American Psychological Association

Psychological Assessment

© 2019 American Psychological Association
1040-3590/19/$12.00

2019, Vol. 31, No. 4, 432–443
http://dx.doi.org/10.1037/pas0000562

Testing the Structure and Process of Personality Using Ambulatory Assessment Data: An Overview of Within-Person and Person-Specific Techniques

Joeri Hofmans
Vrije Universiteit Brussel

Barbara De Clercq
Ghent University

Peter Kuppens
KU Leuven

Lize Verbeke
Ghent University

Thomas A. Widiger
University of Kentucky

In the present article, we discuss the potential of ambulatory assessment for an idiographic study of the structure and process of personality. To this end, we first review important methodological issues related to the design and implementation of an ambulatory assessment study in the personality domain, including methods of ambulatory assessment, frequency of measurement and duration of the study, ambulatory assessment scales and questionnaires, participant selection, training and motivation, and ambulatory assessment hard- and software. Next, we provide a detailed outline of available analytical approaches that can be used to analyze the intensive longitudinal data generated by an ambulatory assessment study. By doing this, we hope to familiarize personality scholars with these methods and to provide guidance for their use in the field of personality psychology and beyond.

Public Significance Statement
We show how ambulatory assessment can be used for an idiographic study of the structure and process of personality.

Keywords: ambulatory assessment, personality, idiographic

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Traditionally, personality psychology has adopted the nomothetic perspective on science. According to this perspective, the goal of personality science is to "make general predictions about the population by examining between-person variation" (Beltz, Wright, Sprague, & Molenaar, 2016, p. 447). In line with this aspiration, personality researchers have strongly focused on studying between-person differences in behavior, affect, and cognition, aiming to identify a limited number of traits allowing for a comprehensive description of personality. From this perspective, each individual is positioned within a basic set of universal personality dimensions, with individual uniqueness resulting from the distinct combination of positions on these basic traits. One of the best-known models embedded within the nomothetic approach to personality is the Five-Factor Model (FFM; Costa & McCrae, 2017), characterizing people by means of their standing on extraversion, openness, agreeableness, conscientiousness, and emotional stability. This five-factor personality structure is reflected at the genetic level (Jarnecke & South, 2017), has been replicated across age (Soto & Tackett, 2015) and culture (Allik & Realo, 2017), and has proven to be relatively stable across time (Roberts & DelVecchio, 2000). A wealth of research in support of the FFM has resulted in a widespread agreement among personality psychologists that these five traits indeed represent the basic and universal structure of between-person differences in personality (Widiger, 2017).

Editor's Note. Yossef S. Ben-Porath, PhD, served as the sole action editor for this submission.

This article was published Online First March 14, 2019.

Joeri Hofmans, Research group of Work and Organizational Psychology, Vrije Universiteit Brussel; Barbara De Clercq, Department of Developmental, Personality and Social Psychology, Ghent University; Peter Kuppens, Research group of Quantitative Psychology and Individual Differences, KU Leuven; Lize Verbeke, Department of Developmental, Personality and Social Psychology, Ghent University; Thomas A. Widiger, Department of Psychology, University of Kentucky.

This research was supported by a Fonds Wetenschappelijk Onderzoek - Vlaanderen (FWO) grant G024615N.

Correspondence concerning this article should be addressed to Joeri Hofmans, Research group of Work and Organizational Psychology, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussel. E-mail: joeri.hofmans@vub.be

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

However, despite major achievements of the nomothetic approach, there is a long-standing awareness that personality also reflects complex dynamic, intraindividual processes that are manifested over time in response to and in interaction with the individual's environment (Fleeson & Noftle, 2012). The acknowledgment of the dynamic nature of personality can conceptually be traced back to Allport (1937, p. 48) who defined personality as "the dynamic organization within the individual of those psychophysical systems that determine his unique adjustments to his environment." Since this original seminal text, substantial empirical evidence has confirmed that the course of personality represents a mixture of stability and variability. When averaged across multiple occasions, individuals show substantial stability in their average level of behavior, affect, and cognition across time. When considering moment-to-moment fluctuations, people's behaviors, feelings, and cognitions show considerable variability. In support of the stability of personality, research shows that the Big Five scores of a single individual are stable from one week to another when averaging across occasions, yielding correlation coefficients of about .80 (Baird, Le, & Lucas, 2006; Fleeson, 2001). At the same time, however, research has also supported the importance of within-person variation in personality by demonstrating that the differences in behavior, feelings, and cognitions within an individual across situations are about as large as the differences between individuals (e.g., Fleeson, 2001). These findings imply that both between-person as well as within-person fluctuations in behavior, affect, and cognition should be equally considered to obtain a full understanding of personality.

Influenced by rapid technological advancements during the last decades, a detailed assessment of personality indicators as they evolve in naturalistic settings of daily life has become feasible. As argued by Trull and Ebner-Priemer (2013), the methods used for collecting such data can be subsumed under the umbrella of ambulatory assessment, including a variety of methods such as experience sampling methodology, daily diary research, observational research, and research using (physiological) sensors (Trull & Ebner-Priemer, 2014). In a typical ambulatory assessment study, participants' behaviors, feelings, and cognitions (along with relevant situational features) are recorded repeatedly during the routine activity of everyday life, after which the researcher looks at within-person fluctuations in the constructs of interest. The merits of ambulatory assessment—as opposed to retrospective cross-sectional survey methods—have been clearly described for clinical research and intervention (Trull & Ebner-Priemer, 2009, 2013, 2014), and have also been outlined for industrial/organizational settings (Beal & Weiss, 2003; Fisher & To, 2012). In the current article, we aim to contribute to both research areas by detailing how ambulatory assessment can be used to study both structural and process-based aspects of personality. To this end, we first discuss a number of important methodological issues related to the design and implementation of an ambulatory assessment study in the personality domain, followed by a detailed outline of available approaches to analyze the intensive longitudinal data generated by ambulatory assessment. We hope to familiarize personality scholars with the collection and analysis of ambulatory assessment data and to provide guidance for their use in the field of personality psychology and beyond.

We first discuss the contributions and pitfalls of using a strict nomothetic approach to studying the structure and process of personality. Next, we demonstrate how ambulatory assessment can be used to examine within-person and person-specific structures and processes, first focusing on important methodological issues, after which we review a series of analytical techniques that can be used for analyzing ambulatory assessment data.[1]

Describing Structural and Process-Based Aspects of Personality: Contributions and Pitfalls of a Strict Nomothetic Approach

Two of the central goals of personality psychology are (a) describing the structure of personality—or describing how the different components of one's personality covary—and (b) understanding the process of personality—or understanding how the components of one's personality influence each other and are influenced by external factors (Conner, Tennen, Fleeson, & Barrett, 2009). Up until now, both goals have predominantly been studied using a nomothetic, between-person perspective on personality.

In particular, studies on the *structure of personality* typically collect data from a (preferably large) sample of individuals. Factor analysis or principal component analysis is then applied to extract common dimensions that explain the majority of the phenotypic variation across individuals. The issue with this approach is that such large-scale, between-person studies tell us how individual differences are structured in the population of individuals, but not how these behaviors, feelings, and cognitions are organized within the individual (Fleeson & Noftle, 2012). Indeed, structures that apply at the between-person level cannot readily be transferred to the within-person level (Molenaar, 2004). This limitation was recognized a long time ago by Allport (1937), who stated that nomothetic methods run the risk of finding structures and processes that fail to apply to any single individual. This issue becomes even more complicated when we consider that there may be between-person differences in person-specific personality structures. For example, research by Borkenau and Ostendorf (1998) showed that, while the factor structure of longitudinal correlations averaged across participants showed a good match with the structure of individual differences, the match with the factor structure of individual participants was substantially worse. In a similar vein, Hamaker, Dolan, and Molenaar (2005), analyzing the same data, showed that for some participants each of the five dimensions of the FFM meaningfully contributed to the description of their personality structure, while other participants' behaviors, feelings, and cognitions are organized according to only two or three broad personality dimensions. Consequently, because between-person structures do not readily translate into person-specific structures, the nomothetic taxonomic work on between-person differences in behaviors, feelings, and cognitions cannot straightforwardly be used when it comes to explaining psychological functioning at the individual level.

