The Use of Intensive Longitudinal Methods in Explanatory Personality Research

HARAN SENED[1]*, [iD], GAL LAZARUS[1], MARCI E.J. GLEASON[2], ESHKOL RAFAELI[1,3] [iD] and WILLIAM FLEESON[4]

[1]*Bar-Ilan University, Ramat Gan, Israel*
[2]*University of Texas at Austin, Austin, TX USA*
[3]*Barnard College, Columbia University, New York, NY USA*
[4]*Wake Forest College, Winston-Salem, NC USA*

Abstract: Intensive longitudinal methods (ILMs), in which data are gathered from participants multiple times with short intervals (typically 24 hours or less apart), have gained considerable ground in personality research and may be useful in exploring causality in both classic personality trait models and more novel contextualized personality state models. We briefly review the various terms and uses of ILMs in various fields of psychology and present five main strategies that can help researchers infer causality in ILM studies. We discuss the use of temporal precedence to establish causality, through both lagged analyses and natural experiments; the use of external measures and peer reports to go beyond self-report data; delving deeper into repeated measures to derive new indices; the use of contextual factors occurring during the measurement period; and combining experimental methods and ILMs. These strategies are illustrated by examples from existing research and by new empirical findings from two dyadic daily diary studies (N = 80 and N = 108 couples) and an experience sampling method study of personality states (N = 52). We conclude by offering a short checklist for designing ILM studies with causality in mind and look at the applicability of these strategies in the intersection of personality psychology and other psychological research domains. Copyright © 2018 European Association of Personality Psychology

Key words: causality; empirical methods; experience sampling methods; intensive longitudinal methods

Intensive longitudinal methods (ILMs; Bolger & Laurenceau, 2013) are research methods in which participants are asked to provide research data multiple times, in short intervals (typically daily or even multiple times a day), as they go about their daily lives. The current article will focus on the potential for using ILMs to support causal inferences in personality research and will demonstrate various adjustments that strengthen such inferences. These methods have been used successfully in many fields of study, including physical health (e.g. for blood pressure monitoring; Hodgkinson et al., 2011), health psychology (e.g. Swendsen et al., 2000), mental health (e.g. Trull & Ebner-Priemer, 2009), social psychology (e.g. Swim, Hyers, Cohen & Ferguson, 2001), close relationships (e.g. Howland & Rafaeli, 2010), occupational research (e.g. Farnworth, Mostert, Harrison, & Worrell, 1996), and of course personality (for a review, see Fleeson & Noftle, 2012).

Intensive longitudinal methods include many different specific data collection techniques that go by a variety of names. For example, experience sampling methods (ESMs; Csikszentmihalyi & Larson, 2014) is a term used mainly in

personality and social psychology, which tends to refer to studies using subjective questionnaires focused on specified events that can occur multiple times in any given day. Ecological momentary assessment (Stone & Shiffman, 1994) is a term used mainly in clinical and medical research; such studies typically add objective measures such as location to an ESM framework. Daily diaries (Bolger, Davis & Rafaeli, 2003) and daily process studies (Tennen, Affleck, Armeli, & Carney, 2000) limit participant input to one or two entries per day. Ambulatory monitoring (e.g. O'Brien et al., 2000) repeatedly or continuously collects health measurements such as blood pressure. All of these methods are often referred to using the catch-all terms of intensive repeated measures in naturalistic settings (IRM-NS; Moskowtiz, Russell, Sadikaj, G., & Sutton, 2009) and/or ambulatory assessment (Fahrenberg, Myrtek, Pawlik, & Perrez, 2007). Although there are some differences between these terms in the exact scope of their referents (e.g. ESM, ecological momentary assessment, daily diaries, and ambulatory monitoring referring more often to the data collection methods, vs ILMs and IRM-NS referring more often to the nature of the obtained data), all share the same core idea—that of repeated, within-person measures, collected outside of the laboratory, and separated by short time intervals.

Intensive longitudinal methods have several qualities that allow them to make unique methodological improvements in the field of personality research. First, ILM studies collect data from participants in close temporal proximity to internal

*Correspondence to: Haran Sened, Department of Psychology, Bar-Ilan University, Ramat-Gan 5290002, Israel.
E-mail: haranse@gmail.com
Haran Sened and Gal Lazarus contributed equally to this article.

or external events and are thus much more 'experience-near' than studies using typical cross-sectional data collection methods. When the data involved are self-reports, this proximity can help reduce memory-related biases (for a review, see Stone, Shiffman, & DeVries, 1999). Additionally, the measurement takes place as participants go about their daily lives, and not in (somewhat artificial) laboratory environments. Moreover, participants' reports capture relatively brief periods of time, increasing their validity, and the multiple repeated measurement of variables allows researchers to examine between-person and within-person variance separately. This sort of immediate reporting process relies more heavily on episodic memory than on semantic memory, thus reducing the risk of interpretation biases (e.g. Robinson & Clore, 2002). Finally, recent experimental research using ILM designs suggests that initial reports of negative mood and health symptoms are exaggerated, which complicates our understanding of any one-time measurement and argues strongly for the importance of repeated measures when studying self-reported outcomes (Shrout et al., 2018).

Intensive longitudinal methods' multiple measurements per individual across time and situations also allow for addressing a neglected but critical aspect of personality: how processes unfold and operate within a person. They allow the identification of person-specific within-person associations between motivations, cognitions, affect, and behaviours (for reviews, see Conner, Tennen, Fleeson, & Barrett, 2009; Hamaker, 2012). These within-person associations are termed *idiographic* and should be distinguished from *nomothetic* associations, which capture the covariation of variables across a population of individuals.

Importantly, although within-person (idiographic) associations may show some similarity to between-person (nomothetic) associations (e.g. Wilt, Noftle, Fleeson, & Spain, 2012), the two cannot be assumed to be always isomorphic (Borsboom, Mellenbergh, & Van Heerden, 2003; Bos et al., 2017; Hamaker, Dolan, & Molenaar, 2005). Empirically, associations between constructs computed across individuals have been shown to be different than ones computed within individuals (for reviews, see Molenaar & Campell, 2009; Molenaar, Huizenga, & Nesselroade, 2003). Conceptually, models that address between-individual differences tend to be classificatory or descriptive in nature: That is, they address *trait-like phenomena*. In contrast, models that address within-individual covariation tend to be explanatory: That is, they address (sometimes causal) *state-like processes* (e.g. Cervone, 2005; Harré, 1998).

Intensive longitudinal methods, which are particularly suited for exploring idiographic hypotheses, also help build a bridge between theories that emphasize personality dispositions and ones that focus on processing dynamics (e.g. Fleeson & Jayawickreme, 2015; Revelle & Condon, 2015; Wright, 2014). Specifically, when idiographic patterns exist and have substantial variability between individuals, ILMs can also help us identify within-person factors (e.g. daily events and momentary salience of certain cognitions) as well as between-person factors (e.g. personality traits and demographic characteristics), which could explain this variability.

Finally, ILMs generate data that are particularly useful for capturing momentary states. This momentary-ness—that is, the tendency of certain constructs to change across time rather than being fixed—is easily understood with regard to affective phenomena (e.g. emotions or moods). But recent years have brought a growing interest in other state-like aspects of psychological constructs, which, historically, were considered to be stable. For example, momentary manifestations of personality traits or of psychopathological characteristics have been found to vary across time and context (e.g. Fleeson & Law, 2015; Roche, Jacobson, & Pincus, 2016; Sherman, Rauthmann, Brown, Serfass, & Jones, 2015; for a review, see Fleeson, 2017). These personality *states* are relatively short-term patterns that characterize individuals' personality at a certain moment and that consist of the salient affect, behaviours, cognitions, and desires of the individual at that moment in time. In other words, these states capture individuals' immediate and fleeting way of being.[1]

Models that address personality states (e.g. Fleeson, 2017) consider these states' *content* to be isomorphic with that of their respective personality traits. For example, *trait* extraversion is marked by a general tendency towards assertiveness, boldness, and activity; correspondingly, *state* extraversion involves current behaviour characterized by assertiveness, boldness, and activity (as well as by their accompanied cognition, affect, and motivation). When it comes to causality, however, personality states and traits are only partially isomorphic. For example, some causal paths may be true only for stable traits (e.g. those that are affected by fixed brain structures) and not of states.

The current article will focus on the advantages of ILMs for the study of causality in personality research. We will review key methods and strategies to strengthen causal inferences that are currently employed in personality research and point to those used in neighbouring research domains that we believe can be adapted to the study of personality. Importantly, while we will touch on statistical issues, a thorough explanation of statistical methods is outside the scope of the current article. For in-depth statistical discussions, we recommend Bolger and Laurenceau (2013, see also Nestler, Grimm, & Schönbrodt, 2015). When applicable, we will include empirical examples of analyses that demonstrate the various methods. These analyses will come from three main datasets, detailed below (see https://osf.io/k67ba/ for full project descriptions and a list of additional publications related to these data).

Power analyses on ILM data are a complex issue; we recommend readers refer to statistical reviews such as Bolger and Laurenceau (2013, ch. 10) and Hox (2010, ch. 12) for detailed discussions. Bolger and Laurenceau note that while various formulae for calculating *a priori* power exist, they require researchers to assume the values of many parameters as ILM studies include multiple sources of randomness. When no previous pilot data exist to inform such assumptions, as in our case, researchers may use general rules of thumb; for example, Hox

[1]Interestingly, *self*-states, which are similar in many respects to personality states, play a central role in several approaches to psychopathology and psychotherapy (e.g. emotion-focused therapy, Greenberg, 2011; schema therapy, Rafaeli, Maurer, Lazarus, & Thoma, 2016; relational psychoanalysis, Bromberg, 1996).

suggests using least 50 Level 2 units (participants, in our case) with 20 Level 1 units in each (measurements, in our case), a rule followed by all three of the included datasets.

To obtain a *post hoc* estimation of whether a study was indeed adequately powered and to help design future studies, Bolger and Laurenceau (2013) recommend performing *post hoc* Monte Carlo power analyses, which assume that all estimated variables are indeed true in the study population, and calculate which percentage of simulated studies with the number of participants given would have resulted in a significant result. We have included such power estimations (β values) for all analyses; that said, some of the analyses are underpowered, as these are meant to demonstrate methods and not to reach comprehensive results.

We will focus on five main issues: the use of temporal precedence to strengthen causal claims; methods that go beyond self-report; extraction of new variables from repeated measurements; the use external contextual factors; and incorporation of experimental designs into ILMs.

DATASET DESCRIPTION

All three studies were approved by the appropriate university institutional review board.

Dataset A Fifty-two (age >18) Israeli university students participated in exchange for course credit. They completed online background questionnaires as well as 10 days of thrice-daily questionnaires, separated by intervals of at least 2 hours.

Dataset B Eighty-six adult (age >18) Israeli heterosexual couples were recruited for a larger project, via online ads and flyers on campus. Only couples cohabiting for at least 6 months were included. Six couples were excluded as one partner failed to complete six or more entries, leaving 80 couples. Couples were paid the equivalent of €100 for their participation. Participants completed background questionnaires, took part in a discussion task in the laboratory (data from the discussion task are not included in the current study) and completed a 35-day diary.

Dataset C One hundred and eight (age >18) Israeli heterosexual couples, expecting their first child, were recruited for a larger project, using social media advertising and flyers posted around the university campus. Participants completed background questionnaires in the third trimester of pregnancy, took part in a laboratory meeting 3 months after birth, and completed a 21-day diary shortly after. Five couples left the study after completing the background questionnaire, one additional couple left before starting the diary portion, and two completed less than six diary entries, leaving 100 couples in the study. Participants received a breakfast gift card worth approximately €20 for completing the background questionnaires and approximately €150 in cash for completing the remaining study parts.

