar networks of processes, but we do not necessarily assume these groupings will represent formal subtypes. Nevertheless, developing and refining clinical phenotypes based on shared patterns of processes is a promising avenue for future work.

On a more basic level, GIMME could be used to investigate heterogeneity in processes of individuals with the same trait or symptom profiles. Currently, it is assumed that individuals who share similar profiles of traits or clinical features are likely to behave similarly. But, as discussed in the introduction, cross-sectional assessments of personality do not directly capture the dynamics that are presumed to give rise to what are recognized as traits. As such, it remains an open question whether and to what degree individuals with similar profiles share similar patterns of dynamic processes. The processes could be examined at a fine-grained level,

incorporating specific indicators from within a single trait domain (e.g., facets). Alternatively, this question could be expanded to incorporate broader range of traits to understand how they interact with each other. This type of approach could unpack gross statements like “narcissists are disagreeable extraverts,” because not all disagreeable extraverts (or extraverted antagonists for that matter) are likely to be narcissistically organized (Wright, 2011).

More broadly, GIMME might be used any time one is interested in studying patterns of dynamic processes in highly multivariate systems. For example, contemporary taxonomies of situations use up to eight situational assessments (Rauthmann et al., 2014), which when included with the five canonical traits, means studying 13 variables over time. Running individual models to examine each of the possible paths would generate a dizzying number of results to be interpreted (e.g., Rauthmann, Jones, & Sherman., 2016). Furthermore, it takes the focus away from the person and places it on the variables. GIMME allows for these to be analyzed simultaneously, retaining the focus on the person.

Implications for Applied Assessment

As we highlighted in the first paragraph, there is now a great deal of interest in using ambulatory assessment to develop personalized models of an individual’s psychological and behavioral processes. This is particularly the case in clinical settings (i.e., mental health), although this could be extended to occupational or other healthcare settings. The main engine of the GIMME approach, uSEM, can be estimated for an individual (For exemplars see Beltz et al., 2016; Fisher et al., in press; or Foster & Beltz, in press), and the GIMME software package includes this capability. This makes it feasible to apply this model using the same tools in practical settings using easy to use and freely available software.

However, a challenge with this approach is that fully idiographic modeling of an individual in a high-stakes setting risks leaving the practitioner adrift when it comes time to

interpret the results. Enthusiasm for understanding the individual should not unmoor fundamental nomothetic principles that play such an important part of practical assessment. Running a full GIMME in a large sample provides the necessary tethers to anchor the interpretation of an individual to a firm foundation. Specifically, if the variables used in the current study were given to someone in an applied setting, we would expect to find that there was a link between daily stress and NA, and we could compare the individual's coefficient to the distribution found here. We could similarly compare any additional emerging paths to those reported in Table 1 to (a) determine how frequently the path is observed in a large sample of patients, and (b) the relative strength of this path. Without these sorts of comparisons, it is difficult to know whether the individual's model is aberrant or merely normative. Many additional ways to compare models might be useful (e.g., overall density of associations).

Practical Guide to using GIMME

GIMME has been tested under a number of conditions to identify when the results are most reliable. Ideally, the researcher will use between 5 and 25 variables (with all individuals required to have the same variables). Simulation studies suggest that reliable results can be obtained with as few as 10 individuals when using regular GIMME (Gates & Molenaar, 2012) or 25 individuals when the subgrouping option is invoked (Gates et al., 2017). GIMME requires the same assumptions be met that exist for typical time series analysis. By that we mean that the data: (1) are collected at approximately equal time intervals; (2) are continuous in nature (i.e., a scale as opposed to categorical); and (3) have constant mean and variance across time (i.e., are stationary). More details on preparing data can be found in Beltz & Gates (2017). At the time of writing this manuscript it is unknown how well GIMME performs when data are not collected at equal intervals or with categorical data. Interestingly, while the GIMME approach assumes that data are stationary across time, it still is able to recover the underlying relations when this

assumption is violated (Gates & Molenaar, 2012). However, it is recommended that if such shifts are expected across time that analysis aimed towards accommodating these changes are conducted. For instance, when the mean level changes across time, this can often be rectified by detrending prior to conducting GIMME analysis.