Apart from examining how the different components of an individual's personality relate to each other (i.e., studying the structure of personality), personality research has also paid considerable attention to studying the directional relationships between those components (i.e., the *process of personality*). Paral-

[1] Note that, because our article does not include empirical data, no research ethics committee approval was applied for.

leling research on the structure of personality, the process of personality has also predominantly been studied using cross-sectional, between-person designs. The problem with this approach is that between-person analyses relate average levels of the predictor to average levels of the outcome, and these average levels "include the sediment of many different processes that have operated over a long period of time" (Fleeson & Noftle, 2012, p. 533). Moreover, as Molenaar (2004) demonstrated, relationships at the between-person level do not necessarily hold at the within-person level (see also Hamaker, 2012). In fact, results obtained from the analysis of between-person data can only be used to explain within-person fluctuations when two very stringent conditions are met: (a) the within-person process needs to be stationary, which means that the statistical characteristics of the process (i.e., mean, variance, and covariances) should be invariant over time, and (b) the process should be homogeneous across subjects, meaning that an identical statistical model should hold for each individual in the population (Gayles & Molenaar, 2013; Molenaar, 2004). As one can readily see, these conditions do not hold for personality. In particular, research on personality development in young (De Clercq, Verbeke, De Caluwé, Vercruysse, & Hofmans, 2017) and adult life (Wille, Hofmans, Feys, & De Fruyt, 2014) demonstrates that there is systematic development in several personality traits over the life span, which in itself implies that personality is nonstationary. With regard to homogeneity, Borkenau and Ostendorf (1998) and Hamaker et al. (2005) demonstrated that Big Five trait adjectives cluster in different numbers of dimensions for different individuals, showing that people differ from each other in the way their behaviors, feelings, and cognitions covary. This lack of stationarity and homogeneity implies that nomothetic methods cannot be used to study person-specific structures and processes. In other words, whereas nomothetic methods are useful when it comes to describing and explaining between-person differences in behavior, affect, and cognition, different methods are needed to study the dynamic structures and processes underling personality.

Idiographic methods are key to this issue because their goal is not to identify those patterns of behaviors, cognitions, and affects that describe the average individual; instead, idiographic methods aim at studying and describing these patterns within one individual across experiences or situations. That is, idiographic approaches do not aim to examine whether certain behaviors, cognitions, and emotions co-occur across individuals, but rather attempt to test whether these behaviors, cognitions, and emotions co-occur in time within one and the same individual. This is important from a theoretical point of view as the aim of the idiographic approach closely aligns with the central goals of personality psychology, namely, examining how different components of an individual's personality influence and relate to one another. This idiographic approach to personality is also consistent with clinical interests, as interventions result in changes within an individual. For of these reasons, Conner et al. (2009) argue that idiographic methods are at the core of personality psychology.

Despite its theoretical and practical appeal, however, the popularity of idiographic methods has been limited, which can be understood from the low statistical power and limited generalizability of traditional idiographic methods. Indeed, idiographic studies relied for a long time on a $N = 1$ approach, yielding data that were well-suited for clinicians or biographers but did not allow hypothesis testing because of the limited number of observations. Moreover, because of the idiographic sample size of one, the findings of early idiographic studies hampered generalization beyond the specific individual under investigation. As a result, idiographic methods have not been considered to fit well with the objectives of psychological sciences.

The Role of Ambulatory Assessment in the Revival of the Ideographic Approach

Recent technological and analytical developments, and specifically the emergence of ambulatory assessment, however, have made a true contemporary idiographic approach to psychological constructs feasible. Ambulatory assessment is specifically designed to collect multiple observations per individual, and the resulting intensive longitudinal data can then be used for hypothesis testing at the level of the single individual. This means that ambulatory assessment has the potential to yield data from the same individual with enough statistical power to perform quantitative hypothesis testing. Moreover, because it also allows collecting data for multiple individuals at the same time, ambulatory assessment creates opportunities for making inferences beyond a single individual. That is, ambulatory assessment allows testing whether the structure and relationships generalize beyond one single individual to a larger population of individuals. As a result, ambulatory assessment rectifies the two main stumbling blocks for the adoption of idiographic methods (i.e., low statistical power and limited generalizability).

Still, compared with other research domains such as the clinical field or the field of emotion research, the assessment of within-person variability has been less widely adopted in the personality domain. Whereas various ambulatory assessment strategies have been regularly applied in research concerning (for instance) mood disorders, anxiety, emotion regulation, alcohol use, or psychotic experiences (for a seminal review see Trull & Ebner-Priemer, 2013), the interest in collecting such data in the personality field has only recently been growing (Wright, Hopwood, & Simms, 2015). The reason for the asynchrony between the clinical/emotion and the personality field may be that for a long time personality was historically believed to be trait-based and, thus, relatively stable (Costa & McCrae, 1994), whereas clinical constructs or emotions were considered to reflect states and, therefore, these constructs qualified more as subject of methodologies aimed at capturing momentary ratings.

Recently, however, studies on dynamic personality processes have evolved within the area of borderline personality disorder research (e.g., Sadikaj, Moskowitz, Russell, Zuroff, & Paris, 2013; Trull et al., 2008), a disorder that is characterized by strong instability. We expect more dynamic within-person research to follow as it recently has been shown that these dynamic processes are likely to be found across all personality disorders (Wright, Beltz, Gates, Molenaar, & Simms, 2015; Wright & Simms, 2016). From a general trait perspective, there have been several ambulatory assessment studies on more adaptive trait tendencies, such as extraversion (Fleeson, Malanos, & Achille, 2002), conscientiousness (Debusscher, Hofmans, & De Fruyt, 2017; Minbashian, Wood, & Beckmann, 2010), emotional stability (Debusscher, Hofmans, & De Fruyt, 2014), and higher-order personality constructs such as core-self evaluations (Debusscher, Hofmans, & De Fruyt,

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

2016a; Hofmans, Debusscher, Dóci, Spanouli, & De Fruyt, 2015). These studies demonstrate that there is also significant within-person variability within the adaptive personality dimensions, and that this within-variability can be predicted from situational triggers while themselves being predictive of important work and life outcomes. Overall, the results of these initial studies confirmed that there is substantial variability in both adaptive and maladaptive personality characteristics, highlighting the need for more research on within-person variability in both adaptive and maladaptive personality dimensions.

Conducting an Ambulatory Assessment Study to Examine Within-Person and Person-Specific Structures and Processes: Methodological Issues Related to its Design and Implementation

Before discussing how one can address two of the central goals of personality psychology with ambulatory assessment data, we will first turn to a number of important methodological issues related to the design and implementation of an ambulatory assessment study in the personality domain. Note that the goal of this section is not to give an exhaustive overview of design-related issues—these issues are at length discussed elsewhere (see the articles by Bolger, Davis, & Rafaeli, 2003; Fisher & To, 2012; and Trull & Ebner-Priemer, 2013)—but rather to bring attention to issues that are critical when designing a study with the specific aim to examine within- and/or person-specific personality structure and/or process. In particular, five issues will be addressed: (a) methods of ambulatory assessment, (b) frequency of measurement and duration of the study, (c) ambulatory assessment scales and questionnaires, (d) participant selection, training and motivation, and (e) ambulatory assessment hard- and software.