USING TEMPORAL PRECEDENCE TO ESTABLISH CAUSALITY (LAGGED ANALYSES, TIME AS A COVARIATE)

Intensive longitudinal methods allow researchers to sample a variety of measures at multiple time points per individual (a design often referred to as panel data), and thus to establish within-person temporal precedence (Bolger & Laurenceau, 2013, ch. 5). In lagged analyses, researchers estimate the effects of independent variables (IVs) assessed at a prior measurement point (e.g. yesterday's conflict) on dependent variables (DVs) assessed at a successive measurement point (e.g. today's mood; Wickham & Knee, 2013). Such analyses are based on the assumption that the former exert their influence over the latter prospectively over time. Of course, such precedence is not a sufficient condition to establish causality, as other possible explanations have to be ruled out (e.g. the omitted-variable problem; Finkel, 1995). Significant sequential associations do allow a weaker form of causal inference, referred to as *Granger causality* (Granger, 1969), which signifies temporal prediction without a strong commitment to the presence of causal relations. To support stronger causal claims, researchers may wish to take the following statistical or methodological steps (for a review, see Hamaker & Wichers, 2017). A first step in strengthening causal claims requires accounting for stable between-person differences that may affect both DVs and IVs. This can be performed by constructing models that include relevant stable factors, by allowing random intercepts that capture stable processes that affect all measurements equally, and by centring the variables around the person's mean, thus separating within-person and between-person variance (e.g. Bolger & Laurenceau, 2013). This step emphasizes a main strength of ILM, specifically that within-person association cannot be explained by features that do not vary within a person, such as most stable personality variables we study. For example, within-person covariation of moral behaviour and happiness could not be explained by differences between people in socially desirable responding. However, even these procedures do not rule out all spurious associations caused by time-varying omitted variables.

A second important step is to include autoregressive coefficients for the DV, which represent the effect of a variable on itself at the next measurement point. This step is important to be sure that the temporal precedence of the IV before the DV is genuine (i.e. that the DV, although unmeasured, did not cause the IV earlier). This step can also be thought of as control for moment-to-moment stability. Such adjustment turns IVs into predictors of change scores (e.g. Finkel, 1995) and reduces the risk for reverse causation (e.g. Shrout et al., 2010; although see Falkenström, Finkel, Sandell, Rubel, & Holmqvist, 2017, for a caveat in this regard).

A third step would be to account for common time trends in both the IVs and DVs, as these too may result in spurious correlations (Granger & Newbold, 1974). This procedure is often referred to as 'time detrending' and is especially necessary in cases in which time is theoretically likely to create third variables. Notably, if the time trend includes the effect

of the putative causal factor, removing it may suppress that effect (e.g. Wang & Maxwell, 2015).

Beyond these three steps, ILM researchers must consider the length of the temporal interval between consecutive measurements. The strength of lagged associations is strongly tied to this length (Dorman & Griffin, 2015; Gollob & Reichardt, 1987). To select the 'optimal time lag' (Dorman & Griffin, 2015), researchers may need to rely on both empirical grounds (e.g. pilot data) and theoretical reasoning; selecting lags that are too short or too long may obscure meaningful temporal associations. Alternatively, when variables are expected to exert continuous effect on other variables, the temporal associations may be better addressed by continuous time models instead of traditional discrete analysis (e.g. de Haan-Rietdijk, Voelkle, Keijsers, & Hamaker, 2017; Deboeck & Preacher, 2016).

Another key issue is the direction of causality between two variables measured across time. As the dynamic causal influence between many variables is reciprocal, researchers may be interested in determining *causal dominance*. This can be performed by comparing within-person standardized cross-lagged coefficients from parallel models (e.g. Schuurman, Ferrer, de Boer-Sonnenschein, & Hamaker, 2016). That said, establishing causal dominance is not always necessary, as reciprocity is principally two causal processes that can simply be addressed independently.

Taking all the above into account, lagged analyses in ILMs may be employed for personality-relevant research in three major ways. First, lagged analyses may be useful in examining internal processes underlying personality. Indeed, the within-person architecture of personality is often manifested in the cross-temporal associations between motivations, affective states, cognitions, and behaviours as they unfold in daily life. For example, Zhang (2009) collected surveys twice weekly for 4 weeks and showed that experiences of interpersonal loss preceded subsequent increases in state attachment *anxiety*. Exploring the relationship between momentary ruminative self-focus and negative affect by assessing them eight times daily for 1 week, Moberly and Watkins (2008) found bidirectional cross-lagged associations, indicating a mutual causal process between the two.

Second, a causal role of personality factors can become evident when they moderate associations between momentary or daily mental states, behaviours, or interpersonal events and subsequent outcomes. Presumably, certain states or events (e.g. conflict) lead some people (but not others) to specific downstream behaviours and/or mental states (e.g. negative mood). Along these lines, several personality researchers have endeavoured to identify higher-order and lower-order personality characteristics that influence such processes. For example, Conway, Rogelberg, and Pitts (2009) collected five surveys per day for five workdays to show that positive affect's association with later helping behaviours was dependent on trait altruism. Similarly, Ilies, Johnson, Judge, and Keeney (2011) sampled work-related interpersonal conflicts for 2 weeks to find that they had a stronger influence on negative affect for more agreeable individuals. Finally, Ford and Collins (2013) employed daily diaries to reveal that previous-day rejection had a stronger association with lower current-day health and well-being for individuals low in self-esteem.

Notably, ILM-based lagged analyses are not always necessary to infer specific processes characterizing certain personality features. Concurrent associations may also be highly instructive in this matter. For example, extensive work has been conducted regarding the mechanisms underlying the association between trait extraversion and trait positive affect (e.g. Howell et al., 2017; Wilt et al., 2012). In particular, Wilt et al. (2012) have identified one such mechanism as the presence of more extraversion states and positive affect states among more extraverted individuals. In another study, Mandel, Dunkley, and Moroz (2015) measured trait self-critical perfectionism at Time 1, daily stress and negative affect for two periods of 2 weeks, and finally depressive and anxious symptoms 4 years later. Daily stress–sadness reactivity explained the association between self-critical perfectionism at Time 1 and subsequent depressive and anxious symptoms.

Third, personality states can be seen not only as manifestations of personality traits but also as causal factors on a small scale within individual processes. For example, the social–cognitive mechanisms underlying personality states can influence the types of situations that individuals experience, encounter, or seek out. Rauthmann, Jones, and Sherman (2016) assessed repeated situational experiences and personality states and found both cross-sectional and lagged associations between the two. In another experience sampling study, Leikas and Ilmarinen (2016) collected measures of Big Five states, mood, stress, and fatigue five times daily for 12 days. They found that while extraverted and conscientious states were concurrently tied to positive mood and lower fatigue, they were tied to higher fatigue after a 3-hour lag, indicating a complex influence.

Another line of research dealing with causality in small-scale within-individual processes is network analysis (Borsboom & Cramer, 2013), a method that aims to capture the dynamic interrelationships of multiple constructs by conceptualizing each as a node in a network. When network analysis is used in personality research (Constantini et al., 2015), each node can represent a personality trait, and analysis of the network structure can lead to insight regarding the optimal clustering of personality traits and dynamics. By targeting inherently dynamic constructs such as personality states or behaviours, network analyses of ILM data can also incorporate temporality, by estimating the lagged (alongside the cross-sectional) associations between nodes (for more details on the practical and statistical aspects of network analysis, see also Epskamp, Borsboom & Fried, 2017).

We used Dataset A to illustrate one possible causal effect of a personality state. These data utilize the schema therapy (Young, Klosko, & Weishaar, 2003) concept of *schema modes*, which capture the predominant emotional and motivational states as well as the activated cognitions and coping reactions for an individual at a particular time (Rafaeli, Bernstein, & Young, 2010). In social–cognitive terms, modes can be considered as the working self-concept, namely, the part of the self that is active or operating at a specific

moment. The presently activated mode underlies individuals' expectations and interpretations of, and responses to, their immediate environment. The measure for assessing schema modes is adapted from the Schema Mode Inventory (SMI; Lobbestael, van Vreeswijk, Spinhoven, Schouten, & Arntz, 2010). It showed adequate internal consistency reliability and construct validity (i.e. moderate correlations with the SMI scores).

We assessed the extent to which the *Detached Protector* mode, a maladaptive coping mode marked by emotional avoidance and numbness, tends to succeed the *Vulnerable Child* mode, a mode characterized by overwhelming anxiety and a sense of helplessness. In theory, individuals are likely to react to the experienced pain of the Vulnerable Child mode, at least in situations when sufficient support is not available, with some sort of emotional detachment. To explore this hypothesis, we followed suggestions made by Bolger and Laurenceau (2013, ch. 5) for analysing within-subject causal processes. Specifically, we used multilevel models to predict participants' current levels of the Detached Protector mode from their previous entry's Vulnerable Child mode score, while adjusting for the previous entry's Detached Protector mode score, which was entered as a covariate into the model; this allowed us to reduce the possibility of reverse causation and to interpret the outcomes as entry-to-entry change scores. Additionally, we person-centred the IVs to statistically remove possible effects of their between-person differences (the data, full results, and the SAS syntax are available at https://osf.io/k67ba/). In line with our hypothesis, we found that the lagged Vulnerable Child mode scores were positively associated with the current Detached Protector mode scores ($b = 0.11$, $SE = 0.05$, $p = .038$, $\beta = .764$). Lagged analyses of this sort, in which the lagged outcome is used as a covariate, can help researchers establish temporal precedence and thus bolster causal accounts.

Beyond the employment of lagged analysis, ILMs allow for a thorough time-sensitive measurement of personality-related processes preceding or following major life events (e.g. transition to parenthood, loss of a loved one, major medical procedure, or bar examination) or periodic events (e.g. menstrual cycle or holidays). Personality characteristics may play a causal role in the time course of various processes preceding or following these events. For example, do extraverted individuals recover faster in the days following a romantic break-up? Do individuals higher in agreeableness associate quicker with colleagues at their new working place?

Several studies have employed ILMs before and/or after major life events such as bar examination (e.g. Shrout, Herman, & Bolger, 2006), major life changes such as quitting smoking (e.g. Cofta-Woerpel et al., 2011), and periodic events such as menstrual cycle (e.g. Nillni, Rohan, Mahon, Pinele, & Zvolensky, 2013), yet to the best of our knowledge, none examined the role of personality characteristics in their DVs' trajectories. Future studies that employ ILMs to explore the place of personality-related variables in the vicinity of important events will provide valuable information to the field.

ELIMINATING CONFOUNDS BY GOING BEYOND SELF-REPORT

Many of the measures employed by ILMs are self-report measures. These measures, while extremely common and useful, also suffer from various biases (J. H. Harvey, Hendrick & Tucker, 1988), which may create problems when attempting to infer causality. For one, self-report measures not only reflect the intended constructs but also assess individuals' perception of these constructs to some extent. For example, Buchanan (2016) demonstrated that self-report Five-Factor Model questionnaires were related to *self-reported* executive function problems, but not to actual executive performance (measured using *objective* computer tasks).

When providing self-reports, respondents may also be swayed by social desirability or experimenter demands. For example, in an ILM study, Wouters et al. (2016) found that high conscientiousness was tied to self-reported medication adherence but was unrelated to objectively monitored antidepressant use.

While researchers in other fields might be able to safely disregard the problems associated with self-report as random measurement bias, personality variables are more likely to be associated with differences between self-reported and objective measures, becoming actual confounds. While ILM studies looking only at the within-person level are less susceptible to self-report bias at the between-person level (as such variance if factored out), ILMs are often used to examine between-person phenomena (e.g. between-person variables moderating within-person associations); additionally, self-report bias can be associated with within-person level variables. For example, a researcher finding that state agreeableness is associated with more self-reported social interactions 3 hours later, might control for variables such as earlier social interactions and conclude that state agreeableness causes social interaction. However, measuring social interaction by other means (e.g. using a roommate's report) might reveal that agreeableness only causes participants to report more social interaction (e.g. because agreeableness is tied to more thorough responses to questionnaires; Bowling et al., 2016), whereas the objective number of interactions stays unchanged.

