Neubauer, Andreas B.; Schmiedek, Florian

Studying within-person variation and within-person couplings in intensive longitudinal data. Lessons learned and to be learned

formal und inhaltlich überarbeitete Version der Originalveröffentlichung in:
formally and content revised edition of the original source in:
Gerontology 66 (2020) 4, S. 332-339, 10.1159/000507993

Viewpoint

Studying within-person variation and within-person couplings in intensive longitudinal data: Lessons learned and to be learned

Andreas B. Neubauer[1] & Florian Schmiedek[1,2]

[1] DIPF | Leibniz Institute for Research and Information in Education, and Center for Research on Individual Development and Adaptive Education of Children at Risk (IDeA), Frankfurt am Main, Germany

[2] Goethe University, Frankfurt am Main, Germany

Short Title: WITHIN-PERSON VARIATION AND COUPLINGS

*Corresponding Author

Andreas B. Neubauer

Education and Human Development

DIPF | Leibniz Institute for Research and Information in Education

Rostocker Straße 6

60323 Frankfurt am Main, Germany

Tel: +49 69 24708-875

Fax: +49 69 24708-444

E-mail: neubauer.andreas@dipf.de

Keywords: ambulatory assessment; within-person processes; intra-individual variability

Abstract

Intensive longitudinal designs (e.g., experience sampling methods, daily-diary studies, or ambulatory assessments) continue to gain importance in psychological aging research. Empirical research using these designs has greatly facilitated our understanding of short-term within-person processes and has started to approach the question how these processes shape long-term development across the life span. The aim of this viewpoint article is to point out four key issues in intensive longitudinal designs that in our opinion require more attention than they are currently given: (a) improvement in measurement reliability, (b) the necessity to investigate inter-individual differences in short-term dynamics, (c) considerations of the time scale across which dynamic effects unfold, and (d) targeting causality by incorporating experimental methods in intensive longitudinal designs. We illustrate these four key issues by referring to a prominent example of within-person dynamics in prior empirical research: the within-person coupling of stressor occurrence and well-being (stress reactivity).

Introduction

For all scientists studying developmental processes, change is an inevitably important concept. Change occurs not only a macro-level across several weeks, months, or years, but also on a micro-level across days, hours, or even seconds. In now-classic conceptualizations of human development, the interwovenness of short-term and long-term dynamics is considered essential to understand the puzzle of human ontology across the life span [1].

In this work, we target a specific aspect of micro-level processes that received an increasing amount of attention in various fields of psychological research, including psychological aging research: within-person dynamics assessed via intensive longitudinal designs (ILDs; e.g., ambulatory assessments, experience sampling methods, or daily diary studies) in people's everyday lives. We do not aim to provide an exhaustive overview of current research in the area of intensive longitudinal methods in aging research (see for example [2], for a current overview), it is rather our intention to briefly discuss some current trends and challenges for future research in this area. Specifically, we will point out four timely issues that we think are important for researchers interested in this field: Reliability, heterogeneity, causality, and timing.

A motivating example: the coupling between stress and well-being

There are many examples in psychological aging research (or other areas of psychology as well) that could be utilized to illustrate the potential and current applications of ILDs to tackle important research questions. For example, ILDs have been utilized to examine whether momentary solitude is associated with affect and cortisol secretion [3], to investigate the interplay between social support, health complaints and negative affect [4], or to examine the within-person association between perceived competence and well-being [5]. In the present work, we chose stress reactivity to exemplify current applications of ILDs in psychological aging research and future challenges in this field. In most studies targeting stress reactivity in individuals' daily lives, participants are asked repeatedly whether a stressful event has occurred (on the current day, e.g., [6] or since the previous measurement occasion, e.g., [7]), and how they feel right now or how they felt today. The within-person association between stressor occurrence and well-being (mostly assessed as negative or positive affect) is then considered an indicator of stress reactivity, that is, the degree to which individuals respond with decreases in affective well-being to the exposure to daily hassles.

In a prototypical setup, participants report their current well-being (here: negative affect; NA) for several measurement occasions (e.g., once per day for several consecutive days). Additionally, participants report whether potentially stressful events have occurred to them today or since the last measurement. The association between stressor occurrence for person *j* on occasion *i*, $stress_{i,j}$, and person *j*'s NA at this occasion, $NA_{i,j}$, can be estimated via multilevel modeling. In order to arrive at a pure and unbiased estimate of the within-person effect of stress on NA, the (dichotomous) predictor $stress_{i,j}$ should be centered on the person mean (i.e., the proportion of occasions at which this individual has reported a stressor). Although it may seem somewhat counterintuitive to center a dichotomous predictor, person-mean centering is expected to yield superior performance compared to alternative parametrizations [8–10]. With $stress.pcen_{i,j}$ representing the (person-mean) centered stress indicator of person *j* at occasion *i*, and $stress.pmean_j$ being person *j*'s proportion of measurement occasions at which he/she reported a stressor (here: centered on the grand mean), the within-person association between stress and NA can be expressed using the following equations:

