University of Nebraska - Lincoln
DigitalCommons@University of Nebraska - Lincoln

Faculty Publications, Department of Psychology | Psychology, Department of

1-1-2009

Persons as Contexts: Evaluating Between-Person and Within-Person Effects in Longitudinal Analysis

Lesa Hoffman
University of Nebraska - Lincoln, lhoffman2@unl.edu

Robert S. Stawski
Pennsylvania State University, rss24@psu.edu

Follow this and additional works at: http://digitalcommons.unl.edu/psychfacpub

Part of the Psychiatry and Psychology Commons

Hoffman, Lesa and Stawski, Robert S., "Persons as Contexts: Evaluating Between-Person and Within-Person Effects in Longitudinal Analysis" (2009). *Faculty Publications, Department of Psychology.* Paper 415.
http://digitalcommons.unl.edu/psychfacpub/415

This Article is brought to you for free and open access by the Psychology, Department of at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Faculty Publications, Department of Psychology by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Published in *Research in Human Development*, (2009) 6(2-3): 97-120. Copyright 2009, Taylor & Francis Group. ISSN: 1542-7609 print/1542-7617 online. Used by permission. DOI: 10.1080/15427600902911189

Persons as Contexts: Evaluating Between-Person and Within-Person Effects in Longitudinal Analysis

Lesa Hoffman (University of Nebraska-Lincoln)
Robert S. Stawski (Pennsylvania State University)

Correspondence to Dr. Lesa Hoffman: lhoffman2@unl.edu.

Relationships among multiple variables over time are of interest in many developmental areas and are frequently examined using time-varying predictors in multilevel models. Yet an incomplete specification of time-varying predictors will usually result in biased model effects. Specifically, the impact of constant, between-person sources of variation must be differentiated from the impact of time-specific, within-person sources of variation -that is, persons should be modeled as contexts. The current didactic article expands upon previous work to address why and how to model persons as contexts in longitudinal analysis. An electronic appendix of syntax for estimating these models is also provided.

Most psychological and developmental theories pertain to processes that transpire within persons over time. The repeated sampling of variables longitudinally (momentarily, daily, weekly, or yearly) creates an abundance of research designs and sampling strategies that are useful for examining within-person associations over multiple time scales (Nesselroade, 1991; Sliwinski, 2008). For instance, research on cognitive aging (*e.g.*, Schaie, 1995) often employs widely spaced measurement occasions designed to capture long-term effects of aging, whereas research on daily stress has employed daily diary designs (*e.g.*, Bolger, Davis, & Rafaeli, 2003) to examine more short-term relationships between daily physical and emotional well-being. Yet no matter what the time scale, collection of longitudinal data per se is insufficient for addressing hypotheses about

within-person processes—one must utilize statistical methods that properly distinguish the multiple sources of information obtained. Specifically, although longitudinal data are collected with the goal of assessing within-person associations, they also provide information about cross-sectional, between-person associations (*e.g.*, relationships among individual differences in overall levels in addition to daily levels of physical and emotional well-being).

The point of this article is to describe the rationale and mechanisms through which one should consider persons as contexts when examining within-person associations. Although the need to incorporate multiple levels of contextual influences (*e.g.*, of stable individual traits, a person's environment, etc.) is certainly not new (*e.g.*, Bronfenbrenner, 1979), what is not often recognized is the extent to which these higher-level influences can permeate measures collected longitudinally as well. Within-person processes do not happen in a vacuum, and the additive and interactive influences of more stable individual differences contained in the longitudinal measures need to be modeled explicitly. Otherwise, associations that reflect solely longitudinal or within-person relationships cannot be distinguished from those that reflect cross-sectional or between-person relationships. Only by formulating statistical models that include the contribution of stable individual differences can within-person associations be elucidated properly. Accordingly, the goal of the current article is twofold: (1) to expand upon previous work to address why persons should be modeled as contexts in longitudinal analysis—that is, the importance of distinguishing between-person effects from within-person effects, and (2) to provide a detailed illustration of how to do so in the framework of multilevel modeling.

Multilevel models are known by a variety of synonyms (*e.g.*, hierarchical linear models, general linear mixed models) and their defining feature is their capacity to provide quantification and prediction of random variance due to multiple sampling dimensions (*e.g.*, across occasions, across persons, or across groups). In general, multiple observations arising from the same sampling unit (*e.g.*, multiple occasions from the same person, or multiple persons from the same group) will have model residuals that are more alike than residuals from different sampling units—that is, the residuals from the same person or group will be dependent, or correlated. Multilevel models provide flexible and advantageous strategies for modeling such dependencies. One such strategy is the use of random effects, or model effects that are permitted to vary across sampling units. For instance, in models of within-person change over time, each person may have his or her own intercept and rate of change (slope), and these individual differences in intercepts and slopes can be quantified and predicted. As will be fully explicated in later sections, between-person effects

and random effects error variances are known as Level-2 model parameters, and within-person effects and residual error variances are known as Level-1 model parameters.

Although multilevel models are commonly employed for longitudinal data, multiple alternative model specifications exist for predictors that vary over time, only some of which explicitly incorporate the contribution of stable individual differences in those time-varying predictors. In the current article we illustrate with a detailed example how these model alternatives result in important differences in the interpretation of model effects. A similar situation occurs in models for cross-sectional clustered data, in which persons are nested within higher-order groups. In that case, although effects at the individual level may be of primary interest, one cannot ignore the possibility that the same predictors may have aggregate effects at the group level—that is, group-level variance permeates the individual measures. For instance, although one may be interested in the effect of student socioeconomic status (SES) on student achievement, it is also possible that the SES of the student's school (*e.g.*, the mean SES across students) would have an additional additive or interactive effect on the student's achievement. The term for an additive effect of the predictor at the group aggregate level of analysis is a contextual effect (*cf.* Raudenbush & Bryk, 2002, p. 140). Accordingly, one may wish to distinguish a between-group effect (*e.g.*, the effect of being in a higher-SES school) from a within-group effect (*e.g.*, the effect of being higher-SES relative to one's school) via the contextual effect (*e.g.*, the incremental effect of school SES after controlling for one's individual SES). Fortunately, there are many helpful treatments of how to distinguish these effects using multilevel models for clustered data (*e.g.*, Enders & Tofighi, 2007; Hofmann & Gavin, 1998).

These same issues apply readily to the examination of within-person processes: one can think of the person as the context in longitudinal data in much the same way that one can think of the group as the context in cross-sectional clustered data. Yet in our work we have found that the extension of these same concepts (*i.e.*, distinguishing effects at multiple levels of analysis) to the longitudinal case is far from intuitive. Although deemed by some as "arguably less complex . . . than in the cross-sectional context" (Enders & Tofighi, 2007, p. 122), in contrast, we have found extension of these issues to be considerably more complex in the longitudinal case. Further, the treatments of this issue in the longitudinal case have been relatively restricted in scope (*cf.* Diggle, Heagerty, Liang, & Zeger, 2002; Hedeker & Gibbons, 2006; Schwartz & Stone, 1998). Thus, we hope this work will help fill in the gap of how existing models can be used to reflect the contribution of persons as contexts in longitudinal data.