One way to overcome the difficulties attendant to self-reports involves technologies incorporating objective data, such as time, location, movement, or physiological input (e.g. heart rate; cf. Harari, Gosling, Wang & Campbell, 2015). For example, Eagle, Pentland, and Lazer (2009) measured proximity between participants during the day using mobile phone sensors and managed to detect over 95% of reported friendships between participants by examining patterns of proximity. Harari et al. (2017) used phone sensors to measure the presence (or absence) of human voices around participants, and the amount of movement (specifically, walking) during the day, as indicators of sociability and activity, respectively. They demonstrated changes in these measures across the school year, as activity gradually reduced during the first term, and sociability increased during the second term. Importantly, as such measures

require little to no input from participants, they can be activated continuously, up to 24 hours/day. ILM studies examining the associations between personality and sensor-recorded measures (e.g. physical activity) could help extend causal inferences beyond subjective perception.

With similar novel tools, researchers have also been able to capture subjective constructs without reliance on self-report. For example, Asselbergs et al. (2016) used sophisticated analytical methods to predict self-reported mood from smartphone activity logs (i.e. records of calls made, messages sent, apps used, and overall phone use duration). Place et al. (2017) were similarly able to predict symptoms related to posttraumatic stress disorder (depressive moods, interest in activities, avoidance of situations, and fatigue) as analysed by clinicians, using phone activity logs, location and movement data, and voice analysis of audio diaries completed by participants at least once a week. While more complicated to implement, ILMs can even include biological assays performed by participants, such as cortisol level tests. For example, Doane and Zeiders (2014) measured cortisol levels among adolescents five times a day, alongside more traditional self-report measures of stress. Nevertheless, researchers using these kinds of measures should make sure that the external measures are indeed related to the subjective phenomena under study, as such associations are not always straightforward (see for example the review by Campbell and Ehlert, 2012, on the imperfect association between cortisol level changes and subjective stress responses). We suggest that such measures should accompany, and not replace, self-reports.

Another way to mitigate problems associated with self-report is by using external observers. While having trained observers monitor participants in the time frame required by ILMs (e.g. once or even multiple times each day for extended periods of time) is likely to be unfeasible, reports from individuals who maintain close contact with participants (e.g. romantic partners, parents, colleagues, friends, or teachers) can complement self-report measures. For example, in an ILM study on personality disorder symptoms among romantic couples, South (2014) utilized partner reports to detect relationship problems that were not necessarily acknowledged by the respondents themselves. Such results increase our confidence that symptoms are associated not only with participants' subjective experience but also with some behavioural differences noticeable to their partners. Of course, as with biological markers, partner reports can themselves be biased and should be used alongside other sources of data (e.g. self-reports).

We used Dataset B to illustrate the use of external observers in going beyond self-reports; in this example, we also employed lagged analysis to achieve temporal precedence. Specifically, we examined the cross-lagged associations between one's partner's reports of providing emotional support and one's own next-day moods. Furthermore, we examined the extent to which this association is moderated by the support recipient's attachment orientation. Emotional and practical support receipt and provision were assessed using the daily support inventory (Bar-Kalifa & Rafaeli, 2013); moods were assessed using a revised daily

version of the profile of moods states (Cranford et al., 2006); and attachment orientation was assessed using the Experiences in Close Relationships revised scale (Brennan, Clark, & Shaver, 1998).

On the basis of the attachment literature (e.g. Simpson, Winterheld, Rholes, & Oriña, 2007), we expected recipients with more avoidant attachment to be characterized by more negative (or less positive) ties between support and mood. Importantly, the support and mood variables are obtained from different sources, thus reducing the risk of some individual-level third variable affecting both. To explore our hypothesis, we followed suggestions detailed in Bolger and Laurenceau (2013; ch. 8) for analysing causal process within distinguishable dyads. Specifically, we used multi-level models to predict participants' daily moods from their partners' previous-day reports of support provision (day-level variable), their attachment orientation (person-level variables), and the (cross-level) interaction between these. Importantly, we adjusted for the previous day's positive mood, which was entered as a covariate into the model. (Both the data and the SAS syntax are available at https://osf.io/k67ba/)

Overall, reports of partners' support provision were not tied to changes in moods the following day. Attachment avoidance ($b = -0.12$, $SE = 0.05$, $p = .008$, $\beta = .58$) was tied to less positive moods; the same was (marginally) true for anxiety ($b = -0.06$, $SE = 0.03$, $p = .073$, $\beta = .215$). Importantly, attachment avoidance moderated the association between support receipt and next-day positive mood ($b = -0.02$, $SE = 0.01$, $p = .026$, $\beta = .801$). Simple slope analyses revealed that individuals who were low in attachment avoidance (1 SD below the sample's mean) showed a positive association between support receipt and next-day positive mood ($b = 0.02$, $SE = 0.01$, $p = .035$, $\beta = .278$), whereas individuals who were moderate (at the sample mean) or high (1 SD above the sample's mean) in attachment avoidance did not show a significant association ($b = 0.004$, $SE = 0.006$, $p = .458$, $\beta = .028$ for moderate avoidance; $b = -0.01$, $SE = 0.01$, $p = .262$, $\beta = .031$ for high avoidance). In summary, with lagged daily partner reports, we were able to show that only nonavoidantly attached individuals appear to benefit emotionally from their partners' earlier emotional support. Dyadic lagged analyses of this sort help make the case that emotional support precedes—and may be a cause of—partners' positive moods (at least for those low in avoidance). The use of partner reports supports this rationale by allaying concerns regarding possible biases in support perception among avoidantly attached individuals (e.g. Barry, Lakey, & Orehek, 2007), although it also raises questions about possible biases in the *support-providing* reports of partners of avoidantly attached individuals.

BYPASSING SELF-REPORT BY DELVING DEEPER INTO REPEATEDLY MEASURED DATA

Adding objective measures or additional observers to overcome self-report bias might not be possible for certain variables or study designs or for assessing variables that

intrinsically involve meaning as part of their nature or that require privileged introspective access. Another strategy altogether to eliminate confounds that may arise when using self-report measures in personality research is to use the rich data generated by ILMs to look for patterns that go beyond respondents' direct awareness. In particular, some personality characteristics can be captured by estimating dynamic patterns in the data, which may include their variability, differentiation, range, temporal stability, inertia, congruence, and so on (e.g. Mejía, Hooker, Ram, Pham, & Metoyer, 2014; Trull, Lane, Koval, & Ebner-Priemer, 2015). Indices of these dynamic patterns may then be employed in several ways.

Certain dynamic parameters, such as variability, differentiation, or range, are *time independent* in that they can be applied to the raw data without requiring time or sequence to be taken into account. When these parameters have good test–retest reliability (as has been shown, for example, with regard to affect, Estabrook, Grimm, & Bowels, 2012; Mejía et al., 2014), they may reflect meaningful individual differences.

Early studies using this approach explored variability in self-esteem (for a review, see Kernis, 2003), which was found to range quite widely. Those with high but variable (as opposed to stable) self-esteem were found to use more self-protective and self-enhancing strategies and to display lower psychological adjustment. More recently, Zeigler-Hill et al. (2015) found variable self-esteem to be associated with low levels of emotional stability, agreeableness, and conscientiousness. Finally, Franck et al. (2016) demonstrated that pre-partum self-esteem variability served as a diathesis for post-partum depression among never-depressed women. Importantly, all of these studies referred to the construct at hand as 'instability' rather than variability; we return to this nomenclature issue shortly.

In other research demonstrating variability, Fleeson and colleagues (Fleeson, 2001; Fleeson & Gallagher, 2009; Fleeson & Law, 2015) repeatedly measured Big-Five states in individuals' daily lives and found the within-person variability in these states to be reliable and high (see also Baird, Le, & Lucas, 2006). Similarly, Moskovitz and Zuroff (2004) asked participants to monitor their own day-to-day behaviour during social interactions along the orthogonal axes of communion and agency. They found that individual differences in variability [both unidimensional (flux) and bidimensional (pulse and spin)] were temporally stable and reliably associated with other personality traits (e.g. neuroticism and extraversion).

Beyond variability, time-independent parameters also include ones reflecting within-person covariation patterns among items that (purportedly) tap the same construct (granularity or differentiation; Kashdan, Barrett, & McKnight, 2015; Erbas, Ceulemans, Lee, Koval, & Kuppers, 2014), as well as within-person covariation between constructs (synchrony or polarization: Coifman, Berenson, Rafaeli, & Downey, 2012; Rafaeli, Rogers, & Revelle, 2007). Importantly, these parameters can be examined vis-à-vis contextual factors; for example Coifman and her colleagues demonstrated that positive and negative affective and relational experiences become more polarized in moments characterized by interpersonal stress—particularly among individuals with borderline personality disorder (BPD).

Individual differences in *time-dependent* parameters of particular constructs may also provide researchers with vital information regarding individuals' personality. One such parameter that has been studied extensively in recent years, especially with regard to affective states, is within-person lability or instability. This construct, which has previously been studied using explicit self-reports (e.g. the Affect Lability Scale; P. D. Harvey, Greenberg, & Serper, 1989), is best assessed by measuring temporal fluctuations as they occur in respondents' day-to-day life (most commonly, using mean-squared successive differences; e.g. Trull et al., 2008). A recent meta-analysis of studies assessing affective instability (Houben, Van Den Noortgate, & Kuppens, 2015) reported a robust negative association between such instability and psychological well-being. Interestingly, the same meta-analysis also reported adverse associations for time-sensitive inertia (i.e. autocorrelation) and for time-insensitive variability.

Instability has also been examined in other, nonaffective, constructs. Beyond work examining self-esteem *variability* (e.g. Franck et al., 2016, cited earlier), recent research has also looked at time-sensitive *(in)stability* in this construct. For example, Farmer and Kashdan (2014) have found higher probability for acute changes and greater instability in self-esteem among individuals with social anxiety disorder in comparison with healthy controls (although the latter effect disappeared when adjusting for mean-level self-esteem). Similarly, Steger and Kashdan (2013) examined instability in individuals' sense of meaning in life. They found that greater instability was negatively associated with various well-being outcomes. Patterns of instability have also been found to differentiate between individuals suffering from personality and mood disorders (Mneimne, Fleeson, Arnold, & Furr, 2017).

Intensive longitudinal methods can also be employed for dyadic research (e.g. romantic partners, close friends, therapist–patient, and mother/father–child) in which they can reveal interesting and unique patterns of *covariation between dyad members*, which may reflect meaningful individual differences in personality characteristics. One such example is individuals' accuracy in assessing their partners' mental states (i.e. empathic accuracy; Ickes & Hodges, 2013). ILM-derived indices of empathic accuracy (e.g. Howland & Rafaeli, 2010; Overall, Fletcher, Simpson & Fillo, 2015), based on comparisons between dyad members' daily experienced affect and its perception, are garnering increased attention. Interestingly, perceivers' self-reported interpersonal sensitivity has been found to be unrelated to real accuracy (Ickes & Hodges, 2013).

Other examples of indices based on dyadic covariation involve the congruence between dyad members' mental states. For example, high congruence of one's daily goals with those of a romantic partner (e.g. Gere, Schimmack, Pinkus, & Lockwood, 2011) may reflect increased flexibility or a tendency to please. In contrast, high congruence between one's positive affect and that of a partner may serve as a compelling window into individual differences in capitalization or savouring (e.g. Gable, Reis, Impett, & Asher, 2004).