Level 1:

$$NA_{i,j} = \beta_{0j} + \beta_{1j} stress.pcen_{i,j} + \varepsilon_{i,j} \qquad (1)$$

Level 2:

$$\beta_{0j} = \gamma_{00} + \gamma_{01} stress.pmean_j + \upsilon_{0j} \qquad (2)$$

$$\beta_{1j} = \gamma_{10} + \upsilon_{1j} \qquad (3)$$

In this model, γ_{00} (the fixed intercept) represents the predicted level of NA for a person, controlling for average stress exposure (i.e., NA experienced on a day with average stress exposure for an individual who experiences an average amount of stressors). The parameter γ_{10} represents the average within-person effect of stress on NA; γ_{01} is the between-person effect of stress on NA. In the present scenario where stress is a dichotomous variable coded as 0 and 1, the within-person effect γ_{10} can be interpreted as the difference between an occasion with versus without a stressor controlling for average stress exposure of this individual. υ_{0j} and υ_{1j} represent person-specific deviations from the average intercept and fixed effect, respectively, and hence, β_{0j} is individual *j*'s

NA intercept (estimated NA of this person when stress is average) and β_{1j} is this individual's within-person effect of stress on NA.

Figure 1 depicts exemplary data that could arise from such a scenario. The red line represents the average within-person coupling of stressor exposure with NA. The single grey lines indicate the respective couplings of all individual participants.

(Insert Fig. 1 about here)

Reliability

The issue of reliability in intensive longitudinal data has recently received an increasing amount of attention. In terms of classical test theory, reliability is defined as the proportion of true score variance to the total variance of observations (= the sum of true score and error variance). Hence, reliability represents the extent to which observations are free of measurement error. In intensive longitudinal data, the variance of observations can be partitioned into two (statistically orthogonal) components: between-person variability and within-person variability. Consequently, when intensive longitudinal measures are collected, two types of reliability need to be considered: between-person reliability (= the reliability of a person's average scores) and within-person reliability (= the reliability of within-person fluctuations, also referred to as the reliability of change; [11]). Statistically, reliability on these two levels is affected by different factors, which requires that reliability needs to be examined separately on each level. When targeting the within-person associations between two variables (e.g., stress and NA), within-person reliability becomes the crucially important reliability estimate. The important question is: are the measurements applied in the present work able to reliably measure true fluctuations in the constructs of interest? Recent advancements have made this topic easily accessible to empirical researchers. Psychometrics in intensive longitudinal data is a very dynamic and productive area of research and many different approaches to estimating reliability in a multilevel context exist [11–14], some of which have been implemented into various software packages (for Mplus code to estimate multilevel reliability, see [13]; for an introduction to multilevel reliability estimation in the R package psych see [15]). We encourage all researchers working with this type of data to report reliability estimates relevant to their research questions.

Reliability estimates can also be obtained for other parameters estimated from intensive longitudinal data. For example, inter-individual differences in within-person couplings have been used to describe differences in the extent to which individuals respond to changes in their environment. Estimating the reliability of these person-level parameters is important to better judge their potential to predict future outcomes (see next section for more details). Additionally, there are several parameters that can be used to quantify an individual's within-person dynamics such as the intra-individual standard deviation (iSD), the mean squared successive difference (MSSD), and the first-order autoregressive effect (AR(1)). In the area of affect dynamics, these parameters have been reported as meaningful correlates of other outcomes. Specifically, high affect variability (= high iSD), high affective instability (= high MSSD), and high affective inertia (= high AR(1)) have been associated with lower overall well-being [16]. Notably, a recent study suggested that such indicators of affect dynamics often add no information for predicting future well-being indicators above and beyond mean levels of affective well-being [17], which might be partially explained by lower reliabilities of affect dynamic parameters. An important issue is therefore the question, how reliably inter-individual differences in these dynamic parameters can be estimated. Du and Wang [18] provide R code that can be used to estimate reliability of these person-level estimates from empirical data.

Heterogeneity

Many empirical studies primarily target the average within-person coupling (see red line in Figure 1). Hence, a crucial question that is often targeted is "is there is a (statistically significant) within-person coupling between stress and NA?". As suggested by Figure 1, focus on the average within-person coupling disregards the oftentimes quite substantial heterogeneity in these effects (note the individual grey lines). In empirical studies, there will often be meaningful random slope variability, indicating that participants differ in the extent to which two variables are coupled. One interesting question might be what accounts for inter-individual differences in these couplings. In developmental research, age has been one candidate that has often been linked with inter-individual differences in the within-person coupling of diverse variables, for example between daily pain and social as well as physical activities [20], or between the use of memory strategies and performance in an associative recall test [21]. For the stress-NA coupling, a potential correlate is the personality trait Neuroticism. Previous research has reported evidence for the hypothesized association between this personality trait and the stress-NA coupling [22]. Age differences in the within-person coupling of stressor exposure and NA have received quite substantial attention as well, though findings have been mixed, possibly due to lack of power to detect age-related differences in this within-person coupling in many studies [23].