Methods of Ambulatory Assessment

Ambulatory assessment methods can broadly be classified within three categories: self-report ambulatory assessment, observational ambulatory assessment, and physiological/biological/behavioral ambulatory assessment (see Trull & Ebner-Priemer, 2013). For self-report ambulatory assessment, respondents are required to respond to questions at prescribed times (interval-contingent sampling), at random times (signal-contingent reporting), or conditional upon a discrete event (event-contingent reporting). Each of these sampling schemes has unique advantages and disadvantages (for an overview, see Fisher & To, 2012; Reis, Gable, & Maniaci, 2014). The major advantage of interval-contingent reporting is that it is easy to implement and relatively unobtrusive for respondents because of the predictable timing of the reports. The downside is that this sampling scheme may disproportionally capture experiences that happen at those specific times of the day. For signal-contingent reporting, the major advantage is obviously that one can obtain a representative sample of experiences, while the downside is that this sampling scheme is somewhat burdensome for participants because of the unpredictable timing of signals. Finally, the advantage of event-contingent reporting is that it allows capturing rare events that are of interest to the researcher, while the disadvantage is that one must train participants to recognize these events and react to them, which may contribute to the reactivity on the part of the respondents (to see how reactivity can be detected in an ambulatory assessment study, see Barta, Tennen, & Litt, 2011).

Observational ambulatory assessment and physiological/biological/behavioral ambulatory assessment are less well-known and are different from self-report ambulatory assessment in the sense that they do not rely on self-reports (Trull & Ebner-Priemer, 2013). Examples of observational ambulatory assessment are the Electronically Activated Recoder (EAR; Mehl & Robbins, 2012)—a portable audio device capturing snippets of ambient sounds in the participant's environment—or the use of global positioning system (GPS) data, photos, or video camera data. Physiological/biological/ behavioral ambulatory assessment, in turn, include the measurement of cardiac activity, heart rate variability, or blood pressure. In the context of the examination of within-person and person-specific structures and processes, these methods can be useful because they are relatively unobtrusive, are less vulnerable to reactivity effects and because they provide objective measures. On the other hand, whereas it is clear that such methods can be used to measure (or complement the subjective measurement of) the behavioral component of personality, the cognitive and affective component are by definition subjective in nature, which makes them less useful to capture these components.

Frequency of Measurement and Duration of the Study

A first important consideration when deciding on the frequency of measurement and the duration of an ambulatory assessment study is the time frame in which the process is expected to occur. The study of directional, idiographic processes requires testing whether or not the predictor at time t predicts the outcome at time $t + 1$. By how many minutes, hours, or days t and $t + 1$ should be separated, is a question that should be answered on theoretical grounds. Note that with signal-contingent reporting (i.e., randomly sampling from one's population of experiences) as well as with event-contingent reporting (i.e., reporting upon the occurrence of a discrete event), the period in between t and $t + 1$ may vary within the individual as well as across individuals, and this variation in time lags may complicate the detection of time-lagged effects. One analytical solution to this issue is to add the length of the interval between t and $t + 1$ as a moderator to the model, thereby testing whether the length of the time lag impacts upon the relationship of interest (see Beal & Weiss, 2003 for a discussion of this approach). However, this obviously complicates the—already complicated—analyses, so it is wise to reflect on this issue beforehand.

Second, the frequency of measurement and the length of the study should also take into account the time frame in which the constructs of interest fluctuate (Reis et al., 2014). For example, when studying neuroticism, it makes sense to have frequent measurements throughout the day because of the high volatility of this particular personality dimension, whereas personality dimensions reflecting more discrete behaviors, such as for instance impulsivity or recklessness, can be adequately measured with one assessment per day as it is unlikely that these specific trait manifestations occur at the exact moment of assessment. In practice, current self-report ambulatory assessment studies on within-person fluctuations in adaptive personality have prompted participants between five times per day (e.g., Fleeson, 2001) and once every day (e.g., Debusscher, Hofmans, & De Fruyt, 2016b; Hofmans et al., 2015). Experience sampling studies on emotion or mood, instead, easily involve from 10 to even 50 assessments per day (e.g., Kuppens, Oravecz, & Tuerlinckx, 2010).

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

More important, the expected frequency of the trait manifestation is not only an issue within days but also across days. This is illustrated by one of the longest daily diary designs currently available (Wright & Simms, 2016), showing considerable differences between personality dimensions in terms of proportion of endorsement across the duration of the study, with dimensions such as hostility reflecting a much lower base rate than for example emotional lability. In other words, very much like frequency of measurement, duration should consider as well the base rate of the targeted behavior/personality dimension.

A final issue that should inform the frequency of measurement and the duration of an ambulatory assessment study is statistical power. When testing within-person structures or processes, statistical power depends—among other things—on the total number of observations, which is determined by the number of participants, the number of repeated measurements per participant and the number of missed signals. From this perspective, a high number of repeated measurements can compensate for low numbers of participants and vice versa. In turn, when one is interested in examining person-specific structures or processes, a high enough number of repeated measurements per individual should be ensured. Although no strict guidelines are available, previous studies that performed idiographic, person-specific analyses included about 75–100 repeated observations per participant (e.g., Wright et al., 2015, 2016). Finally, when the research question concerns a mixture of within- and between-person effects—that is essentially the case when exploring between-person differences in person-specific personality structures or processes—both the number of repeated measurements and the number of participants matter.

Scales and Questionnaires

As self-report ambulatory assessment studies ask participants to repeatedly report on their cognitions, feelings, and behaviors, participation can be burdensome. Keeping this in mind, researchers should try to keep the length of the survey to a strict minimum. Regarding survey length, Reis et al. (2014) suggest that studies requiring up to five signals per day should not exceed 5 min per questionnaire, while studies that require only one daily report should not exceed 15 min. To keep questionnaire length to a minimum, and because there are very few validated multi-item scales specifically developed for ambulatory assessment studies, it is common for researchers to shorten existing scales. A guideline that may help with the selection of items is to select those items with the highest factor loadings. The reasoning is that, statistically speaking, items with high factor loadings capture the core of the construct and can, therefore, be considered the best indicators of the underlying construct. It is important to note, however, that when these factor loadings are obtained from between-person studies, one needs to be willing to assume that items that are central for capturing between-person differences also constitute the core of the construct when the focus is on within-person fluctuations.

A related issue pertains to the number of items needed to capture this core. Whereas some authors have proposed that each construct should be measured by at least three items (Shrout & Lane, 2011), several self-report ambulatory assessment studies have used single-item measures (see Debusscher et al., 2014). Of course, decisions regarding the number of items in a measure should be guided by several considerations. First, one should make sure that all relevant facets of the construct are measured; a condition that is not met when measuring multidimensional constructs with a single item. A second issue with single-item measures is that their internal consistency reliability cannot be tested. Whereas this may seem problematic at first glance, one should keep in mind that high internal consistency is never a goal in itself; internal consistency matters because it serves validity (readers interested in the calculation of internal consistency for ambulatory assessment scales are referred to Nezlek (2017), who proposed an approach using multilevel regression analysis, and to Geldhof, Preacher, & Zyphur (2014), using multilevel confirmatory factor analysis). Therefore, when the single item measures a unidimensional, simple construct, shows face and content validity and correlates with other variables in the expected direction, the single-item measure should probably be considered acceptable (Fisher & To, 2012).