Notably, assessing personality-related variables (whether individual or dyadic) using dynamic data patterns requires researchers to address several principal issues. First, an appropriate index, accurately representing the theoretical construct of interest, has to be chosen; for example, *change resistance*, captured by an inertia index, is not equivalent to *stability* (e.g. Jahng, Wood, & Trull, 2008). Second, appropriate between-measurement time lags need to be determined; for example, day-to-day fluctuations are not the same as hour-to-hour fluctuations (for a review, see Hollenstein, Lichtwarck-Aschoff, & Potworowski, 2013). Third, when examining complex indices' associations with other constructs, proper adjustment for more basic constituents or features of the variables is needed; for example, studies of variability or instability in particular constructs often require adjustment for the constructs' mean (e.g. Trull et al., 2015).

To conclude, ILMs allow researchers to produce a variety of indices that go beyond simple self-reports to reflect, in a valid manner, patterns of change or covariation manifested in individuals' real life. Notably, this approach has been applied mostly to affective and to Big-Five personality states; of course, it can be applied to a much wider range of personality features including motivation and cognition. In this regard, initial work regarding attachment-style variability (Haak, Keller, & Dewall, 2016), characteristics of encountered situations (Jones, Brown, Serfass, & Sherman, 2017), and narrative features (McLean, Pasupathi, Greenhoot, & Fivush 2017) appears promising. So do recent studies exploring sophisticated within-person patterns such as diversity (e.g. Benson, Ram, Almeida, Zautra, & Ong, 2017) and flexibility (e.g. Hollenstein, 2015).

We used Dataset B to illustrate how dynamic patterns found in repeated measurements can capture meaningful individual differences. Specifically, we examined the protective role of the differentiation among negative relationship feelings (NRFs) in conflict days. Differentiation between affective states reflects individuals' tendency to experience (and report) their emotions in high degree of complexity. Recent findings indicate that in the face of stress, individuals characterized by higher levels of negative emotion differentiation are more resilient (for a review, see Kashdan et al., 2015).

Our investigation focused on a novel kind of affective differentiation, namely, the one among a romantic partner's NRFs. To obtain indices of differentiation, we calculated the average inter-item correlation of same-valenced items. The absolute values of average inter-item correlations range between 0 and 1, with higher scores representing lower differentiation. For the ease of results' interpretation, we reversed their scores.

On the basis of the literature in the field (e.g. Pond et al., 2012), we expected that higher NRF differentiation will be associated with smaller declines in intimacy on conflict days. We used multilevel models to predict participants' daily intimacy from their conflict reports (day-level variable), their NRF differentiation indices (person-level variable), and the (cross-level) interaction between the two. Importantly, we adjusted for mean levels of NRFs as well as for their interaction with conflict. Finally, because gender differences were found, we ran the models for men and women separately

using dummy variables. (The data, full results, and SAS syntax are available at https://osf.io/k67ba/)

We found that conflict days were indeed characterized by lower intimacy for both men ($b = -0.38$, $SE = 0.07$, $p < .001$, $\beta > .999$) and women ($b = -0.42$, $SE = 0.08$, $p < .001$, $\beta > .999$). Furthermore, NRF differentiation was not tied to intimacy for either men or women. Finally, and partially in line with our expectation, we found that women's ($b = 0.75$, $SE = 0.26$, $p = .006$, $\beta = .959$) but not men's ($b = 0.28$, $SE = 0.27$, $p = .289$, $\beta = .298$) NRF differentiation moderated conflict's effect on intimacy. Specifically, conflict had a weaker association with intimacy among women high in NRF differentiation ($b = -0.22$, $SE = 0.10$, $p = .036$, $\beta > .999$) than among those low in NRF differentiation ($b = -0.61$, $SE = 0.10$, $p < .001$, $\beta > .999$). These analyses illustrate how indices (e.g. differentiation) derived from repeatedly assessed data without direct reliance on self-reports help make a stronger case for the validity of the measure as a causal agent.

OVERCOMING PITFALLS IN CAUSAL ANALYSIS OF EXTERNAL CONTEXTUAL FACTORS

By looking into participants' lives over a period of time, ILMs can help identify the influence of different contextual conditions on the extent to which the behavioural, cognitive, affective, and motivational aspects of certain traits are enacted. In this vein, a classic study by Bolger and Schilling (1991) measured participants' neuroticism and then asked them to complete questionnaires every day for 6 weeks, reporting on both external stressors and mood. Neuroticism was linked both to experiencing more stressors and to greater stress reactivity following such stressors, with reactivity having a larger influence on total distress.

External factors are likely to exert particularly strong effects on (transient) personality states. For example, Fournier, Moskowitz, and Zuroff (2008) used ILMs to look at the expression of personality states in various interpersonal situations (in which respondents' interaction partners varied in agency and communion). They found both nomothetic associations and idiographic ones. For example, in one nomothetic finding, participants reported being in more agreeable personality states themselves when interacting with an agreeable person. At the same time, certain participants showed distinct idiographic patterns that went beyond the nomothetic effects (e.g. more submissiveness when interacting with an agreeable–submissive person than when interacting with a quarrelsome–submissive one).

In a similar manner, Geukes, Nestler, Hutteman, Küfner, and Back (2016) demonstrated both nomothetic and idiographic associations between specific contexts (i.e. the current location: at work, among friends, and at home) and variability in personality states, depending on personality traits. For example, neuroticism was associated with higher variability within contexts in expressive behaviours, but with lower such variability across contexts. Sherman et al. (2015) found that dominance behaviour was associated with trait extraversion and also with the degree of adversity in specific

situations. The association between dominance and adversity itself differed among participants.

Most of the studies described above did not explicitly attempt to establish causality. Indeed, inferring causality from associations between situations and personality states (e.g. Sherman et al., 2015) or from the moderating effect of personality traits on associations between situations and other outcomes (e.g. Bolger & Schilling, 1991) can present unique challenges. There is ample evidence that personality may influence the situations in which people find themselves, either through situation selection (cf. Gross & Thompson, 2007; Ickes, Synder & Garcia, 1997) or through situation interpretation (Allport, 1961; Gross & Thompson, 2007; Rauthmann, 2016). Moreover, 'third variables' (e.g. culture; Ching et al., 2014) may drive both situations and personality in tandem, leading to spurious (rather than causal) associations.

Three main strategies may be employed to deal with these problems. One strategy examines variables that cannot depend on participants' personalities. These may involve external factors such as time (e.g. the day of the week and the comparison of weekdays to weekends), the weather, or community-wide events (e.g. holidays, natural disasters, or elections). Alternatively, researchers may recruit participants who are all scheduled to experience some event. For example, Bolger, Zuckerman and Kessler (2000) asked law students and their romantic partners to complete daily questionnaires for 32 days before and 3 days after they took the New York State Bar Examination, a test considered stressful.

A second strategy involves the use of specific statistical methods that help bolster the case for inferring causality from ILM data. One such procedure adjusts for the effects of personality on external situations. For example, Bolger and Schilling (1991) calculated the amount of variance in negative mood accounted for by stressors, while holding the total (individual differences in) stressor levels constant. Another procedure involves person-centring the situation-level variable; such centring has the effect of removing any trait-level variance in the variable, leaving only situation-level variance to be explained.

A third strategy capitalizes on the sequential order of the repeated measures, by looking at outcome variables before, during, and after the occurrence of the situation.[2] In a way, this method treats the fluctuation of external stressors along the diary period as a natural ABA experiment (Barlow, Nock, & Hersen, 2006), in which participants begin from a baseline and then encounter a naturally occurring stressor (the 'treatment'), which later recedes or terminates.

We used Dataset C to demonstrate this third strategy. To do so, we examined the effects of disturbed sleep on new parents' negative moods, as well as the possible buffering role played by trait hope against such effects. Lack of sleep is a common problem for new parents, which has been associated with depressive symptoms for mothers (for a review, see Ross, Murray, & Steiner, 2005). Hope as a personality

trait has been shown to serve as a buffer against the effects of various stressors (Snyder, 2000), and thus, we hypothesized that parents with high trait hope would have less negative moods after sleepless nights.

We assessed hope using the trait hope questionnaire (Snyder et al., 1991). Negative moods were assessed using an adapted Profile of Mood States questionnaire (Cranford et al., 2006). Parents were asked how many times they woke up on the night before completing each questionnaire. Following the centring procedure recommended by Bolger and Laurenceau (2013), all IVs were person-centred to statistically remove possible effects of person-level variance in the trait in question (hope) on other IVs (e.g. effects of hope on times woken up).

We performed a mixed-model analysis similar to the analyses presented in previous sections; as the experiences of mothers and fathers are likely to be quite different, we obtained separate estimates for women and men by using a two-intercept model (Kenny, Kashy & Cook, 2006). We examined the association between disturbed sleep, measured by self-reported number of times woken up, and next-day negative mood. As before, we adjusted for lagged negative moods by including them as a covariate. We also included the number of times woken up on the following day to help establish the directionality of the association.

Results showed that men (but not women) with higher trait hope had less negative moods ($b = -0.24$, $SE = 0.08$, $p = .002$, $\beta = .62$). Disturbed sleep on the previous or next day was not associated with negative moods for either men or women. Trait hope moderated women's association between disturbed sleep and negative moods on the next day ($b = -0.07$, $SE = 0.02$, $p < .001$, $\beta = .137$), confirming our hypothesis for women, but the moderation was not found for men. Simple slope analyses revealed that for women who were low in trait hope, disturbed sleep *was* positively associated with next-day negative mood ($b = 0.04$, $SE = 0.01$, $p < .001$, $\beta > .999$), whereas for women who were moderate (at the sample mean) or high in trait hope, disturbed sleep *was not* associated with next-day negative mood ($b = 0.01$, $SE = 0.01$, $p = .271$, $\beta = .827$ for moderate hope; $b = -0.02$, $SE = 0.01$, $p = .064$, $\beta > .999$ for high hope). The moderation by hope of the association between disturbed sleep and negative mood on the previous day was significant for women ($b = -0.04$, $SE = 0.02$, $p = .035$, $\beta = .085$), suggesting that beyond the hypothesized effect there might also be a reversed causal effect (i.e. negative mood affecting sleep) or a third variable affecting both. (The data, full results, and the SAS syntax are available at https://osf.io/k67ba/)

INTEGRATING EXPERIMENTAL ELEMENTS INTO INTENSIVE LONGITUDINAL METHOD DESIGNS

One of the classic ways to demonstrate causality in psychological (and other) research is through randomized experiments. Importantly, ILM studies can be used in tandem with experimental methods to answer new research questions in at least four ways.

---

[2]We made mention of this procedure earlier [in our discussion of temporal precedence; see Section on on Using Temporal Precedence to Establish Causality (Lagged Analyses, Time as a Covariate)] but will go into more detail now on how it may be applied when external situations are of interest.

One way involves the administration of ILMs before performing a manipulation (to improve assessment of the participants' baseline), after it (to improve assessment of long-term, ecological effects), or at both times. This method has been used extensively in psychotherapy studies. For example, Borkovec and Costello (1993) asked participants to complete daily diaries while receiving one of three types of therapy for generalized anxiety disorder, but also for two 2-week periods before and after the therapy, and again for 1-week periods at the 6- and 12-month follow-up points. Using ILMs enabled the researchers to assess daily life symptoms such as anxiety episodes as they happened, both before therapy (as a measure of severity) and after therapy (as a measure of improvement).

Another way of integrating experimental methods into ILM studies involves incorporating an experimental manipulation into the ILMs themselves. In such studies, different participants can be provided with different daily or momentary instructions. Finkel, Slotter, Luchies, Walton, and Gross (2013) demonstrated the effects of one such manipulation in a repeated-measures long-term study. Married couples in this study completed Internet questionnaires at fixed intervals (every 4 months); midway through the study, some couples were asked—as part of the questionnaire—to think of a previous disagreement with their spouse and to try to reappraise their perception of it. Whereas the relationship quality of couples in a control condition declined over time, that of couples who received this manipulation did not.