In this work we illustrate the modeling of persons as contexts via an extended example from a diary study of cognition, health, and aging (see

Sliwinski, Smyth, Hofer, & Stawski, 2006) in which daily stressful experiences and daily negative mood are used to predict daily physical symptoms in younger and older adults. We first introduce a baseline longitudinal model and distinguish among types of predictors and the roles that they can play in longitudinal models. We then focus on the two main ways of distinguishing these roles in the multilevel modeling framework. We follow the extended example with some suggestions for alternative modeling approaches that may also be useful in a longitudinal context.

Although the material that follows is applicable to longitudinal studies that focus on systematic and durable within-person change as well as those that focus on more transient and reversible within-person fluctuation (*e.g.*, Nesselroade, 1991), our example is limited to within-person fluctuation. Our rationale in doing so is that models for within-person change over time have been well explicated, whereas models for within-person fluctuation have received less attention. Daily diary data, such as those we present here, are becoming increasingly popular in social science research, in which investigators are interested in understanding and predicting within-person fluctuations (intraindividual variability), and for which multilevel models are frequently used. Further, longitudinal predictors that change systematically over time require an even more complex parameterization (see Selig & Preacher, this issue), and thus to keep our exposition manageable we focus on outcomes and predictors that are primarily expected to vary, as opposed to change, over time given the 2-week duration of the example study.

ILLUSTRATIVE EXAMPLE

Example Data

The example data were taken from the first assessment burst (6 days over a 2-week period) of the Cognition, Health, and Aging Project (CHAP; Sliwinski *et al.*, 2006). The sample of 1,082 observations was taken from 68 undergraduates (M = 20.5 years, SD = 1.2, range = 18-25) and 113 older adults (M = 80.2 years, SD = 6.4, range = 66-95). Men comprised 24% and 28% of the younger and older adults, respectively. The outcome variable was a sum checklist of daily physical symptoms participants reported experiencing in the past 24 hours (from Larsen & Kasimatis, 1991) including aches/pain, gastrointestinal, cardiovascular, upper respiratory, and "other" physical symptoms. Younger adults reported significantly more physical symptoms on average (M = 2.4, SD = 1.5) than older adults (M = 1.4, SD = 1.2), $p < 0.01$.

The two other predictors besides age were daily stress and daily negative mood. Daily stress was assessed using a version of the Daily Inven-

tory of Stressful Events (DISE: Almeida, Wethington, & Kessler, 2002) and was defined as a dichotomous variable reflecting whether any of five stem questions about interpersonal relationships, events, and health had been endorsed (0 = stressor-free day, 1 = stressor day). Younger adults reported a significantly higher proportion of stressor days (M = 74%, SD = 23%) than did older adults (M = 46%, SD = 33%), $p < 0.01$. Finally, current negative mood was assessed using a sum score from a version of Philadelphia Geriatric Center Positive and Negative Affect Scales (Lawton, Kleban, Dean, Rajagopal, & Parmelee, 1992) with 5 items: sad, annoyed, worried, irritated, and depressed, each rated on a 5-point scale (not at all to extremely). Younger adults reported significantly higher negative mood on average (M = 6.8, SD = 1.6) than older adults (M = 6.0, SD = 1.2), $p < 0.01$.

All multilevel models were estimated in SAS PROC MIXED using full information maximum likelihood (ML). Although ML can underestimate variance components in small samples (*i.e.*, < 50 persons), ML is required to compare the fit of models that differ in fixed and random effects. Nested models can be compared using the difference in the model -2 Log Likelihood (LL) values (*i.e.*, c2 values) as a function of degrees of freedom equal to the difference in the number of model parameters (see Snijders & Bosker, 1999). Finally, ESTIMATE statements were used to produce linear contrasts of fixed effects and their associated standard errors where needed. An electronic appendix of the SAS code used to fit each reported model is also provided.

A Baseline Longitudinal Model

Understanding the role of predictors in a longitudinal model relies on distinguishing its two distinct sides. One side is the model for the means (fixed effects), or how the outcome will vary as a function of values on the predictor variables. The model for the means includes the fixed intercept and any fixed main effects and interactions among predictors that contribute to the expected mean for a given person on a given occasion, and its parameters are usually of primary empirical interest. The other side is the model for the variances (random effects and residuals), or how the model errors are distributed and related over time and persons. Proper specification of the model for the variances is necessary for obtaining standard errors and significance tests of fixed effects that are as accurate as possible. Further, individual differences or group differences in distinct parts of the model for the variances can be of substantive interest in and of themselves (*i.e.*, individual differences in variability, *cf.* Hoffman, 2007).

The simplest possible model includes only a fixed intercept and a residual. That is, the model for the means would predict the sample grand mean for each observation, and the model for the variances would consist

of a single residual for the difference between each outcome and the grand mean. The residuals are then assumed uncorrelated with equal variance across persons and occasions. Such a model does not take into account that residuals from the same person are more likely to be related than residuals from different persons. To address this dependency, we add to the model for the variances, as shown in the empty longitudinal model in Equation 1:

$$\begin{aligned} &\text{Level 1 (L1):} && \text{Symptoms}_{di} = \beta_{0i} + e_{di} \\ &\text{Level 2 (L2):} && \beta_{0i} = \gamma_{00} + U_{0i} \\ &\text{Combined:} && \text{Symptoms}_{di} = \gamma_{00} + U_{0i} + e_{di}, \end{aligned} \tag{1}$$

in which Symptoms_{di} is the outcome on day *d* for person *i*. The Level-1 model describes within-person variation in symptoms as a function of a person-specific intercept (B_{0i}) and a day- and person-specific residual deviation from that intercept (e_{di}). The Level-2 model describes between-person variation in the mean symptoms across days as a function of a fixed intercept (γ_{00}), which is the grand mean for the sample, and a person-specific random intercept (U_{0i}), which is the difference between the grand mean and person *i*'s mean across days. Figure 1 displays the fitted means (dashed lines) of daily physical symptoms for nine randomly selected persons. Between-person variation is shown by the variation of the person means; within-person variation is shown by the deviation of each occasion from the person mean. Thus, the idea of persons as contexts has been implemented by explicitly modeling person mean differences in physical symptoms.

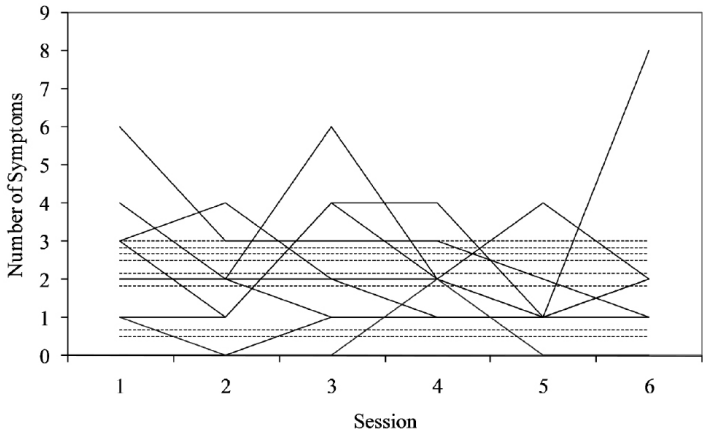

FIGURE 1 Random subsample of nine individual trajectories of daily symptoms during the study. Differences between the fitted person means (dashed lines) represent between-person variation; deviation of each observation from the fitted person means represents within-person variation.