In summary, deciding on the number of items and selecting these items requires that the researcher is aware of the dimensionality and complexity of the construct to be measured. For personality researchers interested in repeatedly measuring one or more of the Big Five dimensions through self-reports, Saucier's (1994) mini markers, the Ten Item Personality Inventory (TIPI; Gosling, Rentfrow, & Swann, 2003) and the Big Five Inventory-10 (BFI-10; Rammstedt & John, 2007) are useful measures for use in ambulatory assessment studies, as they are short while at the same time capturing several facets of each of the Big Five personality dimensions. Important to note here is that, despite their potential for ambulatory assessment studies, these measures were never designed for this type of research. In fact, whereas in ambulatory assessment studies participants are instructed to report on their immediate experiences, virtually all existing personality measures are assessing for retrospective appearance of the traits over the course of one's life. Because of this reason, one often must revise the instructions and/or the items to make the scales suited for an ambulatory assessment format.

Participant Selection, Training, and Motivation

Another important issue that needs to be considered when designing an ambulatory assessment study is participant selection, training, and motivation (Fisher & To, 2012). As noted earlier, participation in an ambulatory assessment study is relatively burdensome. Indeed, because of the time and commitment required it might be difficult to find persons who are willing to participate. To minimize nonresponse and participant dropout, researchers need to invest in participant selection, training, and motivation. Regarding selection, it is important to provide a realistic study preview. This means that participants should know before enrolling in the study how frequently they will be required to respond to signals to occur, when exactly they can expect signals, and how long the surveys will take to complete. In terms of training, participants need to understand that it is important to respond to as many signals as possible. Moreover, they need to know what to do when something unexpected happens (such as equipment malfunctions or inability to respond to the questions because of illness). Also, the researcher needs to make sure that all participants understand the questions being asked. Finally, researchers can boost participant responsiveness by giving incentives or rewards. One way to do this is by adequately compensating persons for their participation; either

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

with a lottery, a fixed amount per participant or a payment per response. As a nonmonetary incentive, the researcher may offer the possibility of personalized feedback at the end of the study, which may increase the motivation to partake in the study. Overall, it is important that researchers develop a warm and friendly relationship with the participants, viewing them, for instance, as partners in the study.

Hard- and Software

Because participants in ambulatory assessment studies are required to report on their behaviors, feelings, and cognitions in situ, the technology used should allow for the collection of responses as participants go through everyday life. Hence, smartphones, with their high level of portability (people carry them everywhere), familiarity (almost everyone has one), high ease of use (allowing display of graphical and textual information), and wireless connectivity (allowing synchronization of the data on the phone with a server) became increasingly popular. Moreover, smartphones are equipped with a range of sensors (e.g., GPS, photo and video cameras, and microphones) that can be used to collect observational or physiological/biological/behavioral data. Because the use of smartphones is widespread, one can use the participants' own phones on which an app can be installed. In recent years, several free and paying apps have been developed. Readers who are interested in an overview of such apps, can consult the chapter of Kubiak and Krog (2012), the overview provided by Conner (2015) or the website of the Society for Ambulatory Assessment (http://www.saa2009.org).

Describing Structural and Process-Based Aspects of Personality: Ambulatory Assessment as a Revival of the Idiographic Approach

It is clear that ambulatory assessment, with its focus on collecting repeated measurements of individuals in their day-to-day lives, has the potential to generate valuable information concerning both the structural as well as the process-based side of personality. In what follows, we will discuss at length a range of available analytical approaches that can be used to address two of the central goals of personality psychology: studying the structure of personality and studying the process of personality.

Describing Structural Aspects of Personality

Regarding the *structure of personality*, an ambulatory assessment study can tell us how the different components of personality are correlated within an individual across different occasions. In what follows, we discuss three ways to study the structure of personality using ambulatory assessment data.

These different approaches can all be placed on a continuum ranging from nonidiographic methods to idiographic methods. All these methods aim at analyzing how the dynamic patterning of behaviors, feelings, and cognitions across time is structured within individuals. However, they do so in a different way. Nonidiographic methods look for the structure of temporal covariation *across* subjects. In other words, the goal of dynamic, nonidiographic methods is to reveal the within-person structure that holds across all individuals in the sample. Because of this reason, Wright et al. (2015) argue that these methods capture the "within-person structure." Idiographic methods, instead, do not consider the temporal covariation pooled across participants, but consider instead each subject's multivariate time-series separately, thereby yielding person-specific models of temporal covariation. Hence, idiographic methods are said to concern the "person-specific structure" (Wright et al., 2015). In what follows, we will discuss three data-analytical methods that differ in their position on the nonidiographic-idiographic continuum. For each method we will discuss advantages and disadvantages.

Multilevel factor analysis. The first option for studying the structure of personality is multilevel factor analysis. Multilevel factor analysis is a nonidiographic method, which means that it can be used to study within-person structures. Generally, there are two ways to perform a multilevel factor analysis, and this holds for both multilevel exploratory factor analysis and multilevel confirmatory factor analysis. The first way is to manually or explicitly decompose the total covariance matrix into two orthogonal covariance or correlation matrices: (a) the between-person matrix and (b) the within-person matrix. With such a manual or explicit decomposition, the between-person matrix contains the covariances or correlations between the person-specific averages. Hence, this matrix reflects whether between-person fluctuations in one variable covary with between-person fluctuations in the other variables. The within-person matrix, in turn, captures the relationships among variables within individuals across time, indicating whether deviations from the person-specific average on one variable covary with deviations from one's average on the other variables. Because there is a different within-person matrix for each person in the dataset, these matrices are pooled across the different individuals in the sample to obtain a single within-matrix. Note that this is the reason why multilevel factor analysis yields a single within-person factor solution that holds across all individuals in the sample. After having decomposed the multilevel data into a between-person matrix and a within-person matrix, an exploratory or confirmatory factor model can be tested on these matrices separately, yielding a between-person factor solution and a within-person factor solution. The second way of testing a multilevel factor model is to directly test the model at both levels, which means that the decomposition of the multilevel data into a between-person and a within-person matrix is done in a latent or implicit way (Muthén, 1994). When this procedure is followed, only one likelihood function is maximized, which means that the within-person and the between-person model are fitted simultaneously to the multilevel data. Note that, because within-person variation is independent from between-person variation (Molenaar, 2004), the factor model does not need to be identical across the different levels. The major advantage of a latent or implicit decomposition of the multilevel data is that it elegantly handles data with a different number of observations per participant (i.e., unbalanced data). Indeed, when the number of observations per participant is equal, the manual/explicit decomposition and the latent/implicit composition are identical up to a scale factor. However, when the data are unbalanced, and particularly when the number of individuals is substantial relative to the number of repeated measurements, the manual/explicit decomposition results in a biased between-person matrix (Hox, 1993; Muthén, 1994). Therefore, the latent/implicit composition is generally superior. Researchers who are interested in multilevel factor analysis can

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

consult the article by Muthén (1994), which offers a stepwise procedure for testing a multilevel factor model. Applying this stepwise procedure, Reise, Ventura, Nuechterlein, and Kim (2005) analyzed ambulatory assessment data of 73 psychiatric patients reporting on 940 large and small negative life events. They showed that, whereas the number of within- and between-person factors was the same, there were substantial between-level differences in the level and pattern of factor loadings, implying that the latent factors had different meanings at the within-person and at the between-person level. Although the multilevel factor model is parsimonious in the sense that it tests a single within-person model rather than a separate model per respondent, it assumes that all individuals are drawn from a single population. As such, the multilevel factor model allows testing within-person, but not person-specific factor structures.