As with any analysis, the number of observations (in this case, time points) is also an important consideration. Much literature within SEM has focused on rules of thumb regarding the minimum number of observations, particularly as it relates to the ratio of the number of variables and the number of cases (e.g., Bentler & Chou, 1987; MacCallum et al. 1999; Wolf et al., 2013). Specific to GIMME, a Monte Carlo simulation study revealed that GIMME is successful for use on data with qualities similar to those seen in daily diary data when time series are at least 60 observations long and 5 or 10 variables (Lane et al., in press). A related aspect concerns that of missing data. GIMME can still be conducted if there are missing time points for a given individual as long as the researcher indicates the missing data with “NA”. In fact, GIMME can be run if individuals have different numbers of observations for any reason. For more details on conducting analysis with the *gimme* R package and interpreting results please refer to Lane and Gates (2017).

The subgrouping capabilities of GIMME are an option, and not required for estimation of a model. Subgrouping does seem to help recovery of the true paths in a model, but as we have shown here it may or may not lead to easily interpretable groupings. As with all modeling choices, researchers may wish to give careful consideration to whether invoking the subgrouping feature would be advantageous or desirable given their investigative aims. In some cases, this may be a motivating feature for selecting GIMME, and in others it may be superfluous to the research questions. In either case, subgrouping can be used to further identify participants who share similar overall patterns of effects for further interrogation, as we did in our empirical

example.

Challenges and Future Directions

Intensive longitudinal data present unique challenges for the study of psychological and behavioral processes. Although GIMME circumvents a number of these to provide highly reliable data-driven results, four main issues with data have yet to be resolved: how to handle variables with low variability, unequal spacing of assessments across time, and errors in measurement. Low variance occurs when a given participant repeatedly select the same response option, resulting in a constant or low-variance variable. Researchers can attempt to prevent low variance at the time of collection with the development of measures sensitive enough to capture day-to-day variability. However, some measures simply will not have variance for some participants as a result of the individuals' behaviors (e.g., someone who does not drink alcohol will report zero drinks every day, someone who has a highly regimented sleep schedule will report the same number of hours slept each night). As seen in the present study, two individuals had to be removed from analysis because their responses on one measure were constant across time. Other studies using time series or network approaches have also reported removing individuals for this reason (e.g., Lebo & Nesselroade, 1978; Whitley, Ford, & Livingstone, 2000). Current developments in GIMME aim at allowing all individuals to be included even if some of their variables are constant or have low variance.

The second challenge, unequal assessment, is at times inherent to the process of completing daily surveys. It can be requested that participants complete the survey in a given window of time (much like the present study requested completion between 8pm and 12am), but realistically, participants cannot be expected to complete a survey at exactly the same time every day. Another source of unequal spacing occurs when data are collected multiple times per day (e.g., ambulatory assessment designs with random assessments throughout the day). Data might

be collected every few hours when the participant is awake, but of course while the participant is asleep there will be no self-report measurements. For these reasons, solutions for dealing with unequal assessment should focus on corrections at the time of analysis. Many approaches for dealing with unequal spacing have been suggested, including marking the data as missing (van der Heijden, 2006), interpolation (Ram & Nesselroade, 2007), induction of equal spaces (Koopman & Ooms, 2003), and an exponential moving average approach (Ellis & Parbery, 2005). These have yet to be fully tested in the GIMME context.

Third, another future direction is to include measurement models within the GIMME search procedure. Oftentimes self-report data contains measurement error and multiple questions are used to quantify one latent construct. While the use of observed variable models such as the ones provided by GIMME is gaining popularity (Fried & Cramer, 2017), given the nature of self-report data, there will likely still be a need for including a measurement component in some cases. Current efforts aim to improve upon much of the published network literature by using scales composed of multiple items instead of single items to represent complex constructs (cf. Fried & Cramer, 2017).