Heterogeneity in couplings cannot always be explained by other person-level characteristics. This may occur for a number of reasons, including low reliability of the measurement [24], low statistical power [23], or simply because the correct predictors/correlates have not been identified yet. Furthermore, the measures used to target the constructs of interest might lack validity. Hence, exploring heterogeneity with ILDs might be an important addition towards examining and challenging the validity of questionnaire measures. Another explanation for a lack of convergence between person-level predictors and within-person couplings might be that within-person couplings are conceptually distinct from their investigated predictors: According to Conner and Barrett [25], momentary assessments as collected in ILDs (e.g., “How do you feel right now?”) differ qualitatively from retrospective assessments of experiences (e.g., “How did you feel last month?”) or general beliefs about experiences (e.g., “How do you feel in general)”. These authors further argue that these various types of assessment differentially predict future outcomes, with trait measures predicting deliberate future behavior and momentary assessments predicting automated and physiological responses. These qualitative differences between assessment types might explain the lack of convergence between within-person couplings (which are derived from momentary assessments) and their postulated predictors on the person level (which are often assessed as traits / beliefs) [26].

Regardless of the specific reasons why inter-individual differences in within-person couplings cannot always be explained, we as well as others [27] argue that even in these cases, estimating and reporting this variability is a valuable piece of information supplementing the results on the average (fixed) effects. In fact, if heterogeneity is large, this might indicate that the coupling between two variables does not only differ in size, but also in direction: For example, in a sample of elementary school children, for some children there was a positive association between current positive affect and working memory performance, whereas for other children, this association was negative, with higher positive affect being associated with slightly worse working memory performance [28]. This pattern of findings suggests that random effects might be useful to detect not only quantitative differences in within-person couplings, but also qualitative differences between groups. In addition to these conceptual considerations, recent methodological work further suggests that explicitly including heterogeneity in within-person associations is beneficial from a statistical perspective: Models that falsely assume that there is no heterogeneity in the effect (i.e., estimating a fixed slope when there is a random slope in the population) may lead to biased results (inflated or deflated standard errors of the fixed effect [29, 30]).

In addition to being outcomes, or theoretically interesting parameters in their own right, between-person differences in within-person couplings have also been shown to predict future outcomes. For the realm of the stress-NA couplings, prior research suggests that differences in these couplings are associated with differences in longitudinal change in well-being and depression [31]: Those individuals who showed particularly strong within-person couplings between stressor exposure and NA showed higher affective distress and higher prevalence of affective disorders ten years later. These findings suggest that heterogeneity in within-person couplings might provide potentially useful diagnostic information, hence allowing for estimating an individual's person-specific coupling of two variables (e.g., stress and NA). Ultimately, this translates into the question regarding the reliability of within-person coupling estimates. Based on considerations by Raudenbush and Bryk [9], Neubauer et al. [24] derived an estimate for this reliability they coined within-person coupling reliability. According to this estimate, reliability increases with increasing number of repeated assessments, increasing true differences in these couplings between individuals, increasing within-person variability of the predictor, and decreasing residual variance in the outcome. R code to estimate this reliability from empirical data is provided in this work, as well as an Excel sheet to estimate the number of required repeated assessments for a desired level of within-person coupling reliability.

In summary, heterogeneity in within-person couplings is a fruitful area for future research. Between-person differences in within-person couplings do not always perfectly overlap with trait measures, which could be due to measurement properties (lack of valid measures; unreliable assessments) or it could represent true differences based on qualitative differences between experiences and beliefs. Estimating heterogeneity in within-person couplings not only yields more precise statistical inferences on the average within-person couplings but can further our knowledge of between-person differences in real-life behavior and experiences.

Timing and Causality

Within-person couplings are typically operationalized as within-person effects of a designated predictor (e.g., stress) on a designated criterion (e.g., NA). Oftentimes, the choice between predictor and criterion is not arbitrary but based on a hypothesized causal effect. For example, it is expected that stress causally affects NA, and therefore stressor occurrence is used as predictor in a multilevel model with well-being as the criterion. However, this coupling is typically obtained from contemporaneous associations. For example, stress and NA might be both assessed at the end of the day [31]. Because stress is an observed variable that has not been manipulated experimentally, the reverse causal direction cannot be ruled out in such a setting. It might be that on days when participants experience higher NA, they behave in certain ways which increases their probability to be exposed to stressor – hence, a reverse causal order of effects.