p-technique factor analysis. The second option—p-technique factor analysis (Cattell, 1963)—is an idiographic method, implying that this method can be used to study person-specific personality structures. In contrast to traditional r-technique factor analysis, which is performed on a Multiperson × Multivariables (× Single Occasion) data matrix, p-technique factor analysis is performed on a (Single Person ×) Multivariables × Multioccasions data matrix. Such a bottom-up approach, in which the factor model is tested on the repeated measures data of each participant separately, allows for true idiosyncrasy in the associations among the variables. Indeed, the goal of p-technique factor analysis is not to provide a parsimonious description of how the scores on a number of variables covary across individuals, but rather how these scores covary within one individual across occasions, which means that "the obtained structure can rightfully be interpreted at the level of the individual" (Brose & Ram, 2012, p. 460). In terms of implementation, p-technique factor analysis involves the same analytic procedures as r-technique factor analysis, implying that it can be done either in an exploratory or a confirmatory manner using standard statistical packages. Typically, once the factor structure has been determined for each of the individuals in the sample, the researcher looks for between-person similarities and differences in the within-person factor structures. This can be done by looking at the number of factors or the pattern of factor loadings (Hamaker, Nesselroade, & Molenaar, 2007). One way to do so is to rely on Tucker's congruency coefficients (see Zevon & Tellegen, 1982 and Wright et al., 2016 for examples). Tucker's congruency coefficient equals the cosine of the angle between two vectors of factor loadings, which means that it represents a standardized measure of proportionality of the factor loadings (Lorenzo-Seva & ten Berge, 2006). By comparing the proportionality of factor loadings between individuals, one can consider the extent to which the factor solutions of different individuals are alike.

In their article on the person-specific structure of borderline personality disorder, Wright and colleagues (2016) combine p-technique factor analysis with the use of Tucker's congruency coefficients. In particular, using event-contingent recording, they invited psychiatric patients to complete an electronic diary registering each interpersonal interaction that lasted more than 10 min, and asked participants to rate their own behavior, the partner's perceived behavior, and several affect adjectives. Next, they subjected each participant's time series on self-dominance, self-affiliation, other-dominance, other-affiliation, positive affect, anxiety, hostility, guilt, and sadness to exploratory principal axis factoring, obtaining a factor solution per participant. To compare these person-specific factor solutions, they used Tucker's congruency coefficient, showing that, although there is considerable heterogeneity in the factor structures (both in the number of factors and in the loading patterns), there are also important similarities. For example, for all but one participant there was a factor on which all negative emotions loaded strongly, implying that negative emotions tend to align together for most individuals. In summary, the combination of p-technique factor analysis and Tucker's congruency coefficient allows one to look for both similarities and differences in person-specific factor structures. At the same time, with large numbers of participants the multitude of person-specific factor structures can get unwieldy, which makes it easy to lose the overview in terms of generalizability.

Mixture simultaneous factor analysis. The third option for determining the factor structure in ambulatory assessment data occupies the middle ground between nonidiographic and idiographic methods in that it builds on the strengths of the multilevel factor model (i.e., parsimony) and those of the p-technique factor technique (i.e., not presupposing that all individuals are drawn from a single population). This method, called mixture simultaneous factor analysis (MSFA; De Roover, Vermunt, Timmerman, & Ceulemans, 2017) combines common factor analysis at the level of the repeated observations with mixture modeling at the level of the different participants, thereby offering a method that looks for clusters of participants with similar factor structures. In other words, MSFA searches for groups of individuals for whom the different behaviors, affects, and cognitions are organized in a similar way. In this way, MSFA can be seen as a mixture of the top-down and the bottom-up approach to modeling personality structures. Obviously, the major advantage of such an approach is that, because of the mixture component, MSFA yields a parsimonious solution that reveals the most important between-person differences in the factor structures using only a few cluster-specific factor loading matrices. The downside is that the number of mixture components and the number of factors are selected by testing and comparing a wide range of models that differ with respect to the number of factors and the number of clusters. Apart from the fact that this requires considerable computation time, model selection criteria for MSFA still need to be further developed (De Roover et al., 2017). To circumvent this issue, prior knowledge on the number of clusters and numbers of factors may be used (if available).

Although this method has not yet been applied to data in the personality domain, De Roover and colleagues (2017) used it to analyze a cross-cultural data set on norms for experienced emotions, including 10,018 participants from 48 countries. In a first step, they performed a multigroup exploratory factor analysis, which showed an excellent fit with a two-factor solution. Next, they compared several two-factor MSFA models with different numbers of clusters. Because these models always reflected the same two extreme factor structures, they proceeded with a two-cluster, two-factor solution. The first cluster was comprised of the less developed, more conservative countries, and in this cluster pride loaded primarily on the negative factor, whereas in the second cluster (made up of more developed, progressive countries), pride primarily loaded on the positive emotions factor. Moreover, De Roover and colleagues (2017) found higher unique variances for all emotions in the more developed, progressive countries, which means that there is more idiosyncratic variability

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

in these countries. Thus, based on MSFA, it was found that there are important between-country differences in the extent to which pride is positively or negatively valued by society and in the extent to which countries within a cluster resemble each other. Although MSFA has not yet been applied to personality data, we believe that it holds great promise because it offers a parsimonious way to consider between-person differences and similarities in factor structures.

An issue with multilevel factor analysis, p-technique factor analysis, and mixture simultaneous factor analysis is that these techniques do not account for the temporal relationships resulting from repeatedly measuring the same individual. For example, whereas multilevel factor analysis accounts for the nesting of measurements within individuals, it assumes *independent* normally distributed errors at the within-person level; an assumption that is typically violated in repeated measures data (Reise et al., 2005). Similarly, also in p-technique factor analysis and MSFA it is assumed that the observations are independent and identically distributed, meaning that the temporal patterning of the data is not taken into account. In response to this limitation, one could consider preprocessing the data to remove such temporal dependencies (Brose & Ram, 2012) or to include more complex error structures that model the dependencies through the residuals (Reise et al., 2005). Also, to address the ignorance of time dependencies, p-technique factor analysis has been extended to dynamic factor analysis, in which these time dependencies are explicitly modeled at the latent factor level, allowing for carryover or spillover effects from one occasion to the next (Molenaar, 1985). However, despite these possible extensions and further developments, we believe that the methods presented here remain useful when the goal is to study the structure of personality using intensive longitudinal ambulatory assessment data.

Describing Process Aspects of Personality

Turning to the study of the *personality processes*—or the study of directional relationships between personality components—we again discuss three analytical approaches that are useful for the analysis of intensive longitudinal data. These methods again range from dynamic, nonidiographic methods to dynamic, idiographic methods. Similar to our prior discussion, the goal of dynamic, nonidiographic methods is to reveal the dynamic, within-person process that characterizes the aggregate of all individuals in the sample, which is why we refer to them as methods that test "within-person processes." Dynamic, idiographic methods, instead, yield person-specific models, which means that they test "person-specific processes." Parallel to our discussion of the structure of personality, we will consider three analytic approaches that differ in their positioning on the dynamic, nonideographic-dynamic idiographic, continuum. For each method we will discuss unique advantages and disadvantages.