Rather than estimating each individual U_{0i} and e_{di} separately as parameters, their variances are estimated instead (as τ_0^2 and σ_e^2, respectively). The U_{0i}'s and edi terms are assumed to be normally distributed, uncorrelated with each other, and independent across persons. Further, this model assumes that the only reason that the residuals edi residuals would be correlated is because of constant mean differences across persons: the random intercept. Once that variation is removed and estimated as a separate variance component (τ_0^2), the e_{di}'s are supposed to be uncorrelated with equal variance over time. This may not be realistic in many longitudinal studies in which residuals from occasions closer together in time may be more closely related. A variety of alternative patterns of within-person correlation are available to address those cases, paired with or without a random intercept (*cf.*, Hoffman, 2007; Singer & Willett, 2003).

In our example data, the empty longitudinal model returns a grand mean of symptoms across days (fixed intercept γ_{00}) of 1.76 with a random intercept variance (τ_0^2) of 1.71 and a residual variance (σ_e^2) of 1.58. These variance components can be used to calculate an intraclass correlation (ICC), as shown in Equation 2:

$$\begin{aligned} \text{ICC} &= \frac{\text{BP variation}}{\text{BP} + \text{WP variation}} = \frac{\text{Var}(U_{0i})}{\text{Var}(U_{0i}) + \text{Var}(e_{di})} \\ &= \frac{\tau_0^2}{\tau_0^2 + \sigma_e^2} = \frac{1.71}{1.71 + 1.56} = .52. \end{aligned} \tag{2}$$

The ICC indicates that 52% of the variance in physical symptoms is between-persons. One can also calculate a 95% random effects confidence interval (Snijders & Bosker, 1999) as the fixed intercept $\pm$ 1.96*SQRT(τ_0^2). Unlike typical confidence intervals that concern the precision of the point estimate (*i.e.*, the fixed effect), a random effects confidence interval conveys the predicted variability of a given effect across persons. As such, the random effects confidence interval for the intercept indicates that 95% of the sample is expected to have a mean of physical symptoms across days between 0 and 4.3 (the lower bound was truncated at 0 given the outcome scale).

Time-Invariant Predictors in Longitudinal Models

One may then add two different kinds of predictors to this empty longitudinal model: time-invariant or time-varying. Time-invariant predictors (*i.e.*, Level-2 predictors) are those measured only once per person. At one end of the spectrum are time-invariant predictors that will never change over time, such as biological sex at birth or race. In the middle of the spectrum would be time-invariant predictors that are not likely to change over

the course of a study, and so may be inquired about only once to save time and resources. For example, in a study of daily behavior, one might measure key personality variables only once under the assumption that, although personality may change over the course of a lifetime, it is not likely to change over the course of a week. The same may hold for educational attainment, SES, or religious affiliation. At the other end of spectrum, however, are time-invariant predictors from a single measurement occasion that are likely to change over time. In a study of adolescents, one might have measured parental attitudes only at a single age. This attitude variable is still a time-invariant predictor, even though parental attitudes are likely to change as their child ages. In that case, one must be careful to interpret its effects conditionally on the measurement occasion (*i.e.*, as effects of "parent attitudes at age 13" rather than "parent attitudes"). In our example, we consider age as a time-invariant predictor given the two-week interval of the study.

Age as a Time-Invariant Predictor

The role of time-invariant predictors by themselves in longitudinal models is relatively straightforward. In the model for the means, they carry between-person, Level-2 effects on the expected person mean outcome across occasions, thereby decreasing the Level-2 random intercept variance. Two-way and higher-order interactions among time-invariant predictors would similarly modify the expected person mean over occasions. In the model for the variances, time-invariant predictors can also predict differential magnitudes of Level-2 between-person variation or Level-1 within-person variation. In our current example, these three possible effects of our time-invariant predictor of age are illustrated in Equation 3:

$$\begin{aligned} &\text{L1: } \text{Symptoms}_{di} = \beta_{0i} + (Y_i)e_{di} + (O_i)e_{di} \\ &\text{L2: } \beta_{0i} = \gamma_{00} + \gamma_{01}(\text{Older}_i) + (Y_i)U_{0i} + (O_i)U_{0i}. \end{aligned} \tag{3}$$

In this equation and those to follow, the first subscript is an index for which term is being modified in the Level-1 equation, and the second subscript is an index for the Level-2 equation. In the model for the means, the fixed intercept (γ_{00}) of 2.37 (SE = 0.18) represents mean symptoms in the younger adults specifically. The main effect of age (γ_{01}) of -0.96 was significant (SE = 0.22, $p < 0.001$), indicating that older adults reported about one fewer daily symptom on average (1.41, SE = 0.11). An additional linear slope of age in the older group was also examined (*cf.* Hoffman & Rovine, 2007) but was not found to be significant in any model, and was thus not included.

In the model for the variances, the Level-2 random intercept variance (τ_0^2) and the Level-1 residual variance (σ_e^2) each differ across younger and older adults (as shown via the dummy codes of Y_i and O_i for younger = 1 and older = 1, respectively). Given the extreme age groups study design, we permitted separate variance components for younger and older adults to account for possible age differences in the magnitude of variability (*cf.* Hofer & Sliwinski, 2001). Age heterogeneity was tested separately for each variance term and was found to be significant for each. The random intercept variance (τ_0^2) was significantly greater in younger adults (3.00, SE = 0.23) than in older adults (0.72, SE = 0.04), $\Delta_\chi^2(1) = 5.0$, $p < 0.026$, indicating that younger adults were more variable as a group in mean physical symptoms (greater between-person variation). The residual variance (σ_e^2) was also significantly greater in younger adults (1.80, SE = 0.40) than in older adults (1.34, SE = 0.19), $\Delta_\chi^2(1) = 221.4$, $p < 0.001$, such that younger adults also showed more symptom fluctuation across days than older adults (greater within-person variation).