A solution sometimes offered to counteract the reverse causality explanation are cross-lagged associations between predictor (e.g., stress) and outcome (e.g., NA). A necessary condition for causality is that the postulated cause temporally precedes the consequence. In this sense, stressor exposure at measurement occasion t should predict NA at a later measurement occasion t+1, whereas NA at t should not predict stressor exposure at t+1. In order to examine the within-person cross-regressive effects, various (related) approaches have been suggested: These include random-intercept cross-lagged models [32], vector autoregressive models [33], and dynamic structural equation models [34]. We note that across-time effects of stress on subsequent NA have a certain appeal when it comes to interpreting the effects as causal, but in general we think that for two reasons they offer no universal solution to the problem at hand.

First, the interpretation of the cross-lagged effects depends on the time frame chosen between two assessments. If this time frame does not match the temporal dynamics of the underlying causal process, interpretations of these effects in causal terms are rendered moot. Especially in daily diary studies, this might be problematic: Cross-lagged models in these designs would examine whether the exposure to a stressor yesterday is associated with higher NA today. While we would at no point reject the possibility that this might be true, we would caution that it is probably more likely that stressor exposure is associated with decreased well-being on the same day, but not necessarily on the next day. Hence, in this case the anticipated temporal dynamics of the causal process (within a day) does not match the sampling and analysis of the data (across days) which makes cross-lagged analyses uninformative. While using repeated samples within a day (e.g., via experience sampling) might help in better approaching the underlying temporal dynamics, we caution that even in this case it is somewhat unclear, across which time frames effects of stressors occur: should the effects occur across the next 2 minutes, 2 hours, or 6 hours? Theoretical work in combination with creative approaches to determining the time delay of within-persons effects [e.g., 7] is required to better understand across which time frames effects might occur [35].

A solution that has been offered is to explicitly estimate the lagged associations as a function of time between adjacent measurement occasions via continuous time models [36–38]. The results of a continuous time model analysis are not only estimates of the auto- and cross-regressive effects for a given time interval, but information on these effects depending on the time interval between assessments. These models are very powerful in that they provide more detailed information on the strength of the effect of a predictor on an outcome. Software to estimate these models has been developed recently [39], which greatly facilitates the application of continuous time models to empirical data. However, continuous time models assume that the analyzed process is continuous throughout the whole observation period, which may not be true for all processes studied using ILDs. Consider an example, in which stress and NA are assessed 10 times per day over 20 consecutive days. Using continuous time models, autoregressive effects of NA and stress can be estimated, as well as cross-regressive effects of NA predicting later stress, and stress predicting later NA. All of these effects can be estimated as a function of time between assessments to account for varying time intervals. In this example, overnight intervals are treated as long measurement intervals, which might not be appropriate in the present context. Whether or not the observed process (affective response to the occurrence of a stressor) continues overnight, is interrupted by sleep, is slowed down, or is qualitatively different from the process during the day is likely unknown for many psychological processes. Hence, models that do not account for potentially qualitative changes for overnight intervals might be over-simplified and not capture the true temporal (or causal) process.

A second issue that remains problematic even when the temporal dynamics are better known is that the confounding influence of a third variable on the (time-lagged) association between predictor and outcome can still not be excluded. It might be that the influence of a third variable on the ostensible predictor occurs faster than the effect of this confounder on the ostensible outcome. In this case, the lagged association does not represent a causal effect of the designated predictor, but is rather a spurious association caused by an omitted third variable. Figure 2 illustrates the implications of an omitted time-varying third variable on the estimate of a cross-lagged effect. It depicts the hypothetical scenario of a predictor X, an outcome Y, and a confounder Z. The upper part of this figure depicts the true causal effect of Z on X (which occurs instantaneously) and the true causal effect of Z on Y (which occurs with the delay of one measurement interval). Note that in this model there is no causal effect of X on Y, or vice versa. The lower part depicts the associations when the variable Z has been omitted (for example because it was not measured). Note that in this case, the results might indicate an effect of X on Y at the next measurement occasion. However, this effect is spurious since it is completely driven by an unmeasured confound with differential effect latencies on the two variables. This (hypothetical) scenario illustrates that temporal precedence is not a sufficient condition for causality.

(Insert Fig. 2 about here)

What are potential approaches to tease apart cause and effects in within-person couplings? As in other areas of scientific inquiry, experimental methods remain the via regia towards determining causality. However, experimental manipulation in daily life comes with additional challenges. For example, for many potential causes, experimental approaches might not be feasible: Even though stress research builds upon a rich experimental tradition with various standardized stress induction paradigms that are used in laboratory studies, transferring these ideas into individuals' daily lives runs into practical and ethical problems (we likely cannot induce stress in individuals' daily lives without supervision of the effects of these stressors). In these cases, combinations of ILDs with laboratory-based studies in the same sample might be a way to somewhat illuminate causal effects. For example, it could be investigated if inter-individual differences in the within-person coupling between stressor occurrence and negative affect in an ILD are related to inter-individual differences in the same individuals' responses to an experimental stress induction in the laboratory.