Fourth, although group- and sub-group level effects are readily interpretable when they emerge, fully exploiting the information contained in the person-specific models is likely to pose challenges for researchers. For instance, one can immediately examine the plots provided by the *gimme* R package to identify nomothetic relations among behaviors of interest, relations that exist among subsets, and the degree to which heterogeneity exists in the individual-level paths. Work is needed on developing useful summary statistics that might allow for feature extraction and association with external variables. One possibility is using graph theory metrics (e.g., network density) when they can be linked with psychologically meaningful constructs. Others have applied these to dynamic data analyzed in different frameworks (e.g., Pe et al., 2015;

Bringmann et al., 2016), but much work remains to be done.

A final consideration is whether an individual's model is representative of his or her behavior beyond the assessment period (i.e., is the model stable across time?). Certainly, questions of measurement invariance are as applicable to intensive samples from an individual as they are from cross-sectional samples of many individuals. Indeed, a main message of the current work is that assessment considerations that are important at the between-person level of analysis remain important when modeling an individual. Whether an individual's model is likely to be more or less invariant over time is likely to track with how more or less representative their assessments were of their general pattern of behavior. For instance, assessing an individual twice while in highly distinct psychological circumstances (e.g., during a manic phase and during a depressed phase) may lead to expectations of non-invariance. The goal might even be to compare the individual's model across assessments to establish the differences. Alternatively, across repeated assessments of the same psychological state (e.g., random bursts of assessing the same individual), the expectation would be that the models would be identical or close to it. We draw the reader's attention to the fact that each individual's coefficient has a standard error, reflective of sampling variability in the estimate. This is unlike MLM, where each individual has an estimate, but no distribution of that estimate. Thus, like between-person statistics, when modeling the individual one must keep in mind whether and how representative the sample is of the population. Future work could implement time-varying effects by design if relevant.

Conclusion

The study of personality and psychopathology, and therefore their assessment, must move beyond a nomothetic approach, whereby individuals scores are only understood as relative to others. With the rapid development of ubiquitous computing and ambulatory assessment techniques, intensive sampling of behavior can now be used to develop personalized (i.e.,

idiographic) models of personality. This is particularly exciting, because these are models of an individual's dynamic processes (Wright & Hopwood, 2016). Dynamic processes are often inferred from static (i.e., cross-sectional assessments), but this is not appropriate and require dynamically sampled behavior (Bos et al., 2017). However, the challenge posed by these models is that they are person-specific, and therefore, can quickly become unwieldy. GIMME provides a way forward, by searching for general, shared, and person-specific contemporaneous and lagged associations in intensive data. Thus, it can be used to build generalizable models from the bottom up. Here we presented a proof-of-concept of GIMME in dynamic personality data in a clinical sample. Although a number of issues remain to address in continued model development, GIMME offers a promising new avenue as personality shifts from the static to dynamic and places the person at the center of personality assessment.

References

- Almeida, D. M., Wethington, E., & Kessler, R. C. (2002). The daily inventory of stressful events: An interview-based approach for measuring daily stressors. *Assessment*, 9(1), 41-55.
- Allport, G. W. (1937). *Personality: A psychological interpretation*. New York, NY: Henry Holt.
- Bak, M., Drukker, M., Hasmi, L., & van Os, J. (2016). An n= 1 Clinical network analysis of symptoms and treatment in psychosis. *PloS one*, 11(9), e0162811.
- Baumert, A., Schmitt, M., Perugini, M., Johnson, W., Blum, G., Borkenau, P., ... & Jayawickreme, E. (2017). Integrating personality structure, personality process, and personality development. *European Journal of Personality*, 31(5), 503-528.
- Beltz, A. M., Wright, A.G.C., Sprague, B., & Molenaar, P.C.M. (2016). Bridging the nomothetic and idiographic approaches to the analysis of clinical data. *Assessment*, 23(4), 447-458.
- Beltz, A. M., & Gates, K. M. (2017). Network Mapping with GIMME. *Multivariate behavioral research*, 52(6), 789-804.
- Bentler, P. M., & Chou, C. P. (1987). Practical issues in structural modeling. *Sociological Methods & Research*, 16(1), 78-117.
- 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(2), 202-221.
- Bos, F. M., Snippe, E., de Vos, S., Hartmann, J. A., Simons, C. J., van der Krieke, L., ... & Wichers, M. (2017). Can We Jump from Cross-Sectional to Dynamic Interpretations of Networks? Implications for the Network Perspective in Psychiatry. *Psychotherapy and Psychosomatics*, 86(3), 175-177.
- Bringmann, L. F., Pe, M. L., Vissers, N., Ceulemans, E., Borsboom, D., Vanpaemel, W., ... &