Time-Varying Predictors in Longitudinal Models

Time-varying predictors (time-level or Level-1 predictors) are those that are measured at each occasion and that vary across occasions. The following section demonstrates alternative methods of specifying models with time-varying predictors and the resulting differences in model parameters. The complexity in doing so can be summarized as follows: because time-varying predictors are usually composed of two sources of variation, they are usually really two variables instead of one. For instance, in our example, although daily negative mood is a time-varying predictor, to the extent that some people are just "grumpier" than others, daily negative mood will contain systematic between-person variation as well as within-person variation. These two sources of variation are likely to have differential effects on the outcome—a between-person effect and a within-person effect, respectively. It is commonly believed that the role of time-varying predictors in the model for the means is to account for within-person, Level-1 residual variation, and this is true. But multiple parameters will be needed for the predictor to the extent that there is systematic between-person variance in the time-varying predictor, and to the extent that this between-person variance has a differentiable effect on the outcome than its within-person variance. Additionally, time-varying predictors can play a role in the model for the variances. They can have Level-2 random effects, or person-to-person differences in the slope for the time-varying predictor. They also might predict the magnitude of Level-1 residual variation, although due to space limitations we do not pursue these models here.

There are (at least) two reasons why between-person and within-person effects of time-varying predictors might differ from each other. The first reason is the theoretical construct measured by the predictor at each level of analysis. For negative mood, a confluence of chronic factors can result in a given person being more or less grumpy, such as personality variables, lifestyle differences, and so forth. Yet different, more acute factors are likely to be the reason why negative mood is worse on some days than others, such as temporally specific deviations from normal routines of work, family, or health. Thus, given that between-person and within-person variation represent two different theoretical constructs, their effects on a given outcome will often be of different magnitudes or even different directions. In our experience, it has been the rule, rather than the exception, that the between-person and within-person effects of time-varying predictors will differ from each other, if not in direction, almost certainly in magnitude.

In addition to differences in the constructs they reflect, the second reason why between-person and within-person effects will differ from each other is because they are unstandardized coefficients on different scales, as is shown in later sections. In summary, the total impact of a single time-varying predictor may require multiple fixed effects in the model for the means and perhaps random effects or parameters for heterogeneity in the model for the variances as well. Consequently, properly specifying and interpreting the overall contribution of a time-varying predictor in a longitudinal model can be a complex endeavor.

Grand-Mean-Centering of a Time-Varying Predictor of Daily Stress

Of interest in the current example is the extent to which daily stressors, as measured by a dichotomous predictor of whether or not a stressor was reported each day, would relate to daily variation in physical symptoms. A naïve approach would be to simply include daily stress in the Level-1 equation predicting daily symptoms, as shown in Equation 4:

$$\begin{aligned} \text{L1:}\ & \text{Symptoms}_{di} = \beta_{0i} + \beta_{1i}(\text{DayStress}_{di}) + (Y_i)e_{di} + (O_i)e_{di} \\ \text{L2:}\ & \beta_{0i} = \gamma_{00} + \gamma_{01}(\text{Older}_i) + (Y_i)U_{0i} + (O_i)U_{0i} \\ & \beta_{1i} = \gamma_{10}. \end{aligned} \tag{4}$$

The fixed intercept (γ_{00}) now represents mean symptoms for a younger adult on days without a stressor ($\text{Older}_i = 0$ and $\text{DayStress}_{di} = 0$). The individual effect of daily stress (β_{1i}) is represented by just a fixed effect (γ_{10}) at Level 2 and is the expected difference in symptoms between days without a stressor and days with a stressor. So far the effect of daily stress is constrained to be the same across persons: It does not differ by age and it does not have a random effect.

This method of specifying Level-1 effects is referred to as *grand-mean centering*. The key idea is that only a constant is subtracted from the Level-1 predictor to include zero in its scale (usually the grand mean is used as the centering constant, thus the name of the method). Centering by subtracting a constant (for any predictor) results in exactly the same model fit and expected values, that is, there is no "wrong" centering constant. However, centering will improve interpretability of model parameters—specifically, the model intercept (the expected value of the outcome when the predictors are at zero), and main effects in the presence of higher-order interactions (which become the simple effect when the interacting predictor is at zero).

In this example, a zero value for daily stress already represents a nonstressor day, so no further centering was needed. The fixed Level-1 effect for stress (γ_{10}) of 0.34 was significant, SE = 0.08, $p < 0.001$. Although this effect would seem to suggest that days with stressors are expected to have 0.34 more symptoms, its actual interpretation is more complicated because we have not yet attended to the person as a context with regard to the effect of daily stress. Some people are just more likely to experience stressors than other people, and so daily stress actually contains between-person variation as well—approximately 31%, as estimated from an empty longitudinal model for daily stress. Thus, the time-varying, within-person effect of daily stress (*i.e.*, the difference in symptoms on stressor days compared to nonstressor days) is confounded with the time-invariant, between-person effect of person mean stress (*i.e.*, the effect on mean symptoms of stable individual differences in stressor exposure). To differentiate these two effects, we introduce a new predictor variable for person mean stress, centered at a proportion of 50% of days with a stressor, given that a person with a mean of 0% of days would have to have all daily stress scores of zero (and thus could not logically show an effect of daily stress). Person mean stress is entered as a main effect in the Level-2 model for the intercept, as shown in Equation 5:

$$\begin{aligned}
&\text{L1: } \text{Symptoms}_{di} = \beta_{0i} + \beta_{1i}(\text{DayStress}_{di}) + (Y_i)e_{di} + (O_i)e_{di} \\
&\text{L2: } \beta_{0i} = \gamma_{00} + \gamma_{01}(\text{Older}_i) + \gamma_{02}(\text{PMstress}_i) + (Y_i)U_{0i} + (O_i)U_{0i} \\
&\qquad \beta_{1i} = \gamma_{10}. \qquad (5)
\end{aligned}$$

Upon doing so, the Level-1 effect upon daily stress (γ_{10}) is reduced to 0.25 (SE = 0.08, $p < 0.003$), whereas the Level-2 effect of person-mean stress (γ_{02}) is 1.35 (SE = 0.30, $p < 0.001$).

But what do these effects actually mean? The short answer is that the Level-1 effect of daily stress is now specifically the within-person effect of stress, and the Level-2 effect of person mean stress is the contextual effect of person mean stress. That is, the Level-1 within-person effect of daily stress indicates that symptoms are expected to be 0.25 higher on days

when a stressor occurs, holding person mean stress constant. The within-person effect operates locally—directly on the expected outcome at that same occasion. The Level-2 contextual effect then indicates the incremental effect of being a "high stress person" on mean symptoms over time, after controlling for today's stress. The context effect operates globally, affecting the outcome across days: after controlling for whether or not a stressor happened on a given day, for every unit increase in person mean stress, a person's symptoms are expected to be 1.35 higher on average. In our example, a unit increase is the difference between someone with 0 stressors and someone with stressors on all 6 days. Thus, we can rescale the coefficient such that for every additional day of stress, a person's symptoms are expected to be 1.35/6 = 0.23 higher on average.