For other potential causes, it might be acceptable to experimentally manipulate a behavior, but treatment fidelity could be a problem. For example, participants might be sent randomized prompts to engage in reappraisal on some occasions when they reported the occurrence of a stressor and to fill in a short questionnaire a couple of minutes later. That is, in 50% of the times an individual reports a stressor, a prompt to engage in a certain behavior is given (vs. no such prompt is given in the remaining 50% of occasions at which the individual reports a stressor). If participants perfectly adhere to such a prompt (i.e., they always use reappraisal when they receive this prompt and they never use reappraisal when they do not receive this prompt), differences in affect between prompted and not-prompted occasions could be attributed to the causal effect of deploying this emotion regulation strategy on affect in the aftermath of a stressor. However, in their daily lives, participants will probably not always perfectly adhere to such prompts: On some occasions they might not engage this emotion regulation strategy even though they are prompted to do so (e.g., because they do not have the capacity to successfully employ this strategy right now), and on some occasions they might of course use reappraisal even if they are not explicitly instructed to do so. For such cases of non-intact experimental designs, instrumental variable approaches have been suggested that allow for determining the causal effect even in situations of non-perfect adherence to the treatment. These designs have recently been expanded to within-person research questions. As we showed elsewhere [40], experimental studies in daily life are feasible with realistic sample sizes (e.g., 50 participants with 50 measurement occasions) and non-perfect adherence to experimental prompts. We consider these designs a fruitful approach towards examining causality in ILDs in a naturalistic setting in individuals' daily lives.

Conclusions

Within-person processes play a prominent role in many psychological theories. They further provide a window into short-term temporal dynamics that might be important building blocks for long-term development [1]. Better understanding the ups and downs of day-to-day lives as well as their dynamic interplay is in our view an important step towards understanding long-term developmental processes including adaptation to normative and non-normative life events, as well as successful aging. With the increasing availability of smartphones, which might also continue to rise in the oldest old, collecting large amounts of data, both self-reports and passive sensor data will become possible to an extent that could not have been foreseen two decades ago.

Psychological aging research has taken an important step towards using this rich information to advance our knowledge of within-person processes. We have outlined four challenges that we think the field at large will need to tackle in this regard in the near future: improving measurement; exploring and understanding heterogeneity; considering temporal dynamics; and determining causality. These issues require further theoretical and psychometric development, new and creative ideas for ways to implement experimental studies in individuals' daily lives, and sophisticated data-analytic approaches. These developments will help us in better understanding changes and dynamic processes on short-term time scales, which may also be helpful to better understand long-term change and human development across the life span.

Acknowledgement

Statement of Ethics

No empirical data was collected in this study. Therefore, ethical approval was not required.

Disclosure Statement

The authors have no conflicts of interest to declare.

Funding Sources

The authors acknowledge support from the Jacobs Foundation grant number 2016-1245-00.

Author Contributions

ABN wrote the manuscript; FS provided feedback, contributed to revising and finalizing the paper, and approved the final version.