- Kuppens, P. (2016). Assessing temporal emotion dynamics using networks. *Assessment*, 23(4), 425-435.
- Bulteel, K., Mestdagh, M., Tuerlinckx, F., & Ceulemans, E. (2017). VAR (1) Based Models do not Outpredict AR (1) Models in Current Psychological Applications. *Manuscript in Preparation*.
- Carpenter, R. W., Wycoff, A. M., & Trull, T. J. (2016). Ambulatory assessment: New adventures in characterizing dynamic processes. *Assessment*, 23(4), 414-424.
- Costa, P. T., Jr., & McCrae, R. R. (2008). The Revised NEO Personality Inventory (NEO-PI-R). In G. Boyle, G. Matthews, & D. Saklofske (Eds.), *Sage handbook of personality theory and assessment* (Vol. 2, pp. 179–198). Los Angeles, CA: Sage Publications.
- Cattell, R. B., Cattell, A. K. S., & Rhymer, R. M. (1947). P-technique demonstrated in determining psychophysiological source traits in a normal individual. *Psychometrika*, 12(4), 267–288.
- Conner, T. S., Tennen, H., Fleeson, W., & Barrett, L. F. (2009). Experience sampling methods: A modern idiographic approach to personality research. *Social and personality psychology compass*, 3(3), 292-313.
- DeYoung, C. (2015). Cybernetic Big Five Theory. *Journal of Research in Personality*, 56, 33-58.
- Ellis, C.A., & Parbery, S.A. (2005). Is smarter better? A comparison of adaptive, and simple moving average trading strategies. *Research in International Business and Finance*, 19, 399-411.
- Fisher, A.J. (2015). Toward a dynamic model of psychological assessment: Implications for personalized care. *Journal of consulting and clinical psychology*, 83(4), 825-836.
- Fisher, A.J., Reeves, J.W., Lawyer, G., Medaglia, J.D., & Rubel, J.A. (2017). Exploring the idiographic dynamics of mood and anxiety via network analysis. *Journal of abnormal*

psychology, 126(8), 1044-1056.

Fleeson, W., & Jayawickreme, E. (2015). Whole Trait Theory. *Journal of Research in Personality*, 56, 82–92.

Foster, K.T., & Beltz, A.M. (in press). Advancing statistical analysis of ambulatory assessment data in the study of addictive behavior: A primer on three person-oriented techniques. *Addictive Behaviors*.

Fried, E.I., & Cramer, A.O. (2017). Moving forward: challenges and directions for psychopathological network theory and methodology. *Perspectives on Psychological Science*, 12(6), 999-1020.

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(2), 129–148.

Gates, K.M., & Molenaar, P. C. M. (2012). Group search algorithm recovers effective connectivity maps for individuals in homogeneous and heterogeneous samples. *NeuroImage*, 63(1), 310–319.

Gates, K.M., Molenaar, P.C.M., 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.

Goldberg, L. R. (1990). An alternative description of personality: the big-five factor structure. *Journal of Personality and Social Psychology*, 59(6), 1216–29.