The possible effects obtained from grand-mean-centering can be summarized as follows. When the person mean of the time-varying predictor is also included in the Level-2 model, the Level-1 effect of the time-varying predictor is the within-person effect, and the Level-2 effect is the contextual effect. The Level-1 and Level-2 effects can then be added together to become the total between-person effect, which is the effect of being a "high stress" person, not controlling for current stress. The between-person effect is elaborated further in the next section. Finally, recall that the effect of the original Level-1 effect of stress when included by itself was 0.34. This is the composite effect of daily stress: an uninterruptible blend of the within-person and between-person effects, weighted by the intraclass correlation of the time-varying predictor. To the extent that more of its variation is within-persons, the composite effect will be closer to the true within-person effect, as is the case for daily stress. Another way of thinking about the Level-2 contextual effect is that it directly represents the difference of the between-person and within-person effects. If the contextual effect of the person mean predictor is significant, then the composite effect (the effect of the Level-1 predictor when by itself) should not be used.

We continue our example by adding interactions of both effects with age to the model for the means. We also examine the role of daily stress in the model for the variances by adding a Level-2 random slope for daily stress, as shown in Equation 6:

$$\begin{aligned}
\text{L1:}\quad & \text{Symptoms}_{di} = \beta_{0i} + \beta_{1i}(\text{DayStress}_{di}) + (Y_i)e_{di} + (O_i)e_{di} \\
\text{L2:}\quad & \beta_{0i} = \gamma_{00} + \gamma_{01}(\text{Older}_i) + \gamma_{02}(\text{PMStress}_i) + \\
& \qquad \gamma_{03}(\text{PMStress}_i{*}\text{Older}_i) + (Y_i)U_{0i} + (O_i)U_{0i} \\
& \beta_{1i} = \gamma_{10} + \gamma_{11}(\text{Older}_i) + (Y_i)U_{1i} + (O_i)U_{1i}. \qquad (6)
\end{aligned}$$

Results for this model are shown under the first column in Table 1. In the Level-1 equation, β_{1i} is the individual within-person effect of daily stress. The Level-2 equation for β_{1i} then specifies its effect for younger adults (γ_{10})

TABLE 1
Results from Multilevel Models Predicting Daily Physical Symptoms

Fixed Effects		*Equation 6* *Est (SE)*	*Equation 8* *Est (SE)*	*Equation 9* *Est (SE)*
Intercept	Young	1.67 (.24)**	1.62 (.24)**	2.20 (.28)**
	Old	1.37 (.10)**	1.38 (.09)**	1.46 (.10)**
	Difference	–0.30 (.26)	–0.24 (.25)	–0.71 (.30)*
Daily stress (within-person)	Young	0.76 (.27)**	0.72 (.25)**	0.59 (.25)**
	Old	0.20 (.09)*	0.19 (.09)*	0.17 (.10)$^{+}$
	Difference	–0.56 (.29)*	–0.53 (.27)*	–0.42 (.27)
Daily stress (contextual)	Young	0.59 (.80)	0.60 (.78)	0.69 (.78)
	Old	1.41 (.30)**	0.98 (.29)**	1.03 (.29)**
	Difference	0.82 (.85)	0.38 (.83)	0.34 (.83)
Daily stress (between-person)	Young	1.35 (.74)$^{+}$	1.32 (.72)$^{+}$	1.28 (.73)$^{+}$
	Old	1.60 (.30)**	1.17 (.28)**	1.20 (.28)**
	Difference	0.25 (.80)	–0.14 (.78)	–0.08 (.79)
Negative mood (within-person)	Young		0.12 (.05)*	0.12 (.05)*
	Old		0.03 (.03)	0.02 (.03)
	Difference		–0.09 (.06)	–0.10 (.06)
Negative mood (contextual)	Young		–0.00 (.12)	0.00 (.11)
	Old		0.33 (.08)**	0.34 (.08)**
	Difference		0.34 (.14)*	0.34 (.14)*
Negative mood (between-person)	Young		0.11 (.11)	0.12 (.11)
	Old		0.36 (.08)**	0.37 (.08)**
	Difference		0.25 (.13)$^{+}$	0.24 (.13)$^{+}$
Linear time (within-person)	Young			–0.21 (.05)**
	Old			–0.04 (.02)*
	Difference			0.17 (.05)**
Residual variance	Young	2.62 (.21)**	2.47 (.21)**	2.35 (.20)**
	Old	0.70 (.05)**	0.65 (.04)**	0.64 (.04)**
Intercept variance	Young	1.11 (.43)**	1.08 (.42)**	1.13 (.42)**
	Old	0.70 (.13)**	0.57 (.11)**	0.56 (.11)**
Daily stress slope variance	Young	1.53 (.73)*	1.41 (.67)*	1.04 (.59)
	Old	0.12 (.12)	0.19 (.14)$^{+}$	0.20 (.14)$^{+}$
Negative mood slope variance	Young		0.03 (.02)$^{+}$	0.03 (.02)$^{+}$
	Old		0.02 (.01)*	0.02 (.01)*
Intercept-stress covariance	Young	–0.11 (.43)	–0.03 (.41)	0.08 (.37)
	Old	0.28 (.08)**	0.25 (.09)**	0.25 (.09)**
Intercept-mood covariance	Young		0.04 (.10)	0.06 (.10)
	Old		0.01 (.03)	0.01 (.03)
Stress-mood covariance	Young		0.18 (.13)	0.16 (.12)
	Old		–0.02 (.03)	–0.03 (.03)
–2 Log Likelihood		3612.5	3653.2	3538.4
AIC		3640.5	3611.2	3590.4
BIC		3686.5	3688.4	3674.0

Note. AIC = Akaike Information Criteria; BIC = Bayesian Information Criteria.
$^{+}p < .10$. $^{*}p < .05$. $^{**}p < .01$.

of daily stress*age). The within-person daily stress effect (γ_{10}) of 0.76 was significant for younger adults and significantly greater by 0.56 (γ_{11}) than the effect for older adults of 0.20, which was also significant. Model effects for older adults were not estimated specifically, but were obtained by separate ESTIMATE statements in PROC MIXED. The Level-2 contextual effect for person mean stress for younger adults (γ_{02}) of 0.59 was not significant and was nonsignificantly smaller by 0.82 (γ_{03}) than the effect for older adults of 1.41, which was significant. A marginally significant between-person effect of 1.35 was found for the younger adults by adding their within-person and contextual effects ($\gamma_{10} + \gamma_{02}$) together, and a significant between-person effect of 1.60 was found for the older adults by also including the interactions of each term with age ($\gamma_{10} + \gamma_{02} + \gamma_{11} + \gamma_{03}$), although the difference of the between-person effects of 0.25 was not significant. In sum, in younger adults the within-person and between-person effects were each significant but equivalent, whereas in older adults, the between-person effect was significantly greater than the within-person effect, and both effects were significant. Younger and older adults differed significantly in their within-person effects, but not in their between-person effects. Finally, we examined the cross-level interaction of daily stress and person mean stress, and whether this interaction differed by age (as would be represented on the Level-2 equation for β_{1i} by γ_{12} and γ_{13}). Neither effect was significant, indicating that the size of the within-person effect of daily stress did not depend on person mean stress, equivalently for both age groups.