Chapman, Rosenthal and Leung (2009) demonstrated a similar technique in a short ILM study. The researchers asked participants with high or low BPD symptoms to complete measures eight times a day for 4 days, while using a different emotion regulation technique each day. Interestingly, and contrary to more classic experimental work on the subject, individuals high in BPD showed increased *positive* emotions and reduced impulsive behaviour on days in which they were asked to suppress emotions, when compared with a baseline day in which they were not given special instructions; conversely, individuals low in BPD had the opposite experience (increased *negative* emotions on days in which they were asked to suppress emotions, compared with baseline).

A third (and related) way to integrate experimental methods into ILMs makes use of information collected within the diary by tying it into the experimental manipulation. For example, Fisher (2015) used the results of a preliminary ILM assessment to tailor psychotherapy interventions. Going one step further, manipulations could be tailored during the ILM period to specific situations. For example, participants could be asked to perform certain coping techniques in response to actual stressors reported in the previous ILM data point. These kinds of interventions have been used in recent years in health research (e.g. Heron & Smyth, 2010; Riley et al., 2011) and are often termed just-in-time adaptive interventions. Klasnja et al. (2015) propose a specific technique to evaluate causality in JITAIs called micro-randomization. Essentially, they suggest that participants can be randomized to different interventions at different points during their ILM period. To illustrate, they discuss a cardiac health study in which participants were randomly assigned twice a day to receive one of two prompts (one suggesting that they take a walk and the other that they move around the room). This technique allows for high efficiency as each participant is randomized multiple times to various conditions.

Finally, ILMs can be combined with experimental methods indirectly, with processes identified by ILMs verified independently in experimental studies. For example, previous ILM research had documented associations between neuroticism and negative affect and between extraversion and positive affect (e.g. David, Green, Martin, & Suls, 1997). McNiel and Fleeson (2006) built on this work and used experimental methods in which state neuroticism and extraversion were manipulated to demonstrate causality in these associations. Similarly, McCabe and Fleeson (2016) examined the role of goals in enactments of conscientiousness and extraversion, first demonstrating an association by using ILMs (Study 1) and then manipulating goals in a laboratory experiment to show that different goals caused different levels of enactment (Study 2).

DESIGNING INTENSIVE LONGITUDINAL METHODS FOR CAUSAL INFERENCES: A METHOD CHECKLIST

To summarize the five issues presented above, we offer a short checklist that may help researchers who wish to use ILM studies to examine causality in the study of personality (see Figure 1). This checklist encompasses decisions that can be made in the design of the study procedures, in the choice of measures, and in the selection of analytic models.

Procedure choices

When researchers begin to plan a new ILM study, some early fundamental choices can have great effects on their ability to draw causal inferences from the obtained data. First, researchers should consider the length of effects they expect to find. A lag that is too long for a fast-occurring process is likely not to reveal an actually existing causal effect. For example, the effects of stressful events on stress might be immediate, and a 3-hour lag, especially if controlling for previous stress, may fail to detect this immediate lag. Oftentimes, lags will be too long in ILMs, making it difficult to detect causal effects. However, if researchers expect a large enough delay between the occurrence of causal variables and their outcomes, they should consider planning the intervals between measurements such that causes and outcomes will occur on separate measurements. For example, if optimists are expected to recover from work-related stress in the morning on the same evening, while pessimists are expected to be affected for a few days, stress should be measured both in the morning and in the evening. Asking participants in the evening about stress in the morning might lead to reverse causal effects.[3]

[3]See Section on Using Temporal Precedence to Establish Causality (Lagged Analyses, Time as a Covariate).
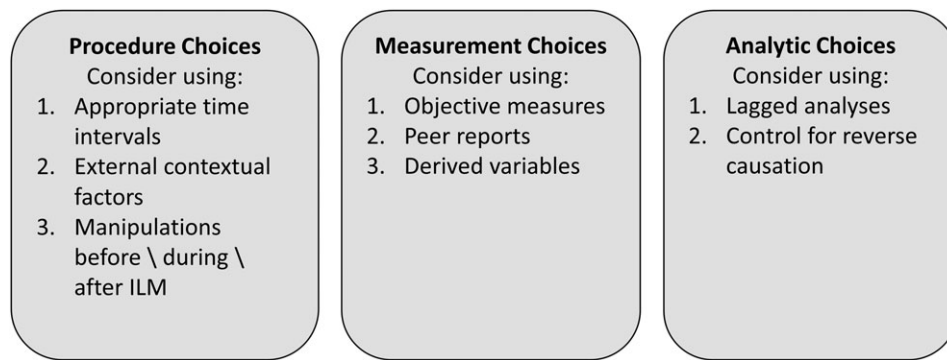
Figure 1.   A basic checklist for strengthening causality in personality intensive longitudinal methods studies

Second, researchers should consider the impact of external contextual factors on the phenomenon they are researching. External context can be utilized to observe changes in a causal variable while ruling out alternative explanations. For example, designing an ILM around a major external stressor (e.g. with students during midterm exams) can enable researchers to argue that any outcomes of the specific stressor period were likely *caused* by the occurrence of the stressor. Other alternatives are using a recurring external condition (e.g. weather patterns, holidays). A somewhat less powerful application of external context is asking participants to self-report contextual factors (e.g. ask them whether they had health-related problems on a specific day). This should be carried out carefully, as it might introduce complications. Importantly, even if not purposefully included in a study, external context can introduce confounding effect; researchers should plan to have some indication of context insofar as they expect it to influence study variables.[4]

Finally, researchers should consider including experimentally manipulating key study variables before the ILM phase (using the ILM to observe outcomes), after the ILM phase (using the ILM to observe the initial context), or during the ILM phase to do both. Another possibility is to use micro-randomization (Klasnja et al., 2015), essentially manipulating variables on a per-measurement basis, to examine manipulation effects continuously over the study period.[5]

Measurement choices

Beyond general procedure, researchers can use alternative measuring methods to reduce self-report biases. First, researchers should consider including objective measures (e.g. location and phone call records), either for key study variables or on an opportunistic basis when the technology allows. For example, most software used in ILMs allow researchers to easily measure participants' response time to various questions, which can help assess participants' investment in answering specific questions or reading study instructions.[6]

Second, researchers should consider including reports from peers. In close relationship studies, this is often as simple as asking each participant to report on themselves and on their relationship partner; other studies could consider recruiting parents, teachers, roommates, and so on.[6]

Finally, researchers should consider using derived indices such as congruence and variance out of self-report data. Such indices are relatively easy to obtain, as they do not require specific technology or involve other participants, but they can still avoid some of the potential problems of complete reliance on self-report.[7]

Importantly, the choice between various methods is not exclusive, and a study combining, for example, self-report, peer report, and objective measures of a key variable can use these different perspectives to sort out possible confounds inherent to each one of them when taken separately.

Analysis choices

When analysing the data, researchers should take care to employ statistical procedures that can take into account the unique characteristics of ILM, such as analysis of lagged data and time course modelling[8], and proper statistical adjustment for alternative causal pathways and for external contextual factors.[4] Importantly, statistical analysis of ILM data is under constant development, and researchers should make sure to stay up to date with recent developments.

BROADER DISCUSSION

In the remainder of this article, we will argue that ILMs offer a useful methodological bridge between trait and social–cognitive approaches to the study of personality. We will also discuss some limitations of ILMs and offer concluding remarks.

[4]See Section on Overcoming Pitfalls in Causal Analysis of External Contextual Factors.
[5]See Section on Integrating Experimental Elements into Intensive Longitudinal Method Designs.
[6]See Section on Eliminating Confounds by Going beyond Self-report.

[7]See Section on Bypassing Self-report by Delving Deeper into Repeatedly Measured Data.
[8]See Section on Using Temporal Precedence to Establish Causality (Lagged Analyses, Time as a Covariate).

Integrating social–cognitive and trait approaches

Use of the methods described above, which support personality-related causal inferences, can be viewed as a part of an ongoing effort to integrate social–cognitive and trait approaches to personality. Historically, social–cognitive approaches (e.g. Cervone & Shoda, 1999) have focused on studying dynamic mental representations, information processing mechanisms, motivational processes, and situational factors to explain relatively narrow slices of behaviours. Such approaches considered within-person cross-situational consistency as rather low. In contrast, trait approaches had focused on describing between-person structural differences in relatively broad behavioural characteristics and considered within-person cross-situational consistency as rather high.

Recently, models integrating the two broad approaches have been proposed by one of us (Fleeson & Jayawickreme, 2015), as well as by others (e.g. Read et al., 2010; Revelle & Condon, 2015; Wright, 2014). Although their specifics are beyond the scope of the present article, it is notable that they all share the goal of explaining between-person personality structure and its state-like behavioural manifestations in terms of dynamic cognitive, affective, and motivational processes. As our review indicates, ILMs can be instrumental in realizing this goal. Specifically, ILMs allow for the identification of various dynamic processes as they occur in individuals' daily lives and help monitor the presence of important situational factors affecting these processes. The continuous nature of ILMs makes the exploration of these processes' behavioural, cognitive, and affective consequences—that is, of manifestations of personality states—possible. Finally, differential within-person patterns may represent differences in between-individual personality traits.

Limitations of intensive longitudinal methods

While we have outlined many benefits of using ILMs, there are also some drawbacks. First, ILM studies can be demanding on both participants' and researchers' resources, although recent guides and software advances make them less and less so. Participants need to be committed to completing measures intensively, following a highly specific schedule, for a long period of time; these measures, while relatively brief, might take up quite some time and effort. Researchers need to design complex procedures, to be available to help participants who might run into difficulties at any point during the study period, to be able to communicate this help to the participants who are typically off-site, and to offer participants adequate compensation, which tends to be greater than in other studies. However, some of these issues arise from our own desires as researchers to collect as much information as possible in each ILM study—we recommend reducing the number of variables assessed per study.

Second, intensive measurement of variables raises issues of reactivity. In a review by Barta, Tennen, and Litt (2012), the authors identify various processes that might lead to reactivity. Specifically, they detail two sources of reactivity that might be especially associated with ILMs: satisficing (participants' tendency to complete measures quickly while paying little attention due to time limits or fatigue) and self-monitoring (participants' excessive attention and focus on the variables measured). Barta and his colleagues suggest various methods to counter these effects. For example, satisficing can be limited by making sure that the measures are not too long or taxing. Self-monitoring can be reduced by preventing participants' ability to review previous entries and by including multiple variables. An alternative method to reduce reactivity problems can be relying less on self-report, by using the methods detailed in previous sections.[9]

Finally, one of the main advantages of ILMs—ecological validity—can also pose problems. ILM studies' participants complete measures while going about their daily life and not under direct supervision by researchers. Consequently, they may fail to follow study instructions for various quotidian reasons. While some aspects of adherence can be monitored (e.g. electronically recording time of completion), it is often unclear how to handle slight deviances from the stated procedure (e.g. how long after a measure was supposed to be completed should it be disregarded?). We encourage researchers to ensure that participants are well trained in the study procedures and to monitor their adherence.

Conclusion

The current article presented various strategies for improving the ability to infer causality in ILM personality studies. We detailed the implications of time series for causality, the use of objective measures and of measures derived indirectly from self-report data, the impact of external factors, and the possibilities of integrating manipulations with ILMs. Importantly, whereas each of these strategies, in isolation, may help researchers draw causal inferences, they can also build upon one another; for example, objective measures such as location or weather can be used to trigger specific manipulations; similarly, external stressors can be measured using peer reports to reduce the probability of reverse causation. We hope the methods presented can help researchers enjoy the advantages of ILMs and improve their capacity to examine causal hypotheses.

[9]Sections on Eliminating Confounds by Going Beyond Self-Report and Bypassing Self-report by Delving into Repeatedly Measured Data.