Multilevel regression analysis. First, multilevel regression (and more generally multilevel structural equations modeling) models are useful when the goal is to test within-person processes because they can simultaneously test fixed effects (representing average sample-level coefficients) and random effects (representing person-specific deviations from these fixed effects). Because these models allow testing between-person differences in within-person associations, it has been claimed that in the multilevel regression model "the psychology of each person is considered separately, preserving much of the goal of idiographic analysis" (Conner et al., 2009, p. 297). However, it is important to realize that multilevel regression and multilevel SEM models test how individual within-person relationships differ with respect to the within-person relationship shown by others in the sample (Beltz et al., 2016). That is, the multilevel regression model assumes that all individuals are drawn from a single population that is fully described by population-level averages and normally distributed differences around these averages (Brose & Ram, 2012). Thus, because in multilevel regression analysis a within-person process is fitted to all participants simultaneously, this method tests within-person processes (Wright et al., 2015). There are numerous examples in personality psychology where researchers have used multilevel regression analysis to study within-person processes. For example, Rauthmann, Jones, and Sherman (2016) used it to study spillovers among and between situational experiences and personality states. Using experience sampling data (eight measurements per day for 7 days) on 210 participants, they tested both contemporaneous as well as cross-lagged relationships between situation experiences and personality states, finding that situation experiences and personality states were contemporaneously related to each other, and that situation experiences predicted personality states as well as the other way around (although the effect sizes were very small). An important quality of this model is its parsimony, because the multilevel regression model fits the same within-person process to all individuals in the sample. However, for the same reason, this model often fails to provide a precise match to any given individual's actual patterning of behaviors, cognitions, and affects (Wright et al., 2015).

***n* = 1 vector autoregressive models.** If the goal is to study person-specific processes, an alternative analytical approach is needed. In what follows, we will discuss a model that is well suited to do so, namely the $n = 1$ vector autoregressive model (VAR; Hamilton, 1994). VAR models are specifically designed to test how a vector or a set of variables affect themselves (i.e., autoregression) as well as the other variables (i.e., cross-regressive effects) across time. In VAR models, the time lag can be varied, with for example a VAR model of order one testing associations between the current and the previous measurement occasion, and a VAR model of order two testing associations with the previous measurement occasion and the measurement occasion before. Generally speaking, there are two types of VAR models: reduced form VAR models and structural VAR (SVAR) models. The difference between these models lies in the way they deal with contemporaneous relationships. Whereas reduced form VAR models do not include contemporaneous relationships, these contemporaneous relationships are explicitly modeled on top of lagged relationships in SVAR models. This difference has some important implications. First, whereas reduced form VAR models are well suited for forecasting, or making predictions about future time points based on previous time points, the coefficients in a reduced form VAR model are not directly interpretable in terms of the causal process (Bulteel, Tuerlinckx, Brose, & Ceulemans, 2016). In other words, reduced form VAR models can be used for description and for prediction of new or future observations based on previous observations, but they do not allow for causal explanation of the process giving rise to these observations (see Shmueli, 2010 for a discussion of the philosophical and statistical differences between causal

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

or explanatory modeling and forecasting). SVAR models, in turn, yield unbiased parameter estimates of the causal relationships (Gates, Molenaar, Hillary, Ram, & Rovine, 2010), but the price to pay is that the additional contemporaneous effects increase model complexity to the extent that several constraints are needed to identify the model (because not all contemporaneous and lagged relationships can be estimated simultaneously). More important, Kim, Zhu, Chang, Bentler, and Ernst (2007) addressed this issue with their unified SEM approach (uSEM). By combining SEM and VAR modeling, uSEM identifies both contemporaneous and lagged relationships. Moreover, to deal with the identifiability issue, Gates and colleagues developed an automatic search procedure in which one starts from an empty model and then uses the modification indices (i.e., generalized Lagrange multiplier tests) provided by standard SEM packages to iteratively add contemporaneous and/or lagged paths to the model (Gates et al., 2010). The result of this procedure is the derivation of the set of contemporaneous and lagged relationships that best matches the data without necessitating prior theory or knowledge on the relationships (that is difficult when testing person-specific processes; Gates et al., 2010). An application of individual-level uSEM can be found in Wright et al. (2015), who applied this method to study how variability in negative affect, detachment, disinhibition, and hostility are influenced by the contemporaneous and lagged variability in these pathology domains. By analyzing the data of four exemplar participants, Wright et al. (2015) showed that there is substantial between-person variability in the patterns of associations between the different pathology domains. For example, whereas negative affect positively predicted detachment for two of the participants, these pathology domains were unrelated for a third participant, and showed a complicated pattern of positive reciprocal contemporaneous relationships combined with a negative lagged relationship from detachment to negative affect for the fourth participant. In summary, both reduced form $n = 1$ VAR models and the uSEM model can be used to test person-specific processes. The most important differences are that uSEM, but not reduced form VAR models, yields unbiased parameter estimates of the causal relationships, while uSEM, but not reduced form VAR models, needs the imposition of parameter restrictions to be able to identify and estimate the model.

Clusterwise VAR modeling and group iterative multiple model estimation. Whereas reduced form $n = 1$ VAR models and the uSEM model allow for a true idiographic analysis of personality processes because of their bottom-up approach, an important issue with testing such person-specific processes is the exponential increase in complexity when the number of participants increases. For this reason, researchers often want to look for groups of individuals that are characterized by similar person-specific processes (e.g., Zheng, Wiebe, Cleveland, Molenaar, & Harris, 2013). In what follows, we discuss two models specifically developed to bridge the nomothetic and idiographic approaches. The first model—clusterwise VAR modeling—does so by looking for subgroups of people with similar reduced form VAR regression weights, thereby examining qualitative between-person differences in person-specific processes (Bulteel et al., 2016). Because it is an extension of the reduced form VAR model, the clusterwise VAR model is useful for forecasting, but its coefficients are not directly interpretable in terms of the causal process. Applying the clusterwise VAR model to repeated measurements of depression-related symptoms in young women, Bulteel and colleagues (2016) found two clusters of individuals characterized by a differential persistence of the previous state. Moreover, the cluster that appeared to be more resistant to change had significantly higher depression scores, thereby supporting the role of an inert affective system for depression.

A second model that bridges the nomothetic and idiographic approaches is the group iterative multiple model estimation, or GIMME model (Gates & Molenaar, 2012). This model is an extension of uSEM in that it uses uSEM to capture the person-specific associations between the study variables, while taking advantage of the nomothetic information by including group-level information in the individual-level solutions (Beltz et al., 2016). In GIMME, the group-level information is captured by retaining those relationships that are robust across participants. By combining information about relationships that replicate across participants with the participant-specific relationships—through the creation of person-specific graphs containing a group-level structure—the GIMME model clearly combines the strengths of both the idiographic and nomothetic approach (Beltz et al., 2016). Moreover, the GIMME model has recently been extended with a variant that offers subgrouping, which allows the grouping of individuals with similar dynamic process models (Gates, Lane, Varangis, Giovanello, & Guiskewicz, 2017). This novel feature is an interesting one because vast heterogeneity in the dynamic process models is more often the rule than the exception, and the GIMME model with subgrouping offers an elegant way to reduce this immense complexity through the simultaneous study of general, shared, and person-specific dynamic processes.

An application of the traditional GIMME model can be found in Beltz et al. (2016), wherein the model is applied to intensive repeated measurements data (a median of 95 observations per individual) of 25 individuals with a personality disorder. GIMME analyses were run on the time series data of negative affect, detachment, disinhibition, and hostility, revealing that the person-specific models contained a mixture of group-level relationships and participant-specific relationships. In particular, the positive contemporaneous relationship between negative affect and detachment, as well as the one between disinhibition and hostility generalized across participants (note that the strength of these relationships did differ across participants). In addition to these group-level relationships, numerous participant-specific relationships were found. For example, for some participants carry-over effects from one day to the next were found for negative affect, whereas for other participants their present day negative affect was not influenced by their negative affectivity of the previous day. An application of the subgrouping feature of the GIMME model can be found in Wright et al. (2017), who applied this model to intensive longitudinal behavioral data collected in a sample of individuals with a personality disorder. In summary, although the clusterwise VAR model and the GIMME model both bridge the idiographic and the nomothetic approach, they differ in the extent to which they weigh both approaches. In particular, the GIMME model has a stronger idiographic basis as it starts from a pure bottom-up approach in which first all person-specific process models are tested after which the paths that generalize across participants are retained for the group-level structure. The clusterwise VAR model, instead, does not rely on such a bottom-up approach, but immediately takes a top-down approach in that it starts from a group-

level solution in each of the clusters that is then iteratively updated to improve the clustering.