We then examined the role of daily stress in the model for the variances. At Level 2, a random deviation from the age group slope for daily stress was included (U_{1i}), again with different variances (τ_1^2) across groups and was found to significantly improve model fit relative to the random intercept-only model, $\Delta_\chi^2(4) = 28$, $p < 0.001$. We computed 95% random effects confidence intervals to describe the slope variation as the fixed slope ± 1.96*SQRT(τ_{1i}^2). Accordingly, 95% of the younger adults were expected to have between 1.66 fewer symptoms and 3.19 more symptoms on stressor days, whereas older adults were expected to have between 0.48 fewer symptoms and 0.87 more symptoms on stressor days. Thus, it appears the younger adults showed more variability in their within-person effect of daily stress on daily symptoms.

Person-Mean-Centering of a Time-Varying Predictor of Daily Negative Mood

Previous daily stress research has shown that negative mood is higher on stressor days as compared to nonstressor days (*e.g.*, Stawski, Sliwinski, Almeida, & Smyth, 2008). Therefore, we continue our example by examining to what extent the observed associations between daily stress and physical symptom reports would remain after controlling for negative

mood. Negative mood contains approximately 33% between-person variation as estimated by an empty longitudinal model, so we need to include its between-person and within-person effects; that is, we need to explicitly consider the person as context when examining the effect of negative mood. To do so we will use an alternative method of model specification for negative mood. In person-mean-centering (otherwise known as group-mean-centering when employed with clustered data), the person mean of the time-varying predictor is subtracted from the original time-varying predictor, such that the new time-varying predictor represents variation about one's own mean level. A predictor for person mean negative mood (centered near the grand mean of 6) is then included in the Level-2 model, as shown in Equation 7:

$$\begin{aligned}
\text{L1:}\quad & \text{Symptoms}_{di} = \beta_{0i} + \beta_{1i}(\text{DayStress}_{di}) + \\
& \qquad \beta_{2i}(\text{DayMood}_{di} - \text{PMMood}_i) + (Y_i)e_{di} + (O_i)e_{di} \\
\text{L2:}\quad & \beta_{0i} = \gamma_{00} + \gamma_{01}(\text{Older}_i) + \gamma_{02}(\text{PMStress}_i) + \\
& \qquad \gamma_{03}(\text{PMStress}_i{*}\text{Older}_i) + \gamma_{04}(\text{PMMood}_i) + (Y_i)U_{0i} + (O_i)U_{0i} \\
& \beta_{1i} = \gamma_{10} + \gamma_{11}(\text{Older}_i) + (Y_i)U_{1i} + (O_i)U_{1i} \\
& \beta_{2i} = \gamma_{20}. \qquad (7)
\end{aligned}$$

In the Level-1 equation, β_{2i} is the individual within-person effect of negative mood, which is then specified solely as a fixed effect (γ_{20}) at level 2. Because daily negative mood has been person-mean-centered, it contains only within-person variation, and its effect in the model (γ_{20}) is the pure within-person effect, regardless of whether person mean negative mood is in the model. Under person-mean-centering, the Level-2 effect of person mean negative mood (γ_{04}) becomes the total between-person effect (unlike under grand-mean-centering, in which the Level-2 effect of the person mean was the contextual effect). These parameters can be interpreted as follows. The significant between-person effect (γ_{04}) of 0.29 indicates that for every unit higher of person mean negative mood, 0.29 more symptoms are reported on average (for both age groups so far). The significant within-person effect (γ_{20}) of 0.05 indicates that for every unit more negative mood *than usual*, 0.05 more symptoms than usual are reported. The between-person effect is the effect of person mean negative mood not controlling for daily negative mood. The corresponding contextual effect is 0.24, which is the effect of person mean negative mood after controlling for daily negative mood. It also indicates that the between-person effect is 0.24 significantly greater than the within-person effect. The fact that the difference of the within-person and between-person effects (*i.e.*, the contextual effect) is significant indicates that the use of a composite effect for negative mood (*i.e.*, not formally differentiating within- and between-person variation) would not yield a correct estimate for either its within-person or between-person effects.

To examine whether these effects differ between younger and older adults, we next add interactions with age to the model for the means and a Level-2 random slope for the effect of daily mood in the model for the variances, as shown in Equation 8:

$$
\begin{aligned}
\text{L1: } & \text{Symptoms}_{di} = \beta_{0i} + \beta_{1i}(\text{DayStress}_{di}) + \\
& \qquad \beta_{2i}(\text{DayMood}_{di} - \text{PMMood}_i) + (Y_i)e_{di} + (O_i)e_{di} \\
\text{L2: } & \beta_{0i} = \gamma_{00} + \gamma_{01}(\text{Older}_i) + \gamma_{02}(\text{PMStress}_i) + \gamma_{03}(\text{PMStress}_i{*}\text{Older}_i) + \\
& \qquad \gamma_{04}(\text{PMMood}_i) + \gamma_{05}(\text{PMMood}_i{*}\text{Older}_i) + (Y_i)U_{0i} + (O_i)U_{0i} \\
& \beta_{1i} = \gamma_{10} + \gamma_{11}(\text{Older}_i) + (Y_i)U_{1i} + (O_i)U_{1i} \\
& \beta_{2i} = \gamma_{20} + \gamma_{21}(\text{Older}_i) + (Y_i)U_{2i} + (O_i)U_{2i}. \qquad (8)
\end{aligned}
$$

Results for this final model are shown under the second column in Table 1. In the Level-1 equation, β_{2i} is the individual within-person effect of daily mood. The Level-2 equation for β_{2i} then specifies its effect for younger adults specifically (γ_{20}) and the cross-level interaction of how that effect differs for older adults (γ_{21}). The Level-1 within-person daily mood effect (γ_{20}) of 0.12 was significant for younger adults, but not significant for older adults (0.03), and the difference in the within-person effects (γ_{11}) of -0.09 was marginally significant. The Level-2 between-person effect for person mean mood for younger adults (γ_{04}) of 0.11 was not significant and was marginally smaller by 0.25 (γ_{05}) than the effect for older adults of 0.36, which was significant. The difference of the between-person and within-person effects, or contextual effect, of -0.003 was not significant for the younger adults, but the contextual effect of 0.33 was significant for the older adults, and significantly greater than that for the younger adults. To summarize, although having more stress than usual relates to more symptoms than usual in younger adults, the older adults did not show this same effect. However, greater mean levels of negative mood were related to greater mean levels of physical symptoms in older adults, but this between-person effect was not significant in younger adults. Older adults also showed a greater difference between the between-person and within-person effects of negative mood. Finally, we examined the cross-level interaction of daily mood and person mean mood, and whether this interaction differed by age (as would be represented on the Level-2 equation for β_{2i} by γ_{22} and γ_{23}). Neither effect was significant, indicating that the within-person effect did not depend on person mean mood, equivalently for both age groups. Notably, the inclusion of fixed effects for mood did not change the stress effects at either the within-or between-person level considerably (see Table 1). Thus, the associations between stress and physical symptoms were largely independent of negative mood at each level. Further, controlling for daily stress and negative mood rendered the age difference in physical symptoms nonsignificant. Thus, a portion of the

age differences in daily physical symptom reports can be attributed to age differences in stress and negative mood (see Selig & Preacher, this issue, for further explication of such mediational findings).