REFERENCES

Allport, G. W. (1961). *Pattern and growth in personality*. Oxford, UK: Holt, Reinhart & Winston.

Asselbergs, J., Ruwaard, J., Ejdys, M., Schrader, N., Sijbrandij, M., & Riper, H. (2016). Mobile phone-based unobtrusive ecological momentary assessment of day-to-day mood: An explorative study. *Journal of Medical Internet Research*, *18*. https://doi.org/10.2196/jmir.5505.

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. https://doi.org/10.1037/0022-3514.90.3.512.

Bar-Kalifa, E., & Rafaeli, E. (2013). Disappointment's sting is greater than help's balm: Quasi-signal detection of daily support matching. *Journal of Family Psychology*, *27*, 956–967. https://doi.org/10.1037/a0034905.

Barlow, D. H., Nock, M., & Hersen, M. (2006). *Single case research designs: Strategies for studying behavior change*. New York, NY: Allyn & Bacon.

Barry, R. A., Lakey, B., & Orehek, E. (2007). Links among attachment dimensions, affect, the self, and perceived support for broadly generalized attachment styles and specific bonds. *Personality and Social Psychology Bulletin*, *33*, 340–353. https://doi.org/10.1177/0146167206296102.

Barta, W. D., Tennen, H., & Litt, M. D. (2012). Measurement reactivity in diary research. In M. R. Mehl, & T. S. Conner (Eds.), *Handbook of research methods for studying daily life* (pp. 108–123). London, UK: Guilford Press.

Benson, L., Ram, N., Almeida, D. M., Zautra, A. J., & Ong, A. D. (2017). Fusing biodiversity metrics into investigations of daily life: Illustrations and recommendations with emodiversity. *The Journals of Gerontology: Series B*. https://doi.org/10.1093/geronb/gbx025.

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

Bolger, N., & Laurenceau, J. P. (2013). *Intensive longitudinal methods: An introduction to diary and experience sampling research*. New York, NY: Guilford.

Bolger, N., & Schilling, E. A. (1991). Personality and the problems of everyday life: The role of neuroticism in exposure and reactivity to daily stressors. *Journal of Personality*, *59*, 355–386. https://doi.org/10.1111/j.1467-6494.1991.tb00253.x.

Bolger, N., Zuckerman, A., & Kessler, R. C. (2000). Invisible support and adjustment to stress. *Journal of Personality and Social Psychology*, *79*, 953–961. https://doi.org/10.1037/0022-3514.79.6.953.

Borkovec, T. D., & Costello, E. (1993). Efficacy of applied relaxation and cognitive–behavioral therapy in the treatment of generalized anxiety disorder. *Journal of Consulting and Clinical Psychology*, *61*, 611–619. https://doi.org/10.1037/0022-006X.61.4.611.

Borsboom, D., & Cramer, A. O. (2013). Network analysis: An integrative approach to the structure of psychopathology. *Annual Review of Clinical Psychology*, *9*, 91–121. https://doi.org/10.1146/annurev-clinpsy-050212-185608.

Borsboom, D., Mellenbergh, G. J., & Van Heerden, J. (2003). The theoretical status of latent variables. *Psychological Review*, *110*, 203–219. https://doi.org/10.1037/0033-295X.110.2.203.

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*, 175–177. https://doi.org/10.1159/000453583.

Bowling, N. A., Huang, J. L., Bragg, C. B., Khazon, S., Liu, M., & Blackmore, C. E. (2016). Who cares and who is careless? Insufficient effort responding as a reflection of respondent personality.

Journal of Personality and Social Psychology, *111*, 218–229. https://doi.org/10.1037/pspp0000085.

Brennan, K. A., Clark, C. L., & Shaver, P. R. (1998). Self-report measurement of adult attachment: An integrative overview. In J. A. Simpson, & W. S. Rholes (Eds.), *Attachment theory and close relationships* (pp. 46–76). New York, NY: Guilford Press.

Bromberg, P. M. (1996). Standing in the spaces: The multiplicity of self and the psychoanalytic relationship. *Contemporary Psychoanalysis*, *32*, 509–535. https://doi.org/10.1080/00107530.1996.10746334.

Buchanan, T. (2016). Self-report measures of executive function problems correlate with personality, not performance-based executive function measures, in nonclinical samples. *Psychological Assessment*, *28*, 372–385. https://doi.org/10.1037/pas0000192.

Campbell, J., & Ehlert, U. (2012). Acute psychosocial stress: Does the emotional stress response correspond with physiological responses? *Psychoneuroendocrinology*, *37*, 1111–1134. https://doi.org/10.1016/j.psyneuen.2011.12.010.

Cervone, D. (2005). Personality architecture: Within-person structures and processes. *Annual Review of Psychology*, *56*, 423–452. https://doi.org/10.1146/annurev.psych.56.091103.070133.

Cervone, D., & Shoda, Y. (Eds) (1999). *The coherence of personality: Social–cognitive bases of consistency, variability, and organization*. New York, NY: Guilford Press.

Chapman, A. L., Rosenthal, M. Z., & Leung, D. W. (2009). Emotion suppression in borderline personality disorder: An experience sampling study. *Journal of Personality Disorders*, *23*, 29–47. https://doi.org/10.1521/pedi.2009.23.1.29.

Ching, C. M., Church, A. T., Katigbak, M. S., Reyes, J. A. S., Tanaka-Matsumi, J., Takaoka, S., Zhang, H., … Ortiz, F. A. (2014). The manifestation of traits in everyday behavior and affect: A five-culture study. *Journal of Research in Personality*, *48*, 1–16. https://doi.org/10.1016/j.jrp.2013.10.002.

Cofta-Woerpel, L., McClure, J. B., Li, Y., Urbauer, D., Cinciripini, P. M., & Wetter, D. W. (2011). Early cessation success or failure among women attempting to quit smoking: Trajectories and volatility of urge and negative mood during the first post cessation week. *Journal of Abnormal Psychology*, *120*, 596–606. https://doi.org/10.1037/a0023755.

Coifman, K. G., Berenson, K. R., Rafaeli, E., & Downey, G. (2012). From negative to positive and back again: Polarized affective and relational experience in borderline personality disorder. *Journal of Abnormal Psychology*, *121*, 668–679. https://doi.org/10.1037/a0028502.

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*, 292–313. https://doi.org/10.1111/j.1751-9004.2009.00170.x.

Conway, J. M., Rogelberg, S. G., & Pitts, V. E. (2009). Workplace helping: Interactive effects of personality and momentary positive affect. *Human Performance*, *22*, 321–339. https://doi.org/10.1080/08959280903120279.

Constantini, G., Epskamp, S., Borsboom, D., Perugini, M., Mõttus, R., Waldorp, L. J., & Cramer, A. O. (2015). State of the aRt personality research: A tutorial on network analysis of personality data in R. *Journal of Research in Personality*, *54*, 13–29. https://doi.org/10.1016/j.jrp.2014.07.003.

Cranford, J. A., Patrick, E. S., Masumi, I., Rafaeli, E., Yip, T., & Bolger, N. (2006). A procedure for evaluating sensitivity to within-person change: Can mood measures in diary studies detect change reliably? *Personality and Social Psychology Bulletin*, *32*, 917–929. https://doi.org/10.1177/0146167206287721.

Csikszentmihalyi, M., & Larson, R. (2014). Validity and reliability of the experience-sampling method. In M. Csikszentmihalyi (Ed.), *Flow and the foundations of positive psychology* (pp (pp. 35–54). Dordrecht, The Netherlands: Springer Netherlands https://doi.org/10.1007/978-94-017-9088-8.

David, J. P., Green, P. J., Martin, R., & Suls, J. (1997). Differential roles of neuroticism, extraversion, and event desirability for mood in daily life: An integrative model of top-down and bottom-up influences. *Journal of Personality and Social Psychology*, *73*, 149–159. https://doi.org/10.1037/0022-3514.73.1.149.

de Haan-Rietdijk, S., Voelkle, M. C., Keijsers, L., & Hamaker, E. L. (2017). Discrete- vs. continuous-time modeling of unequally spaced experience sampling method data. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.01849.

Deboeck, P. R., & Preacher, K. J. (2016). No need to be discrete: A method for continuous time mediation analysis. *Structural Equation Modeling*, *23*, 61–75. https://doi.org/10.1080/10705511.2014.973960.

Doane, L. D., & Zeiders, K. H. (2014). Contextual moderators of momentary cortisol and negative affect in adolescents' daily lives. *Journal of Adolescent Health*, *54*, 536–542. https://doi.org/10.1016/j.jadohealth.2013.10.007.

Dormann, C., & Griffin, M. A. (2015). Optimal time lags in panel studies. *Psychological Methods*, *20*, 489–505. https://doi.org/10.1037/met0000041.

Eagle, N., Pentland, A. S., & Lazer, D. (2009). Inferring friendship network structure by using mobile phone data. *Proceedings of the National Academy of Sciences*, *106*, 15274–15278. https://doi.org/10.1073/pnas.0900282106.

Epskamp, S., Borsboom, D., & Fried, E. I. (2017). Estimating psychological networks and their accuracy: A tutorial paper. *Behavior Research Methods*, *50*, 1–18. https://doi.org/10.3758/s13428-017-0862-1.

Erbas, Y., Ceulemans, E., Lee Pe, M., Koval, P., & Kuppens, P. (2014). Negative emotion differentiation: Its personality and well-being correlates and a comparison of different assessment methods. *Cognition and Emotion*, *28*, 1196–1213. https://doi.org/10.1080/02699931.2013.875890.

Estabrook, R., Grimm, K. J., & Bowles, R. P. (2012). A Monte Carlo simulation study of the reliability of intraindividual variability. *Psychology and Aging*, *27*, 560–576. https://doi.org/10.1037/a0026669.

Fahrenberg, J., Myrtek, M., Pawlik, K., & Perrez, M. (2007). Ambulatory assessment-monitoring behavior in daily life settings. *European Journal of Psychological Assessment*, *23*, 206–213. https://doi.org/10.1027/1015-5759.23.4.206.

Falkenström, F., Finkel, S., Sandell, R., Rubel, J. A., & Holmqvist, R. (2017). Dynamic models of individual change in psychotherapy process research. *Journal of Consulting and Clinical Psychology*, *85*, 537–549. https://doi.org/10.1037/ccp0000203.

Farmer, A. S., & Kashdan, T. B. (2014). Affective and self-esteem instability in the daily lives of people with generalized social anxiety disorder. *Clinical Psychological Science*, *2*, 187–201. https://doi.org/10.1177/2167702613495200.

Farnworth, L., Mostert, E., Harrison, S., & Worrell, D. (1996). The experience sampling method: Its potential use in occupational therapy research. *Occupational Therapy International*, *3*, 1–17. https://doi.org/10.1002/oti.23.

Finkel, E. J., Slotter, E. B., Luchies, L. B., Walton, G. M., & Gross, J. J. (2013). A brief intervention to promote conflict reappraisal preserves marital quality over time. *Psychological Science*, *24*, 1595–1601. https://doi.org/10.1177/0956797612474938.

Finkel, S. E. (1995). *Causal analysis with panel data*. Thousand Oaks, CA: Sage https://doi.org/10.4135/9781412983594.

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

Fleeson, W. (2001). Toward a structure- and process-integrated view of personality: Traits as density distributions of states. *Journal of Personality and Social Psychology*, *80*, 1011–1027. https://doi.org/10.1037/0022-3514.80.6.1011.

Fleeson, W. (2017). The production mechanisms of traits: Reflections on two amazing decades. *Journal of Research in Personality*, *69*, 4–12. https://doi.org/10.1016/j.jrp.2017.07.003.