Granger, C. W. J. (1969). Investigating Causal Relations by Econometric Models and Cross-

- spectral Methods. *Econometrica*, 37(3), 424–438.
- Hallquist, M.N. & Wright, A.G.C. (2014). Mixture modeling methods for the assessment of normal and abnormal personality part I: Cross-sectional models. *Journal of Personality Assessment*, 96(3), 256-268.
- 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(1), 10-15.
- Hamburg, M. A., & Collins, F. S. (2010). The path to personalized medicine. *New England Journal of Medicine*, 363, 301-304.
- Henry, T., & Gates, K. (2017). Causal search procedures for fMRI: review and suggestions. *Behaviormetrika*, 44(1), 193-225.
- 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.
- Koopman, S. J., & Ooms, M. (2003). Time-Series Modelling of Daily Tax Revenues. *Statistica Neerlandica*, 57(4), 439–469.
- Kotov, R., Krueger, R.F., Watson, D., Achenbach, T.M., Althoff, R.R., ... & Zimmerman, M. (2017). The Hierarchical Taxonomy of Psychopathology (HiTOP): A dimensional alternative to traditional nosologies. *Journal of Abnormal Psychology*, 126(4), 454-477.
- Lane, S., Beltz, A., Gates, K., & Wright, A.G.C. (in press). Uncovering general, shared, and unique temporal patterns in ambulatory assessment data. *Psychological Methods*.
- 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*, 1-15.

- 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(2), 205–224.
- Litten, R. Z., Ryan, M. L., Falk, D. E., Reilly, M., Fertig, J. B., & Koob, G. F. (2015). Heterogeneity of alcohol use disorder: understanding mechanisms to advance personalized treatment. *Alcoholism: Clinical and Experimental Research*, 39(4), 579-584.
- Liu, S. (2017). Person-specific versus multilevel autoregressive models: Accuracy in parameter estimates at the population and individual levels. *British Journal of Mathematical and Statistical Psychology*.
- MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods*, 4(1), 84-99.
- Molenaar, P. C. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement*, 2(4), 201-218.
- Molenaar, P. C., & Campbell, C. G. (2009). The new person-specific paradigm in psychology. *Current directions in psychological science*, 18(2), 112-117.
- Moskowitz, D. S., Russell, J. J., Sadikaj, G., & Sutton, R. (2009). Measuring people intensively. *Canadian Psychology/Psychologie Canadienne*, 50(3), 131–140.
- Mumford, J. A., & Ramsey, J. D. (2014). Bayesian networks for fMRI: A primer. *NeuroImage*, 86, 573–582.
- Pincus, A.L., Hopwood, C.J., & Wright, A.G.C. (in press). The interpersonal situation: An integrative framework for the study of personality, psychopathology, and psychotherapy. In D. Funder, J.F. Rauthmann, & R. Sherman (Eds), *Oxford handbook of Psychological Situations*.

Pons, P., & Latapy, M. (2005). Computing communities in large networks using random walks.

Journal of Graph Algorithms and Applications, 10(2), 191-218.

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(4), 347-357.

Ram, N., & Nesselroade, J. R. (2007). Modeling intraindividual and intracontextual change:

Rendering developmental contextualism operational. Modeling contextual effects in longitudinal studies, 325-342.

Rauthmann, J. F. F., Gallardo-Pujol, D., Guillaume, E. M. M., Todd, E., Nave, C. S. S., Sherman,

R. A. A., ... Funder, D. C. C. (2014). The Situational Eight DIAMONDS: A taxonomy of major dimensions of situation characteristics. *Journal of Personality and Social Psychology*, 107(4), 677-718.

Rauthmann, J. F., Jones, A. B., & Sherman, R. A. (2016). Directionality of Person-Situation

Transactions: Are There Spillovers Among and Between Situation Experiences And Personality States?. *Personality and Social Psychology Bulletin*, 42(7), 893-909.

Roche, M. J., Pincus, A. L., Rebar, A. L., Conroy, D. E., & Ram, N. (2014). Enriching

Psychological Assessment Using a Person-Specific Analysis of Interpersonal Processes in Daily Life. *Assessment*, 21(5), 515-528.

Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological momentary assessment. *Annu.*

Rev. Clin. Psychol., 4, 1-32.

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(2), 875-891.

Spencer, J.P., & Schöner, G. (2003). Bridging the representational gap in the dynamic systems

- approach to development. *Developmental Science*, 6, 392-412.
- Thompson, E. R. (2007). Development and validation of an internationally reliable short-form of the positive and negative affect schedule (PANAS). *Journal of Cross-Cultural Psychology*, 38(2), 227-242.
- Trull, T. J., & Ebner-Priemer, U. (2013). Ambulatory assessment. *Annual review of clinical psychology*, 9, 151-176.
- van der Heijden, G. J., Donders, A. R. T., Stijnen, T., & Moons, K. G. (2006). Imputation of missing values is superior to complete case analysis and the missing-indicator method in multivariable diagnostic research: a clinical example. *Journal of clinical epidemiology*, 59(10), 1102-1109.
- van de Leemput, I.A., Wichers, M., Cramer, A.O., Borsboom, D., Tuerlinckx, F., Kuppens, P., ... & Derom, C. (2014). Critical slowing down as early warning for the onset and termination of depression. *Proceedings of the National Academy of Sciences*, 111, 87-92.
- van Os, J., Delespaul, P., Wigman, J., Myin-Germeys, I., & Wichers, M. (2013). Beyond DSM and ICD: introducing “precision diagnosis” for psychiatry using momentary assessment technology. *World Psychiatry*, 12(2), 113–117.
- Whitley, D. C., Ford, M. G., & Livingstone, D. J. (2000). Unsupervised forward selection: a method for eliminating redundant variables. *Journal of Chemical Information and Computer Sciences*, 40(5), 1160-1168.
- Wichers, M., Groot, P. C., Psychosystems, E. S. M., & EWS Group. (2016). Critical slowing down as a personalized early warning signal for depression. *Psychotherapy and psychosomatics*, 85(2), 114-116.
- Wichers, M., Simons, C.J.P., Kramer, I.M.A., Hartmann, J.A., Lothmann, C., Myin-Germeys, I., . . . van Os, J. (2011). Momentary assessment technology as a tool to help patients with

- depression help themselves. *Acta Psychiatrica Scandinavica*, 124(4), 262–272.
- Wiggins, J. S. (1995). *The Interpersonal Adjective Scales (IAS) manual*. Odessa, FL: Psychological Assessment Resources.
- Wright, A.G.C. (2011). Quantitative and qualitative distinctions in personality disorder. *Journal of Personality Assessment*, 93(4), 370-379.
- Wright, A.G.C. (2017). Factor analytic support for the five-factor model. In T.A. Widiger (Ed.) *Oxford handbook of the five-factor model* (pp. 217-242). Oxford, UK: Oxford University Press.
- 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.
- Wright, A.G.C., Hallquist, M.N., Stepp, S.D., Scott, L.N., Beeney, J.E., Lazarus, S.A., & Pilkonis, P.A. (2016). Modeling heterogeneity in momentary interpersonal and affective dynamic processes in borderline personality disorder. *Assessment*, 23(4), 484-495.
- Wright, A.G.C. & Hopwood, C.J. (2016). Advancing the assessment of dynamic psychological processes. *Assessment*, 23(4), 399-403.
- Wright, A.G.C., Hopwood, C.J., & Simms, L.J. (2015). Daily interpersonal and affective dynamics in personality disorder. *Journal of Personality Disorders*, 29(4), 503-525.
- Wright, A.G.C. & Simms, L.J. (2015). A metastructural model of mental disorders and pathological personality traits. *Psychological Medicine*, 45(11), 2309-2319.
- Wright, A.G.C., & Simms, L.J. (2016). Stability and fluctuation of personality disorder features in daily life. *Journal of Abnormal Psychology*. 125(5), 641-656.
- Zevon, M. A., & Tellegen, A. (1982). The structure of mood change: An idiographic/nomothetic analysis. *Journal of Personality and Social Psychology*, 43(1), 111–122.