We then examined the role of mood in the model for the variances. At Level 2, a random deviation from the age group slope for daily negative mood was included (U_{2i}), again with different variances (τ_2^2) across groups, and was found to significantly improve model fit relative to the random intercept and daily stress slope model, $\Delta_\chi^2(6) = 20$, $p < 0.003$. We again computed 95% random effects confidence intervals to describe the size of the slope variation as the fixed slope ± 1.96*SQRT(τ_{2i}^2). Accordingly, 95% of the younger adults were expected to have within-person effects of mood from 0.19 fewer symptoms to 0.45 more symptoms per unit increase in negative mood relative to the person's mean, whereas older adults were expected have within-person effects of mood from 0.24 fewer symptoms to 0.29 more symptoms. Thus, the within-person effect of negative mood is only somewhat more variable in younger than older adults.

A Pure Within-Person Effect of a Time-Varying Predictor for Session

So far our model has included contextual and within-person effects of daily stress using grand-mean-centering, as well as between-person and within-person effects of daily negative mood using person-mean-centering. Our last model includes an example of a time-varying predictor that contains only within-person variation: "time" itself, or the session of measurement. In contrast to studies in which systematic change over time is of primary interest, however, "time" is included here only to control for reactivity—there is no inherent reason why physical symptoms would change over the course of 2 weeks in individuals who are otherwise healthy. Further, because time was balanced in this design (*i.e.*, everyone was measured on the same schedule), "time" contains only within-person variation— there is no between-person variation in session number. Because we do not need to worry about a potential between-person effect of time on physical symptoms, a linear trend for the within-person effect of time (centered such that zero is the first session) can be included directly in the Level-1 model, as shown in Equation 9:

$$
\begin{aligned}
\text{L1:}\quad & \text{Symptoms}_{di} = \beta_{0i} + \beta_{1i}(\text{DayStress}_{di}) + \beta_{2i}(\text{DayMood}_{di} - \text{PMMood}_i) + \\
& \qquad \beta_{3i}(\text{Time}_{di}) + (Y_i)e_{di} + (O_i)e_{di} \\
\text{L2:}\quad & \beta_{0i} = \gamma_{00} + \gamma_{01}(\text{Older}_i) + \gamma_{02}(\text{PMStress}_i) + \gamma_{03}(\text{PMStress}_i{*}\text{Older}_i) + \\
& \qquad \gamma_{04}(\text{PMMood}_i) + \gamma_{05}(\text{PMMood}_i{*}\text{Older}_i) + (Y_i)U_{0i} + (O_i)U_{0i} \\
& \beta_{1i} = \gamma_{10} + \gamma_{11}(\text{Older}_i) + (Y_i)U_{1i} + (O_i)U_{1i} \\
& \beta_{2i} = \gamma_{20} + \gamma_{21}(\text{Older}_i) + (Y_i)U_{2i} + (O_i)U_{2i} \\
& \beta_{3i} = \gamma_{30} + \gamma_{31}(\text{Older}_i). \qquad (9)
\end{aligned}
$$

Results for this final model are shown under the third column in Table 1. In the Level-1 equation, β_{3i} is the individual linear effect of time (session). The Level-2 equation for the within-person effect of time specifies the effect for younger adults (γ_{30}) and the cross-level interaction of how that effect differs for older adults (γ_{31}). The effect of time (γ_{30}) of -0.21 was significant for younger adults, significant for older adults (-0.04), and the difference in the time effects (γ_{31}) of 0.17 was also significant. A random deviation from the age group slope for time was then included (U_{3i}), but the matrix of random effects variances and covariances was not positive definite upon doing so, and thus the random slope for time was not retained. Although the intercepts are now conditional on time, and thus represent expected symptoms at the first occasion, controlling for a fixed linear trend for time resulted in similar effect sizes for the other predictors. The largest change was for the within-person effect of daily stress in the younger adults (from 0.72 to 0.59), but the effect was still significant.

Summary of Possible Effects of Time-Varying Predictors

Figure 2 shows the three different possible effects of the time-varying predictor of negative mood. For simplicity we plot only the effects for older adults (shown in the 3rd column of Table 1), and all nonpictured model effects have been set to their centered zero values. In Figure 2, the separate lines show the within-person effect: the expected difference in daily symptoms as a function of within-person variation in daily negative mood for persons with a mean negative mood of 5, 6 (the centering point), or 7. The x-axis shows daily negative mood ± 2 of the person mean, and the y-axis shows predicted symptoms. The within-person slope of 0.02 for the individual lines was not significant and did not vary across person mean negative mood. Thus, there is no effect on daily symptoms of being in a worse mood than usual in the older adults. The slope of the line connecting the person means (0.38) was significant and is the between-person effect: the expected difference in mean symptoms as a function of between-person variation in person mean negative mood. Thus, persons with worse negative mood (on average) have more symptoms (on average). The difference between these slopes of 0.36 is the contextual effect: the effect on mean symptoms of between-person variation person mean mood after controlling for current mood. The context effect is represented by the vertical difference between the lines for each person at the same value on the *x*-axis for negative mood.

The alternative methods of modeling a time-varying predictor can be summarized as follows. Under grand-mean-centering (*i.e.*, subtracting a constant from the time-varying predictor), its Level-1 effect is the composite effect when included by itself: it represents a weighted blend of its