Fleeson, W., & Gallagher, P. (2009). The implications of Big Five standing for the distribution of trait manifestation in behavior: Fifteen experience-sampling studies and a meta-analysis. *Journal of Personality and Social Psychology*, *97*, 1097–1114. https://doi.org/10.1037/a0016786.

Fleeson, W., & Jayawickreme, E. (2015). Whole trait theory. *Journal of Research in Personality*, *56*, 82–92. https://doi.org/10.1016/j.jrp.2014.10.009.

Fleeson, W., & Law, M. K. (2015). Trait enactments as density distributions: The role of actors, situations, and observers in explaining stability and variability. *Journal of Personality and Social Psychology*, *109*, 1090–1104. https://doi.org/10.1037/a0039517.

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

Ford, M. B., & Collins, N. L. (2013). Self-esteem moderates the effects of daily rejection on health and well-being. *Self and Identity*, *12*, 16–38. https://doi.org/10.1080/15298868.2011.625647.

Fournier, M. A., Moskowitz, D. S., & Zuroff, D. C. (2008). Integrating dispositions, signatures, and the interpersonal domain. *Journal of Personality and Social Psychology*, *94*, 531–545. https://doi.org/10.1037/0022-3514.94.3.531.

Franck, E., Vanderhasselt, M. A., Goubert, L., Loeys, T., Temmerman, M., & De Raedt, R. (2016). The role of self-esteem instability in the development of postnatal depression: A prospective study testing a diathesis-stress account. *Journal of Behavior Therapy and Experimental Psychiatry*, *50*, 15–22. https://doi.org/10.1016/j.jbtep.2015.04.010.

Gable, S. L., Reis, H. T., Impett, E. A., & Asher, E. R. (2004). What do you do when things go right? The intrapersonal and interpersonal benefits of sharing positive events. *Journal of Personality and Social Psychology*, *87*, 228–245. https://doi.org/10.1037/0022-3514.87.2.228.

Gere, J., Schimmack, U., Pinkus, R. T., & Lockwood, P. (2011). The effects of romantic partners' goal congruence on affective well-being. *Journal of Research in Personality*, *45*, 549–559 https://doi.org/10.1016/j.jrp.2011.06.010.

Geukes, K., Nestler, S., Hutteman, R., Küfner, A. C., & Back, M. D. (2016). Trait personality and state variability: Predicting individual differences in within-and cross-context fluctuations in affect, self-evaluations, and behavior in everyday life. *Journal of Research in Personality*, online publication before print. https://doi.org/10.1016/j.jrp.2016.06.003, *69*, 124–138.

Gollob, H. F., & Reichardt, C. S. (1987). Taking account of time lags in causal models. *Child Development*, *58*, 80–92. https://doi.org/10.2307/1130293.

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

Granger, C. W. J., & Newbold, P. (1974). Spurious regressions in econometrics. *Journal of Econometrics*, *2*, 111–120. https://doi.org/10.1016/0304-4076(74)90034-7.

Greenberg, L. S. (2011). *Emotion-focused therapy*. Washington, DC: American Psychological Association.

Gross, J. J., & Thompson, R. A. (2007). Emotion regulation: Conceptual foundations. In J. J. Gross (Ed.), *Handbook of emotion regulation* (pp. 3–24). New York, NY: Guilford Press.

Haak, E. A., Keller, P. S., & DeWall, C. N. (2016). Daily variations in attachment anxiety and avoidance: A density distributions approach. *Journal of Research in Personality*. Advance online publication. https://doi.org/10.1016/j.jrp.2016.08.002, *69*, 218–224.

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

Hamaker, E. L. (2012). Why researchers should think "within-person": A paradigmatic rationale. In: M. R., Mehl, T. S., Conner

(Eds.), *Handbook of research methods for studying daily life* (pp. 43–61). New York, NY: Guilford Press.

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

Harari, G. M., Gosling, S. D., Wang, R., & Campbell, A. T. (2015). Capturing situational information with smartphones and mobile sensing methods. *European Journal of Personality*, 29, 509–511. https://doi.org/10.1002/per.2032.

Harari, G. M., Müller, S. R., Mishra, V., Wang, R., Campbell, A. T., Rentfrow, P. J., & Gosling, S. D. (2017). An evaluation of Students' interest in and compliance with self-tracking methods: Recommendations for incentives based on three smartphone sensing studies. *Social Psychological and Personality Science*, online publication before prin. https://doi.org/10.1177/1948550617712033.

Harré, R. (1998). *The singular self: An introduction to the psychology of personhood*. Thousand Oaks, CA: Sage.

Harvey, J. H., Hendrick, S. S., & Tucker, K. (1988). Self-report methods in studying personal relationships. In S. Duck, D. F. Hay, S. E. Hobfoll, W. Ickes, & B. M. Montgomery (Eds.), *Handbook of personal relationships: Theory, research and interventions* (pp. 99–113). Oxford, UK: John Wiley.

Harvey, P. D., Greenberg, B. R., & Serper, M. R. (1989). The affective lability scales: Development, reliability, and validity. *Journal of Clinical Psychology*, 45, 786–793. https://doi.org/10.1002/1097-4679(198909)45:5%3C786::AID-JCLP2270450515%3E3.0.CO;2-P.

Heron, K. E., & Smyth, J. M. (2010). Ecological momentary interventions: Incorporating mobile technology into psychosocial and health behaviour treatments. *British Journal of Health Psychology*, 15, 1–39. https://doi.org/10.1348/135910709X466063.

Hodgkinson, J., Mant, J., Martin, U., Guo, B., Hobbs, F. D. R., Deeks, J. J., Heneghan, C., … McManus, R. J. (2011). Relative effectiveness of clinic and home blood pressure monitoring compared with ambulatory blood pressure monitoring in diagnosis of hypertension: Systematic review. *BMJ*, 342, d3621. https://doi.org/10.1136/bmj.d3621.

Hollenstein, T. (2015). This time, it's real: Affective flexibility, time scales, feedback loops, and the regulation of emotion. *Emotion Review*, 7, 308–315. https://doi.org/10.1177/1754073915590621.

Hollenstein, T., Lichtwarck-Aschoff, A., & Potworowski, G. (2013). A model of socioemotional flexibility at three time scales. *Emotion Review*, 5, 397–405. https://doi.org/10.1177/1754073913484181.

Houben, M., Van Den Noortgate, W., & Kuppens, P. (2015). The relation between short-term emotion dynamics and psychological well-being: A meta-analysis. *Psychological Bulletin*, 141, 901–930. https://doi.org/10.1037/a0038822.

Howell, R. T., Ksendzova, M., Nestingen, E., Yerahian, C., & Iyer, R. (2017). Your personality on a good day: How trait and state personality predict daily wellbeing. *Journal of Research in Personality*, 69, 250–263. https://doi.org/10.1016/j.jrp.2016.08.001.

Howland, M., & Rafaeli, E. (2010). Bringing everyday mind reading into everyday life: Assessing empathic accuracy with daily diary data. *Journal of Personality*, 78, 1437–1468. https://doi.org/10.1111/j.1467-6494.2010.00657.x.

Hox, J. J., Moerbeek, M., & van de Schoot, R. (2010). *Multilevel analysis: Techniques and applications*. New York, NY: Routledge.

Ickes, W., & Hodges, S. D. (2013). Empathic accuracy in close relationships. In J. A. Simpson, & L. Campbell (Eds.), *The Oxford handbook of close relationships* (pp. 116–136). New York, NY: Oxford University Press.

Ickes, W., Snyder, M., & Garcia, S. (1997). Personality influences on the choice of situations. In R. Hogan, J. A. Johnson, & S. R. Briggs (Eds.), *Handbook of personality psychology*

(pp. 165–195). https://doi.org/10.1016/B978-012134645-4/50008-1.

Ilies, R., Johnson, M. D., Judge, T. A., & Keeney, J. (2011). A within-individual study of interpersonal conflict as a work stressor: Dispositional and situational moderators. *Journal of Organizational Behavior*, 32, 44–64. https://doi.org/10.1002/job.677.

Jahng, S., Wood, P. K., & Trull, T. J. (2008). Analysis of affective instability in ecological momentary assessment: Indices using successive difference and group comparison via multilevel modeling. *Psychological Methods*, 13, 354–375. https://doi.org/10.1037/a0014173.

Jones, A. B., Brown, N. A., Serfass, D. G., & Sherman, R. A. (2017). Personality and density distributions of behavior, emotions, and situations. *Journal of Research in Personality*, 69, 225–236. https://doi.org/10.1016/j.jrp.2016.10.006.

Kashdan, T. B., Barrett, L. F., & McKnight, P. E. (2015). Unpacking emotion differentiation: Transforming unpleasant experience by perceiving distinctions in negativity. *Current Directions in Psychological Science*, 24, 10–16. https://doi.org/10.1177/0963721414550708.

Kenny, D. A., Kashy, D. A., & Cook, W. L. (2006). *Dyadic data analysis*. New York, NY: Guilford Press.

Kernis, M. H. (2003). Toward a conceptualization of optimal self-esteem. *Psychological Inquiry*, 14, 1–26. https://doi.org/10.1207/S15327965PLI1401_01.

Klasnja, P., Hekler, E. B., Shiffman, S., Boruvka, A., Almirall, D., Tewari, A., & Murphy, S. A. (2015). Microrandomized trials: An experimental design for developing just-in-time adaptive interventions. *Health Psychology*, 34, 1220–1228. https://doi.org/10.1037/hea0000305.

Leikas, S., & Ilmarinen, V. J. (2016). Happy now, tired later? Extraverted and conscientious behavior are related to immediate mood gains, but to later fatigue. *Journal of Personality*, 85, 603–615. https://doi.org/10.1111/jopy.12264.

Lobbestael, J., van Vreeswijk, M., Spinhoven, P., Schouten, E., & Arntz, A. (2010). Reliability and validity of the short Schema Mode Inventory (SMI). *Behavioural and Cognitive Psychotherapy*, 38, 437–458. https://doi.org/10.1017/S1352465810000226.

Mandel, T., Dunkley, D. M., & Moroz, M. (2015). Self-critical perfectionism and depressive and anxious symptoms over 4 years: The mediating role of daily stress reactivity. *Journal of Counseling Psychology*, 62, 703–717. https://doi.org/10.1037/cou0000101.

McCabe, K. O., & Fleeson, W. (2016). Are traits useful? Explaining trait manifestations as tools in the pursuit of goals. *Journal of Personality and Social Psychology*, 110, 287–301. https://doi.org/10.1037/a0039490.

McLean, K. C., Pasupathi, M., Greenhoot, A. F., & Fivush, R. (2017). Does intraindividual variability in narration matter and for what? *Journal of Research in Personality*, 69, 55–66. https://doi.org/10.1016/j.jrp.2016.04.003.

McNiel, J. M., & Fleeson, W. (2006). The causal effects of extraversion on positive affect and neuroticism on negative affect: Manipulating state extraversion and state neuroticism in an experimental approach. *Journal of Research in Personality*, 40, 529–550. https://doi.org/10.1016/j.jrp.2005.05.003.

Mejía, S., Hooker, K., Ram, N., Pham, T., & Metoyer, R. (2014). Capturing intraindividual variation and covariation constructs: Using multiple time-scales to assess construct reliability and construct stability. *Research in Human Development*, 11, 91–107. https://doi.org/10.1080/15427609.2014.906728.

Mneimne, M., Fleeson, W., Arnold, E. M., & Furr, R. M. (2017, May 29). Differentiating the everyday emotion dynamics of borderline personality disorder from major depressive disorder and bipolar disorder. *Personality Disorders: Theory, Research, and Treatment*, Advance online publication. https://doi.org/10.1037/per0000255.

Moberly, N. J., & Watkins, E. R. (2008). Ruminative self-focus and negative affect: An experience sampling study. *Journal of*

Abnormal Psychology, *117*, 314–323. https://doi.org/10.1037/0021-843X.117.2.314.