References

1. Nesselroade JR. The warp and woof of the developmental fabric. In: Downs R, Liben L, Palermo DS, editors. Visions of aesthetics, the environment, and development: The legacy of Joachim F. Wohwill. Hillsdale, NJ: Lawrence Erlbaum Associates; 1991. p. 213–240.
2. Diehl M, Hooker K, Sliwinski MJ, editors. Handbook of intraindividual variability across the life span. New York, London: Routledge; 2015.
3. Pauly T, Lay JC, Nater UM, Scott SB, Hoppmann CA. How We Experience Being Alone: Age Differences in Affective and Biological Correlates of Momentary Solitude. Gerontology. 2017;63:55–66. doi:10.1159/000450608.
4. Wolff JK, Lindenberger U, Brose A, Schmiedek F. Is Available Support Always Helpful for Older Adults? Exploring the Buffering Effects of State and Trait Social Support. J Gerontol B Psychol Sci Soc Sci. 2016;71:23–34. doi:10.1093/geronb/gbu085.
5. Neubauer AB, Schilling OK, Wahl H-W. What Do We Need at the End of Life? Competence, but not Autonomy, Predicts Intraindividual Fluctuations in Subjective Well-Being in Very Old Age. J Gerontol B Psychol Sci Soc Sci. 2017;72:425–35. doi:10.1093/geronb/gbv052.
6. Almeida DM. Resilience and vulnerability to daily stressors assessed via diary methods. Current Directions in Psychological Science. 2005;14:64–8. doi:10.1111/j.0963-7214.2005.00336.x.
7. Scott SB, Ram N, Smyth JM, Almeida DM, Sliwinski MJ. Age differences in negative emotional responses to daily stressors depend on time since event. Dev Psychol. 2017;53:177–90. doi:10.1037/dev0000257.
8. Enders CK, Tofighi D. Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. Psychological Methods. 2007;12:121–38. doi:10.1037/1082-989X.12.2.121.
9. Raudenbush SW, Bryk AS. Hierarchical linear models: Applications and data analysis methods. 2nd ed. Thousand Oaks: Sage Publications; 2002.
10. Wang L, Maxwell SE. On disaggregating between-person and within-person effects with longitudinal data using multilevel models. Psychological Methods. 2015;20:63–83. doi:10.1037/met0000030.
11. Cranford JA, Shrout PE, Iida M, Rafaeli E, Yip T, Bolger N. A procedure for evaluating sensitivity to within-person change: can mood measures in diary studies detect change reliably? Pers Soc Psychol Bull. 2006;32:917–29. doi:10.1177/0146167206287721.
12. Brose A, Schmiedek F, Gerstorf D, Voelkle MC. The measurement of within-person affect variation. Emotion 2019. doi:10.1037/emo0000583.
13. Geldhof GJ, Preacher KJ, Zyphur MJ. Reliability estimation in a multilevel confirmatory factor analysis framework. Psychological Methods. 2014;19:72–91. doi:10.1037/a0032138.
14. Nezlek JB. A practical guide to understanding reliability in studies of within-person variability. Journal of Research in Personality. 2017;69:149–55. doi:10.1016/j.jrp.2016.06.020.
15. Revelle W, Wilt J. Analyzing dynamic data: A tutorial. Personality and Individual Differences. 2019;136:38–51. doi:10.1016/j.paid.2017.08.020.
16. Houben M, van den Noortgate W, Kuppens P. The relation between short-term emotion dynamics and psychological well-being: A meta-analysis. Psychol Bull. 2015;141:901–30. doi:10.1037/a0038822.

17. Dejonckheere E, Mestdagh M, Houben M, Rutten I, Sels L, Kuppens P, Tuerlinckx F. Complex affect dynamics add limited information to the prediction of psychological well-being. Nature Human Behaviour. 2019;3:478–91. doi:10.1038/s41562-019-0555-0.
18. Du H, Wang L. Reliabilities of Intraindividual Variability Indicators with Autocorrelated Longitudinal Data: Implications for Longitudinal Study Designs. Multivariate Behavioral Research. 2018;53:502–20. doi:10.1080/00273171.2018.1457939.
19. Zhang S, Gamaldo AA, Neupert SD, Allaire JC. Predicting Control Beliefs in Older Adults: A Micro-longitudinal Study. J Gerontol B Psychol Sci Soc Sci 2019. doi:10.1093/geronb/gbz001.
20. Curtis RG, Windsor TD, Mogle JA, Bielak AAM. There's More than Meets the Eye: Complex Associations of Daily Pain, Physical Symptoms, and Self-Efficacy with Activity in Middle and Older Adulthood. Gerontology. 2017;63:157–68. doi:10.1159/000450786.
21. Hertzog C, Lövdén M, Lindenberger U, Schmiedek F. Age differences in coupling of intraindividual variability in mnemonic strategies and practice-related associative recall improvements. Psychol Aging. 2017;32:557–71. doi:10.1037/pag0000177.
22. Bolger N, Schilling EA. Personality and the problems of everyday life: The role of neuroticism in exposure and reactivity to daily stressors. J Personality. 1991;59:355–86.
23. Stawski RS, Scott SB, Zawadzki MJ, Sliwinski MJ, Marcusson-Clavertz D, Kim J, et al. Age differences in everyday stressor-related negative affect: A coordinated analysis. Psychol Aging 2018. doi:10.1037/pag0000309.
24. Neubauer AB, Voelkle MC, Voss A, Mertens UK. Estimating reliability of within-person couplings in a multilevel framework. Journal of Personality Assessment 2019. doi:10.1080/00223891.2018.1521418.
25. Conner TS, Barrett LF. Trends in ambulatory self-report: The role of momentary experience in psychosomatic medicine. Psychosom Med. 2012;74:327–37. doi:10.1097/PSY.0b013e3182546f18.
26. Neubauer AB, Lerche V, Koehler F, Voss A. What do you (think you) need? Perceived vs. experienced effects of need fulfillment on well-being. Journal of Research in Personality. 2020;86:103938. doi:10.1016/j.jrp.2020.103938.
27. Bolger N, Zee KS, Rossignac-Milon M, Hassin RR. Causal processes in psychology are heterogeneous. J Exp Psychol Gen. 2019;148:601–18. doi:10.1037/xge0000558.
28. Neubauer AB, Dirk J, Schmiedek F. Momentary working memory performance is coupled with different dimensions of affect for different children: A mixture model analysis of ambulatory assessment data. Dev Psychol. 2019;55:754–66. doi:10.1037/dev0000668.
29. Baird R, Maxwell SE. Performance of time-varying predictors in multilevel models under an assumption of fixed or random effects. Psychological Methods. 2016;21:175–88. doi:10.1037/met0000070.
30. Wang L, Yang M, Liu X. The Impact of Over-Simplifying the Between-Subject Covariance Structure on Inferences of Fixed Effects in Modeling Nested Data. Structural Equation Modeling: A Multidisciplinary Journal. 2018;26:1–11. doi:10.1080/10705511.2018.1489725.
31. Charles ST, Piazza JR, Mogle J, Sliwinski MJ, Almeida DM. The wear and tear of daily stressors on mental health. Psychol Sci. 2013;24:733–41. doi:10.1177/0956797612462222.
32. Hamaker EL, Kuiper RM, Grasman RPPP. A critique of the cross-lagged panel model. Psychological Methods. 2015;20:102–16. doi:10.1037/a0038889.