Conclusion

In the present article, we discussed the potential of ambulatory assessment for studying the structure and process of personality. Methods that assess how different components of personality dynamically operate within one and the same individual are not only essential from a theoretical point of view, but also for clinical practice, in which the target of any treatment or intervention is always located at the individual level. Because such interventions typically not only target changes in individual variables, but also in the dynamic associations between these variables, it is important to not only conceptualize, but also assess psychopathology as a process. In such a situation, the revelation of person-specific structures and processes can be a tremendous help because they show how the different elements of an individual's personality relate to each other, thereby offering a roadmap that may efficiently guide intervention toward the intended change. Hence, we believe that ambulatory assessment, paired with the analytical techniques discussed in this article, represents a much-needed methodology that has the potential to deepen our understanding of human nature in general, and the nature and development of individual-level constellations of personality components in particular.

From this perspective, the present article provides an introduction to issues in designing and implementing ambulatory assessment in the personality domain. By doing so, we hope to familiarize personality scholars with these methods and to provide guidance for their use in the field of personality psychology and beyond. Moreover, with this article, we also hope to inspire scholars to conduct research on the expression of personality in everyday life. Such research is crucial and may help to further develop the much-needed empirically based standards to rely on when implementing ambulatory assessment research in the personality domain.

References

Allik, J., & Realo, A. (2017). Universal and specific in the five-factor model of personality. In T. A. Widiger (Ed.), *The Oxford handbook of the five factor model* (pp. 173–190). New York, NY: Oxford University Press.

Allport, G. W. (1937). *Personality: A psychological interpretation*. New York, NY: Holt.

Baird, B. M., Le, K., & Lucas, R. E. (2006). On the nature of intraindividual personality variability: Reliability, validity, and associations with well-being. *Journal of Personality and Social Psychology, 90,* 512–527. http://dx.doi.org/10.1037/0022-3514.90.3.512

Barta, W. D., Tennen, H., & Litt, M. D. (2011). Measurement reactivity in diary research. In M. Mehl & T. Conner (Eds.), *Handbook of research methods for studying daily life* (pp. 108–123). New York, NY: Guilford Press.

Beal, D. J., & Weiss, H. M. (2003). Methods of ecological momentary assessment in organizational research. *Organizational Research Methods, 6,* 440–464. http://dx.doi.org/10.1177/1094428103257361

Beltz, A. M., Wright, A. G., Sprague, B. N., & Molenaar, P. C. (2016). Bridging the nomothetic and idiographic approaches to the analysis of clinical data. *Assessment, 23,* 447–458. http://dx.doi.org/10.1177/1073191116648209

Bolger, N., Davis, A., & Rafaeli, E. (2003). Diary methods: Capturing life as it is lived. *Annual Review of Psychology, 54,* 579–616. http://dx.doi.org/10.1146/annurev.psych.54.101601.145030

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

Brose, A., & Ram, N. (2012). Within-person factor analysis: Modeling how the individual fluctuates and changes across time. In M. R. Mehl & T. S. Conner (Eds.), *Handbook of research methods for studying daily life* (pp. 43–61). New York, NY: Guilford Press.

Bulteel, K., Tuerlinckx, F., Brose, A., & Ceulemans, E. (2016). Clustering vector autoregressive models: Capturing qualitative differences in within-person dynamics. *Frontiers in Psychology, 7,* 1540.

Cattell, R. B. (1963). The structuring of change by P-technique and incremental R-technique. In C. W. Harris (Ed.), *Problems in measuring change* (pp. 167–198). Madison, WI: University of Wisconsin Press.

Conner, T. S. (2015, May). *Experience sampling and ecological momentary assessment with mobile phones*. Retrieved from http://www.otago.ac.nz/psychology/otago047475.pdf

Conner, T. S., Tennen, H., Fleeson, W., & Barrett, L. F. (2009). Experience sampling methods: A modern idiographic approach to personality. *Social and Personality Psychology Compass, 3,* 292–313. http://dx.doi.org/10.1111/j.1751-9004.2009.00170.x

Costa, P. T., & McCrae, R. R. (1994). 'Set like plaster?' Evidence for the stability of adult personality. In T. Heatherton & J. Weinberger (Eds.), *Can personality change?* (pp. 21–40). Washington, DC: American Psychological Association. http://dx.doi.org/10.1037/10143-002

Costa, P. T., & McCrae, R. R. (2017). The NEO inventories as instruments of psychological theory. In T. A. Widiger (Ed.), *The Oxford handbook of the five factor model* (pp. 11–37). New York, NY: Oxford University Press.

Debusscher, J., Hofmans, J., & De Fruyt, F. (2014). The curvilinear relationship between state neuroticism and momentary task performance. *PLoS ONE, 9,* e106989. http://dx.doi.org/10.1371/journal.pone.0106989

Debusscher, J., Hofmans, J., & De Fruyt, F. (2016a). The effect of state Core Self-Evaluations on task performance, organizational citizenship behaviour, and counterproductive work behaviour. *European Journal of Work and Organizational Psychology, 25,* 89–104. http://dx.doi.org/10.1080/1359432X.2014.983085

Debusscher, J., Hofmans, J., & De Fruyt, F. (2016b). Do personality states predict momentary task performance? The moderating role of personality variability. *Journal of Occupational and Organizational Psychology, 89,* 330–351. http://dx.doi.org/10.1111/joop.12126

Debusscher, J., Hofmans, J., & De Fruyt, F. (2017). The multiple face(t)s of state conscientiousness: Predicting task performance and organizational citizenship behavior. *Journal of Research in Personality, 69,* 78–85. http://dx.doi.org/10.1016/j.jrp.2016.06.009

De Clercq, B., Verbeke, L., De Caluwé, E., Vercruysse, T., & Hofmans, J. (2017). Understanding adolescent personality pathology from growth trajectories of childhood oddity. *Development and Psychopathology, 29,* 1403–1411.

De Roover, K., Vermunt, J. K., Timmerman, M. E., & Ceulemans, E. (2017). Mixture simultaneous factor analysis for capturing differences in latent variables between higher level units of multilevel data. *Structural Equation Modeling, 24,* 506–523. http://dx.doi.org/10.1080/10705511.2017.1278604

Fisher, C. D., & To, M. L. (2012). Using experience sampling methodology in organizational behavior. *Journal of Organizational Behavior, 33,* 865–877. http://dx.doi.org/10.1002/job.1803

Fleeson, W. (2001). Toward a structure- and process-integrated view of personality: Traits as density distribution of states. *Journal of Person-*

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

ality and Social Psychology, 80, 1011–1027. http://dx.doi.org/10.1037/0022-3514.80.6.1011

Fleeson, W., Malanos, A. B., & Achille, N. M. (2002). An intraindividual process approach to the relationship between extraversion and positive affect: Is acting extraverted as "good" as being extraverted? *Journal of Personality and Social Psychology, 83,* 1409–1422. http://dx.doi.org/10.1037/0022-3514.83.6.1409

Fleeson, W., & Noftle, E. E. (2012). Personality research. In M. R. Mehl & T. S. Conner (Eds.), *Handbook of research methods for studying daily life* (pp. 525–538). New York, NY: Guilford Press.