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

Molenaar, P. C. M., Huizenga, H. M., & Nesselroade, J. R. (2003). The relationship between the structure of interindividual and intraindividual variability: A theoretical and empirical vindication of developmental systems theory. In U. M. Staudinger, & U. Lindenberger (Eds.), *Understanding human development: Dialogues with life-span psychology* (pp. 339–360). Dordrecht, The Netherlands: Kluwer.

Moskowitz, D. S., & Zuroff, D. C. (2004). Flux, pulse, and spin: Dynamic additions to the personality lexicon. *Journal of Personality and Social Psychology*, *86*, 880–893. https://doi.org/10.1037/0022-3514.86.6.880.

Moskowtiz, D. S., Russell, J. J., Sadikaj, G., & Sutton, R. (2009). Measuring people intensively. *Canadian Psychology*, *50*, 131–140. https://doi.org/10.1037/a0016625.

Nillni, Y. I., Rohan, K. J., Mahon, J. N., Pineles, S. L., & Zvolensky, M. J. (2013). The role of anxiety sensitivity in the experience of menstrual-related symptoms reported via daily diary. *Psychiatry Research*, *210*, 564–569. https://doi.org/10.1016/j.psychres.2013.07.016.

O'Brien, E., Coats, A., Owens, P., Petrie, J., Padfield, P. L., Littler, W. A., … Mee, F. (2000). Use and interpretation of ambulatory blood pressure monitoring: Recommendations of the British Hypertension Society. *BMJ [British Medical Journal]*, *320*, 1128–1134.

Overall, N. C., Fletcher, G. J. O., Simpson, J. A., & Fillo, J. (2015). Attachment insecurity, biased perceptions of romantic partners' negative emotions, and hostile relationship behavior. *Journal of Personality and Social Psychology*, *108*, 730–749. https://doi.org/10.1037/a0038987.

Place, S., Blanch-Hartigan, D., Rubin, C., Gorrostieta, C., Mead, C., Kane, J., Marx, B. P., … Azarbayejani, A. (2017). Behavioral indicators on a mobile sensing platform predict clinically validated psychiatric symptoms of mood and anxiety disorders. *Journal of Medical Internet Research*, *19*, e75. https://doi.org/10.2196/jmir.6678.

Pond, R. S., Kashdan, T. B., Dewall, C. N., Savostyanova, A. A., Lambert, N. M., & Fincham, F. D. (2012). Emotion differentiation buffers aggressive behavior in angered people: A daily diary analysis. *Emotion*, *12*, 326–337. https://doi.org/10.1037/a0025762.

Rafaeli, E., Bernstein, D. P., & Young, J. (2010). *Schema therapy: Distinctive features*. London, UK: Routledge.

Rafaeli, E., Maurer, O., Lazarus, G., & Thoma, N. C. (2016). The self in schema therapy. In M. Kyrios, R. Moulding, G. Doron, N. Nedeljkovic, S. S. Bhar, & M. Mikulincer (Eds.), *The Self in Understanding and Treating Psychological Disorders* (pp. 59–48). Cambridge, United Kingdom: Cambridge University Press.

Rafaeli, E., Rogers, G. M., & Revelle, W. (2007). Affective synchrony: Individual differences in mixed emotions. *Personality and Social Psychology Bulletin*, *33*, 915–932. https://doi.org/10.1177/0146167207301009.

Rauthmann, J. F. (2016). Motivational factors in the perception of psychological situation characteristics. *Social and Personality Psychology Compass*, *10*, 92–108. https://doi.org/10.1111/spc3.12239.

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. https://doi.org/10.1177/0146167216647360.

Read, S. J., Monroe, B. M., Brownstein, A. L., Yang, Y., Chopra, G., & Miller, L. C. (2010). A neural network model of the structure and dynamics of human personality. *Psychological Review*, *117*(1), 61–92. https://doi.org/10.1037/a0018131.

Revelle, W., & Condon, D. M. (2015). A model for personality at three levels. *Journal of Research in Personality*, *56*, 70–81. https://doi.org/10.1016/j.jrp.2014.12.006.

Riley, W. T., Rivera, D. E., Atienza, A. A., Nilsen, W., Allison, S. M., & Mermelstein, R. (2011). Health behavior models in the age of mobile interventions: Are our theories up to the task? *Translational Behavioral Medicine*, *1*, 53–71. https://doi.org/10.1007/s13142-011-0021-7.

Robinson, M. D., & Clore, G. L. (2002). Episodic and semantic knowledge in emotional self-report: Evidence for two judgment processes. *Journal of Personality and Social Psychology*, *83*, 198–215. https://doi.org/10.1037//0022-3514.83.1.198.

Roche, M. J., Jacobson, N. C., & Pincus, A. L. (2016). Using repeated daily assessments to uncover oscillating patterns and temporally-dynamic triggers in structures of psychopathology: Applications to the DSM-5 alternative model of personality disorders. *Journal of Abnormal Psychology*, *125*, 1090–1102. https://doi.org/10.1037/abn0000177.

Ross, L. E., Murray, B. J., & Steiner, M. (2005). Sleep and perinatal mood disorders: A critical review. *Journal of Psychiatry and Neuroscience*, *30*, 247–256.

Schuurman, N. K., Ferrer, E., de Boer-Sonnenschein, M., & Hamaker, E. L. (2016). How to compare cross-lagged associations in a multilevel autoregressive model. *Psychological Methods*, *21*, 206–221. https://doi.org/10.1037/met0000062.

Sherman, R. A., Rauthmann, J. F., Brown, N. A., Serfass, D. G., & Jones, A. B. (2015). The independent effects of personality and situations on real-time expressions of behavior and emotion. *Journal of Personality and Social Psychology*, *109*, 872–888. https://doi.org/10.1037/pspp0000036.

Shrout, P. E., Bolger, N., Iida, M., Burke, C., Gleason, M. E. J., & Lane, S. P. (2010). The effects of daily support transactions during acute stress: Results from a diary study of bar exam preparation. In K. T. Sullivan & J. Davila (Eds.), *Support processes in intimate relationships* (pp. 175–199). New York, NY: Oxford University Press https://doi.org/10.1093/acprof:oso/9780195380170.003.0007.

Shrout, P. E., Herman, C. M., & Bolger, N. (2006). The costs and benefits of practical and emotional support on adjustment: A daily diary study of couples experiencing acute stress. *Personal Relationships*, *13*, 115–134. https://doi.org/10.1111/j.1475-6811.2006.00108.x.

Shrout, P. E., Stadler, G., Lane, S. P., McClure, M. J., Jackson, G. L., Clavél, F. D., … Bolger, N. (2018). Initial elevation bias in subjective reports. Proceedings of the National Academy of Sciences, *115*(1), E15–E23. https://doi.org/10.1073/pnas.1712277115.

Simpson, J. A., Winterheld, H. A., Rholes, W. S., & Oriña, M. M. (2007). Working models of attachment and reactions to different forms of caregiving from romantic partners. *Journal of Personality and Social Psychology*, *93*, 466–477. https://doi.org/10.1037/0022-3514.93.3.466.

Snyder, C. R. (2000). The past and possible futures of hope. *Journal of Social and Clinical Psychology*, *19*, 11–28. https://doi.org/10.1521/jscp.2000.19.1.11.

Snyder, C. R., Harris, C., Anderson, J. R., Holleran, S. A., Irving, L. M., Sigmon, S. T., Yoshinobu, L., … Harney, P. (1991). The will and the ways: Development and validation of an individual-differences measure of hope. *Journal of Personality and Social Psychology*, *60*, 570–585. https://doi.org/10.1037/0022-3514.60.4.570.

South, S. C. (2014). Personality pathology and daily aspects of marital functioning. *Personality Disorders: Theory, Research, and Treatment*, *5*, 195–203. https://doi.org/10.1037/per0000039.

Steger, M. F., & Kashdan, T. B. (2013). The unbearable lightness of meaning: Well-being and unstable meaning in life. *The Journal of Positive Psychology*, *8*, 103–115. https://doi.org/10.1080/17439760.2013.771208.

Stone, A. A., & Shiffman, S. (1994). Ecological momentary assessment (EMA) in behavorial medicine. *Annals of Behavioral Medicine*, *16*, 199–202.

Stone, A. A., Shiffman, S. S., & DeVries, M. W. (1999). Ecological momentary assessment. In D. Kahneman, E. Diener, & N. Schwarz (Eds.), *Well being: The foundations of hedonic psychology* (pp. 26–39). New York, NY: Russell Sage Foundation.

Swendsen, J. D., Tennen, H., Carney, M. A., Affleck, G., Willard, A., & Hromi, A. (2000). Mood and alcohol consumption: An experience sampling test of the self-medication hypothesis. *Journal of Abnormal Psychology*, *109*, 198–204. https://doi.org/10.1037/0021-843X.109.2.198.

Swim, J. K., Hyers, L. L., Cohen, L. L., & Ferguson, M. J. (2001). Everyday sexism: Evidence for its incidence, nature, and psychological impact from three daily diary studies. *Journal of Social Issues*, *57*, 31–53. https://doi.org/10.1111/0022-4537.00200.

Tennen, H., Affleck, G., Armeli, S., & Carney, M. A. (2000). A daily process approach to coping: Linking theory, research, and practice. *American Psychologist*, *55*, 626–636. https://doi.org/10.1037/0003-066X.55.6.626.

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. https://doi.org/10.1037/a0017653.

Trull, T. J., Lane, S. P., Koval, P., & Ebner-Priemer, U. W. (2015). Affective dynamics in psychopathology. *Emotion Review*, *7*, 355–361. https://doi.org/10.1177/1754073915590617.

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. https://doi.org/10.1037/a0012532.

Wang, L. P., & Maxwell, S. E. (2015). On disaggregating between-person and within-person effects with longitudinal data using multilevel models. *Psychological Methods*, *20*, 63–83. https://doi.org/10.1037/met0000030.

Wickham, R. E., & Knee, C. R. (2013). Examining temporal processes in diary studies. *Personality and Social Psychology Bulletin*, *39*, 1184–1198. https://doi.org/10.1177/0146167213490962.

Wilt, J., Noftle, E. E., Fleeson, W., & Spain, J. S. (2012). The dynamic role of personality states in mediating the relationship between extraversion and positive affect. *Journal of Personality*, *80*, 1205–1236. https://doi.org/10.1111/j.1467-6494.2011.00756.x.

Wouters, H., Amin, D. F., Taxis, K., Heerdink, E. R., Egberts, A. C., & Gardarsdottir, H. (2016). Associations between personality traits and adherence to antidepressants assessed through self-report, electronic monitoring, and pharmacy dispensing data: A pilot study. *Journal of Clinical Psychopharmacology*, *36*, 465–471. https://doi.org/10.1097/JCP.0000000000000541.

Wright, A. G. C. (2014). Integrating trait and process based conceptualizations of pathological narcissism in the DSM-5 era. In A. Besser (Ed.), *Handbook of psychology of narcissism: Diverse perspectives* (pp. 153–174). Hauppauge, NY: Nova Science.

Young, J. E., Klosko, J., & Weishaar, M. (2003). *Schema therapy: A practitioner's guide*. New York, NT: Guildford Press.

Zeigler-Hill, V., Holden, C. J., Enjaian, B., Southard, A. C., Besser, A., Li, H., & Zhang, Q. (2015). Self-esteem instability and personality: The connections between feelings of self-worth and the Big Five dimensions of personality. *Personality and Social Psychology Bulletin*, *41*, 183–198. https://doi.org/10.1177/0146167214559719.

Zhang, F. (2009). The relationship between state attachment security and daily interpersonal experience. *Journal of Research in Personality*, *43*, 511–515. https://doi.org/10.1016/j.jrp.2008.12.026.