33. Bringmann LF, Vissers N, Wichers M, Geschwind N, Kuppens P, Peeters F, et al. A network approach to psychopathology: New insights into clinical longitudinal data. PLoS One. 2013;8:e60188. doi:10.1371/journal.pone.0060188.
34. Hamaker EL, Asparouhov T, Brose A, Schmiedek F, Muthén B. At the Frontiers of Modeling Intensive Longitudinal Data: Dynamic Structural Equation Models for the Affective Measurements from the COGITO Study. Multivariate Behavioral Research. 2018;53:820–41. doi:10.1080/00273171.2018.1446819.
35. Neubauer AB, Voss A, Ditzen B. Exploring need dynamics within and across days in everyday life: A three-level analysis. Journal of Research in Personality. 2018;77:101–12. doi:10.1016/j.jrp.2018.10.001.
36. Voelkle MC, Oud JHL, Davidov E, Schmidt P. An SEM approach to continuous time modeling of panel data: Relating authoritarianism and anomia. Psychological Methods. 2012;17:176–92. doi:10.1037/a0027543.
37. Driver CC, Voelkle MC. Hierarchical Bayesian continuous time dynamic modeling. Psychological Methods. 2018;23:774–99. doi:10.1037/met0000168.
38. Kuiper RM, Ryan O. Drawing Conclusions from Cross-Lagged Relationships: Re-Considering the Role of the Time-Interval. Structural Equation Modeling: A Multidisciplinary Journal. 2018;25:809–23. doi:10.1080/10705511.2018.1431046.
39. Driver CC, Oud JHL, Voelkle MC. Continuous Time Structural Equation Modeling with R Package ctsem. J. Stat. Soft. 2017. doi:10.18637/jss.v077.i05.
40. Schmiedek F, Neubauer AB. Experiments in the Wild: Introducing the Within-Person Encouragement Design. Multivariate Behavioral Research. 2019:1–21. doi:10.1080/00273171.2019.1627660.

Figures

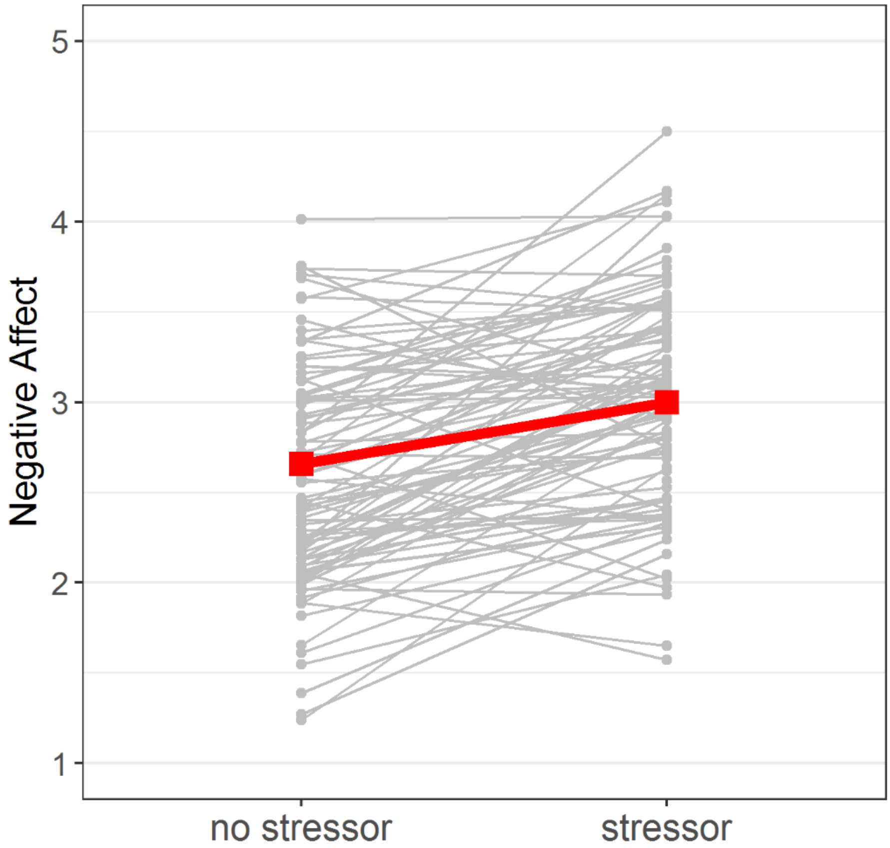

Fig. 1. Figure depicts the exemplary data that could arise from a daily diary study. Each grey line represents one individual. The red line represents the average association between stressor exposure and negative affect.