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. M. (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. 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. http://dx.doi.org/10.1016/j.neuroimage.2009.12.117

Gayles, J. G., & Molenaar, P. C. M. (2013). The utility of person-specific analyses for investigating developmental processes: An analytic primer on studying the individual. *International Journal of Behavioral Development, 37,* 549–562. http://dx.doi.org/10.1177/0165025413504857

Geldhof, G. J., Preacher, K. J., & Zyphur, M. J. (2014). Reliability estimation in a multilevel confirmatory factor analysis framework. *Psychological Methods, 19,* 72–91. http://dx.doi.org/10.1037/a0032138

Gosling, S. D., Rentfrow, P. J., & Swann, W. B., Jr. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality, 37,* 504–528. http://dx.doi.org/10.1016/S0092-6566(03)00046-1

Hamaker, E. L. (2012). Why researchers should think "within-person": A paradigmatic rationale. In M. R. Mehl & T. S. Conner (Eds.), *Handbook of methods for studying daily life* (pp. 43–61). New York, NY: Guilford Press Publications.

Hamaker, E. L., Dolan, C. V., & Molenaar, P. C. M. (2005). Statistical modeling of the individual: Rationale and application of multivariate stationary time series analysis. *Multivariate Behavioral Research, 40,* 207–233. http://dx.doi.org/10.1207/s15327906mbr4002_3

Hamaker, E. L., Nesselroade, J. R., & Molenaar, P. C. M. (2007). The integrated state-trait model. *Journal of Research in Personality, 41,* 295–315. http://dx.doi.org/10.1016/j.jrp.2006.04.003

Hamilton, J. D. (1994). *Time series analysis*. Princeton, NJ: Princeton University Press.

Hofmans, J., Debusscher, J., Dóci, E., Spanouli, A., & De Fruyt, F. (2015). The curvilinear relationship between work pressure and momentary task performance: The role of state and trait core self-evaluations. *Frontiers in Psychology, 6,* 1680. http://dx.doi.org/10.3389/fpsyg.2015.01680

Hox, J. J. (1993). Factor analysis of multilevel data: Gauging the Muthén model. In J. H. L. Oud & R. A. W. van Blokland-Vogelesang (Eds.), *Advances in longitudinal and multivariate analysis in the behavioral sciences* (pp. 141–156). Nijmegen, NL: ITS.

Jarnecke, A. M., & South, S. C. (2017). Behavior and molecular genetics of the five factor model. In T. A. Widiger (Ed.), *The Oxford handbook of the five factor model* (pp. 301–318). New York, NY: Oxford University Press.

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

Kubiak, T., & Krog, K. (2012). Computerized sampling of experiences and behavior. In M. R. Mehl & T. S. Conner (Eds.), *Handbook of research methods for studying daily life* (pp. 160–175). New York, NY: Guilford Press.

Kuppens, P., Oravecz, Z., & Tuerlinckx, F. (2010). Feelings change: Accounting for individual differences in the temporal dynamics of affect. *Journal of Personality and Social Psychology, 99,* 1042–1060. http://dx.doi.org/10.1037/a0020962

Lorenzo-Seva, U., & ten Berge, J. M. F. (2006). Tucker's congruence coefficient as a meaningful index of factor similarity. *Methodology, 2,* 57–64. http://dx.doi.org/10.1027/1614-2241.2.2.57

Mehl, M. R., & Robbins, M. L. (2012). Naturalistic observation sampling: The electronically activated recorder (EAR). In M. R. Mehl & T. S. Conner (Eds.), *Handbook of research methods for studying daily life* (pp. 176–192). New York, NY: Guilford Press.

Minbashian, A., Wood, R. E., & Beckmann, N. (2010). Task-contingent conscientiousness as a unit of personality at work. *Journal of Applied Psychology, 95,* 793–806. http://dx.doi.org/10.1037/a0020016

Molenaar, P. C. M. (1985). A dynamic factor model for the analysis of multivariate time series. *Psychometrika, 50,* 181–202. http://dx.doi.org/10.1007/BF02294246

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

Muthén, B. (1994). Multilevel covariance structure analysis. *Sociological Methods & Research, 22,* 376–398. http://dx.doi.org/10.1177/0049124194022003006

Nezlek, J. B. (2017). A practical guide to understanding reliability in studies of within-person variability. *Journal of Research in Personality, 69,* 149–155. http://dx.doi.org/10.1016/j.jrp.2016.06.020

Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10 item short version of the Big Five Inventory in English and German. *Journal of Research in Personality, 41,* 203–212. http://dx.doi.org/10.1016/j.jrp.2006.02.001

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,* 893–909. http://dx.doi.org/10.1177/0146167216647360

Reis, H. T., Gable, S. L., & Maniaci, M. R. (2014). Methods for studying everyday experience in its natural context. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (2nd ed., pp. 373–403). New York, NY: Cambridge University Press.

Reise, S. P., Ventura, J., Nuechterlein, K. H., & Kim, K. H. (2005). An illustration of multilevel factor analysis. *Journal of Personality Assessment, 84,* 126–136. http://dx.doi.org/10.1207/s15327752jpa8402_02

Roberts, B. W., & DelVecchio, W. F. (2000). The rank-order consistency of personality traits from childhood to old age: A quantitative review of longitudinal studies. *Psychological Bulletin, 126,* 3–25. http://dx.doi.org/10.1037/0033-2909.126.1.3

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

Saucier, G. (1994). Mini-markers: A brief version of Goldberg's unipolar big-five markers. *Journal of Personality Assessment, 63,* 506–516. http://dx.doi.org/10.1207/s15327752jpa6303_8

Shmueli, G. (2010). To explain or to predict? *Statistical Science, 25,* 289–310. http://dx.doi.org/10.1214/10-STS330

Shrout, P. E., & Lane, S. P. (2011). Psychometrics. In M. R. Mehl & T. A. Conner (Eds.), *Handbook of research methods for studying daily life* (pp. 302–320). New York, NY: Guilford Press.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Soto, C. J., & Tackett, J. L. (2015). Personality traits in childhood and adolescence: Structure, development, and outcomes. *Current Directions in Psychological Science, 24,* 358–362. http://dx.doi.org/10.1177/0963721415589345

Trull, T. J., & Ebner-Priemer, U. W. (2009). Using experience sampling methods/ecological momentary assessment (ESM/EMA) in clinical assessment and clinical research: Introduction to the special section. *Psychological Assessment, 21,* 457–462. http://dx.doi.org/10.1037/a0017653

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., & Ebner-Priemer, U. (2014). The Role of Ambulatory Assessment in Psychological Science. *Current Directions in Psychological Science, 23,* 466–470. http://dx.doi.org/10.1177/0963721414550706

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

Widiger, T. A. (2017). *The Oxford handbook of the five factor model*. New York, NY: Oxford University Press.

Wille, B., Hofmans, J., Feys, M., & De Fruyt, F. (2014). Maturation of work attitudes: Correlated change with Big Five personality traits and reciprocal effects over 15 years. *Journal of Organizational Behavior, 35,* 507–529. http://dx.doi.org/10.1002/job.1905

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., Gates, K., Arizmendi, C., Lane, S., Woods, W., & Edershile, E. A. (2017, July 22). *Focusing personality assessment on the person: Modeling general, shared, and person specific processes in personality and psychopathology*. Retrieved from osf.io/nf5me

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,* 484–495. http://dx.doi.org/10.1177/1073191116653829

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

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

Zheng, Y., Wiebe, R. P., Cleveland, H. H., Molenaar, P. C., & Harris, K. S. (2013). An idiographic examination of day-to-day patterns of substance use craving, negative affect and tobacco use among young adults in recovery. *Multivariate Behavioral Research, 48,* 241–266. http://dx.doi.org/10.1080/00273171.2013.763012

Received June 4, 2017
Revision received November 4, 2017
Accepted November 6, 2017 ■

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

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!