Table 1: Number of individual models with each path present and average strength and standard deviation of present.

	Lagged Effects (<i>N/M/SD</i>)					
	Dominance	Affiliation	Positive Affect	Negative Affect	Stress	Functioning
Dominance	94/.12/.17	5/.04/.36	1/.41/-	4/.02/.47	4/.01/.31	2/-.23/.02
Affiliation	6/.11/.26	94/.18/.15	1/35/-	4/-.03/.39	5/-.09/.34	5/.01/.35
Positive Affect	2/-.01/.41	6/.18/.38	94/.24/.20	6/.11/.40	3/.07/.56	6/.01/.28
Negative Affect	3/-.25/.03	1/.19/-	2/.02/.35	94/.20/.17	9/-.18/.27	3/.26/.06
Stress	3/.83/1.96	3/-.01/.47	4/.37/.18	10/.31/.30	94/.14/.21	4/-.31/.06
Functioning	5/.09/.30	4/-.04/.25	3/-.08/.29	3/-.10	3/.30/.02	94/.19/.19

	Contemporaneous Effects					
	Dominance	Affiliation	Positive Affect	Negative Affect	Stress	Functioning
Dominance	0	16/-.07/.49	31/.40/.20	10/.16/.30	14/.33/.21	7/-.12/.36
Affiliation	6/.35/.10	0	36/.35/.30	70/-.45/.31	5/-.27/.04	3/.44/1.34
Positive Affect	22/.48/.11	20/.45/.13	0	13/.30/.25	10/.14/.26	13/-.40/.25
Negative Affect	3/.03/.38	9/.08/.51	10/.22/.25	0	94/.50/.26	7/.07/.41
Stress	12/.35/.11	6/-.14/.60	7/.21/.29	4/-1.04/3.09	0	9/.20/.50
Functioning	4/-.33/.09	16/-.39/.28	23/-.44/.21	18/.46/.22	13/.40/.20	0

Note: Cells reflect the column variable predicting the row variable.

Table 2. Means and standard deviations of group-level beta coefficients by subgroup.

	Subgroup 1		Subgroup 2		Subgroup 3	
	Mean	SD	Mean	SD	Mean	SD
Dominance (AR)	.14 ^a	.15	.10 ^a	.21	.06 ^a	.18
Stress (AR)	.10^a	.15	.31^b	.23	.19 ^{ab}	.28
PA (AR)	.20^a	.18	.37^b	.26	.29 ^{ab}	.17
NA (AR)	.19 ^a	.15	.30 ^a	.20	.18 ^a	.17
Functioning (AR)	.16 ^a	.17	.28 ^a	.27	.23 ^a	.20
Affiliation (AR)	.13^a	.13	.19 ^{ab}	.16	.29	.16
STR→NA	.55^a	.22	.12 ^b	.16	.52^a	.23

Note. Subscripts that differ indicate significant differences at $p < .05$; significant differences bolded.

Figure 1. Diagram of all general (group), shared (subgroup), and person-specific (individual) paths estimated in group iterative multiple model estimation. Group level paths are in black, subgroup paths are in green, and individual paths are in gray. Line thickness indicates the number of individuals with the path. Contemporaneous effects are represented with solid lines, and lagged effects are represented with dashed lines.

Figure 2. Diagrams of group iterative multiple model estimation results by subgroup (excluding two ungrouped individuals). Group level paths are in black, subgroup paths are in green, and individual paths are in gray. Line thickness indicates the number of individuals with the path proportionate to the total group size. Contemporaneous effects are represented with solid lines, and lagged effects are represented with dashed lines.

Figure 3. Example diagrams of person-specific (i.e., individual-level) models from group iterative multiple model estimation. Contemporaneous effects are represented with solid lines, and lagged effects are represented with dashed lines. Positive effects are in red, negative effects are in blue. Line thickness denotes strength of effect.