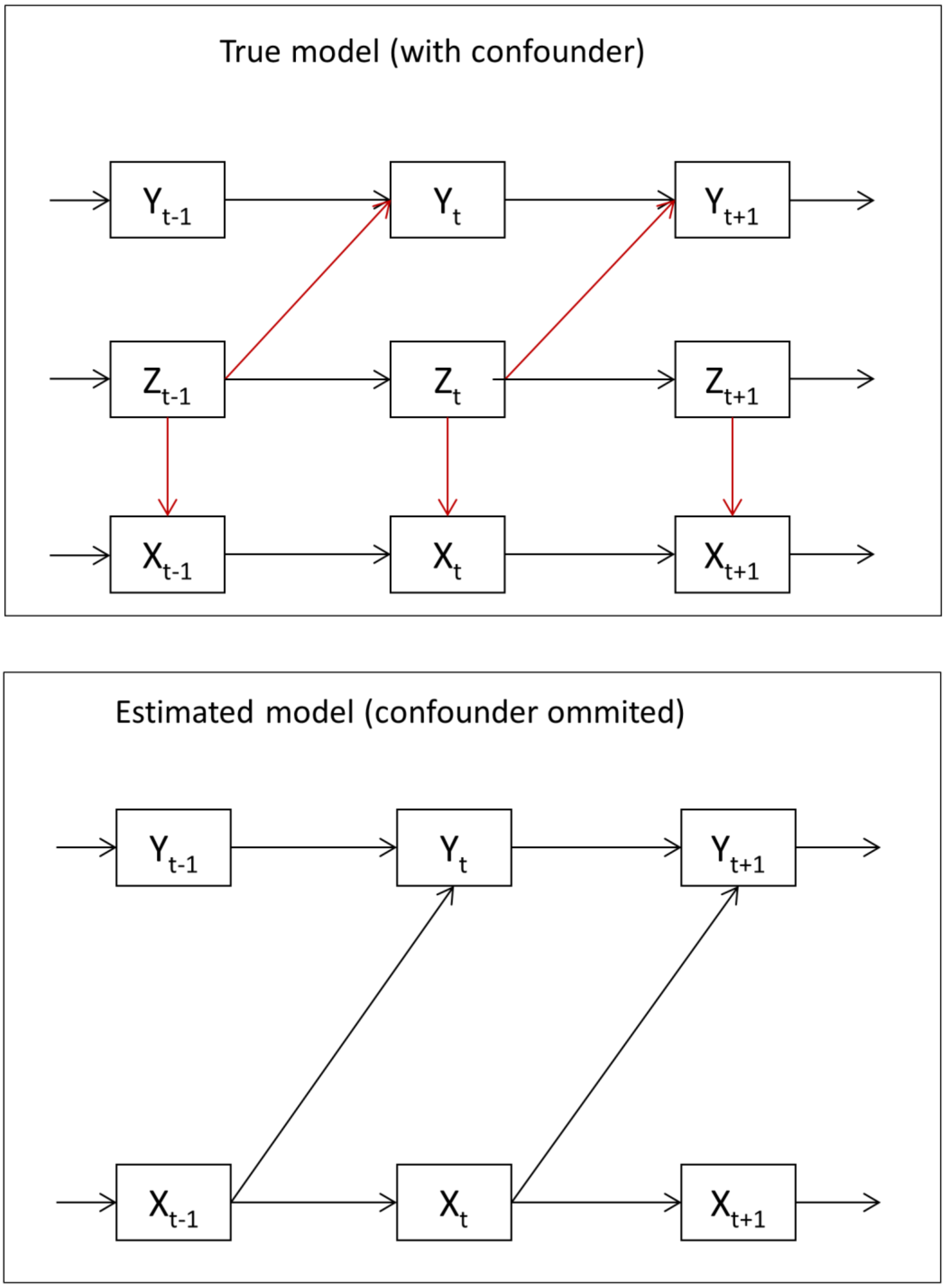

Fig. 2. Upper panel depicts the true model in which variable Z predicts X at the same measurement occasion and Y one measurement occasion later. Lower panel depicts hypothetical results from this model when the variable Z is omitted.