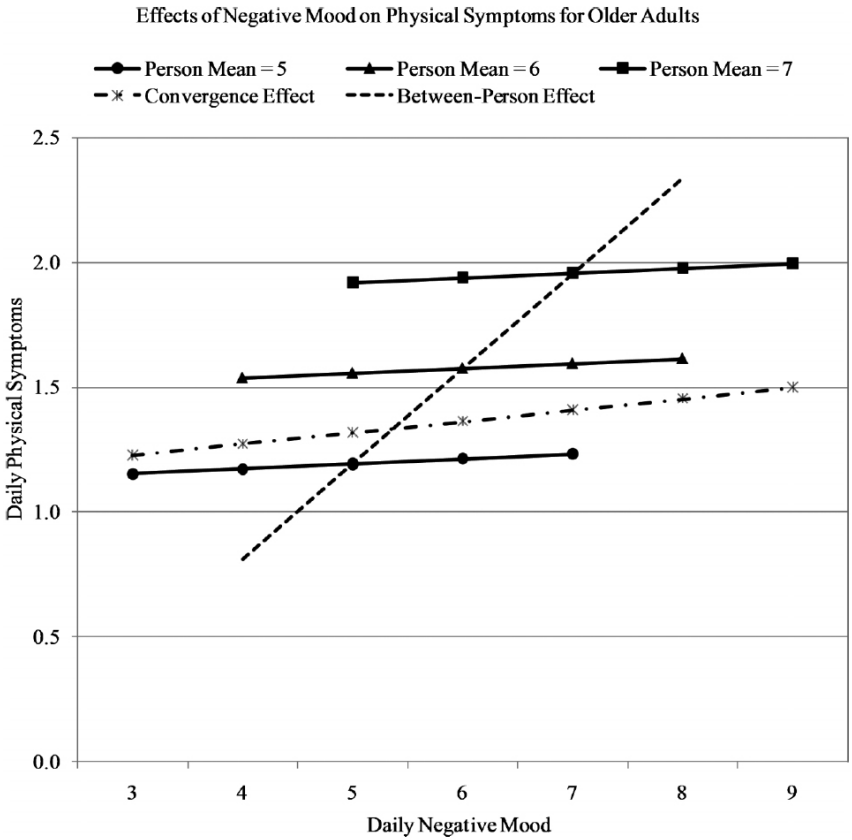

FIGURE 2 Illustration of within-person, between-person, and contextual effects of negative mood for older adults. Observed daily negative mood is on the x-axis, with values corresponding to person-mean-centered versions of negative mood of –2, –1, 0, 1, and 2 for each level of mean negative mood. A weak positive within-person effect is observed (slope of individual lines), along with a stronger positive between-person effect (slope of person means) and a positive contextual effect (distance between lines at the same point on the x-axis).

within-person and between-person effects because the time-varying predictor contains both within-person and between-person variation. Once the person mean of the time-varying predictor is included with the grand-mean-centered time-varying predictor, the Level-1 effect of the time-varying predictor is its within-person effect. The Level-2 effect of the person mean predictor is its contextual effect, or the incremental influence of between-person variation controlling for the current value of the time-varying predictor. In contrast, under person-mean-centering (*i.e.*, subtracting the person mean from the time-varying predictor), the Level-1 effect of

the new time-varying predictor is the within-person effect because it only contains within-person variation, regardless of whether the person mean predictor is included. The Level-2 effect of the person mean predictor is its between-person effect, or the total influence of between-person variation, not controlling for current value of the time-varying predictor.

Although Figure 2 illustrates the interpretation of the between-person and within-person effects, it does not explain why they are often of such different magnitudes. As discussed earlier, one cause for this is that the reasons why persons differ from one another may not be the same reasons why any given person varies from one point in time to another. Thus, what predicts within-person variability may be completely different than what predicts between-person variability because they reflect different theoretical constructs. Yet a simpler reason why between-and within-person effects are likely to differ from each other is that they are unstandardized regression coefficients on different scales. In linear regression, standardized coefficients can be calculated as $\beta_{\text{unstandardized}} * \text{SD(X)/SD(Y)}$. There is no clear analog in multilevel models with random slopes because there are multiple interdependent variance components to consider. One approach could be to use the standard deviations at each level of analysis to calculate the standardized coefficients so that they can be compared directly. Equation 10 shows how one might do so for the effects of negative mood in older adults:

$$\begin{aligned}
\text{WP}_{\text{std}} &= \beta_{\text{unstd}} * \frac{\text{SD}(\text{X}_{\text{di}})}{\text{SD}(\text{Y}_{\text{di}})} = \text{WP}_{\text{unstd}} * \frac{\text{SD}(\text{DayMood}_{\text{di}} - \text{PMMood}_{\text{i}})}{\text{SD}(\sigma_{\text{e}}^2)} \\
&= .02 * \frac{1.372}{.797} = .034 \\
\text{BP}_{\text{std}} &= \beta_{\text{unstd}} * \frac{\text{SD}(\text{X}_{\text{i}})}{\text{SD}(\text{Y}_{\text{i}})} = \text{BP}_{\text{unstd}} * \frac{\text{SD}(\text{PMMood}_{\text{i}})}{\text{SD}(\tau_{\text{U}_{0\text{i}}}^2)} \\
&= .38 * \frac{1.233}{1.023} = .461. \qquad (10)
\end{aligned}$$

Although we used the original SD of the predictors, we used the variance components for the outcome from a model not including mood (not shown). We used unaccounted for variation, not the total original variation, to represent all the variance currently available to be captured by between-person and within-person variation. Many reasonable alternative approaches to standardization are possible—our calculations here serve only to illustrate the difference in the scales of these effects. As shown in Equation 10, the standardized within-person and between-person effects are 0.034 and 0.461, respectively. Thus in this example, even when corrected onto the same scale, these effects still differ from each other. In

summary, to the extent that variation in the predictors and the outcomes is not equally distributed across the between-person and within-person levels, effects at each level are likely to differ numerically simply for this reason, above and beyond any more theoretical sources for the differences in these effects.

Finally, we note that although grand-mean-centering and person-mean-centering provide direct estimates of the fixed within-person effect, their corresponding random effects are not equivalent. Specifically, under grand-mean-centering, the random effect of a time-varying predictor is based on all of its variation, whereas under person-mean-centering the random effect is based on only its within-person variation. In choosing between these options one should incorporate substantive considerations about the source of the random effect (*i.e.*, whether the effect of more stress or more stress than usual should be allowed to vary over persons) as well as empirical considerations (*i.e.*, these represent non-nested models whose fit can be compared via ML information criteria).

Structural Equation Modeling and Other Approaches

The example analysis thus far has been presented via multilevel modeling. One might wonder, given the isomorphism of many multilevel and structural equation models (SEM: Bauer, 2003; Curran, 2003), can one estimate these same models in SEM? The answer is yes and no. The problem of how to separate between-person from within-person effects in SEM has received very limited attention (but see Curran, Lee, & MacCallum, 2007). As shown earlier, the person mean of the time-varying predictor must be included to distinguish its between- and within-person effects. Simply put, there is no direct way to include the person mean in a single-level SEM because it is redundant with the original time-varying predictor (as represented by separate variables across occasions). Further, if correlations of the time-varying predictor with the random intercept are included, its estimated effect will be the within-person effect as desired; if not, its effect will be the composite (blended) effect instead (Curran *et al.*, 2007).

There are (at least) two possible alternative specifications of a time-varying predictor in SEM to obtain its between- and within-person effects. One approach is a direct analog to the multilevel model in which person-mean-centered time-varying predictors are used instead of the original versions so that its person mean can then be included. An alternative approach suggested by Curran *et al.* (2007; see also Lüdtke *et al.*, 2008) is to model a time-varying predictor as another longitudinal outcome by estimating a random intercept for it as well. The random intercept for the "predictor" would thus represent the model-based analog of the person mean, permit-

ting a disaggregation of the time-varying predictor's between- and within-person variance and thus of its between- and within-person effects.

The work of Lüdtke *et al.* (2008) suggests that under certain circumstances the model-based approach may have distinct advantages over an observed person mean approach. In the model-based approach, the "true" between-person variance is estimated by reducing the observed between-person variance by a factor of σ_e^2/n, or the residual variance divided by the number of occasions per person. Although the resulting disattenuated between-person effects may be stronger, the extent to which this is reasonable depends on whether one views the within-person variance as an appropriate metric with which to estimate and correct for the reliability of the person mean. Although in some settings the Level-1 measures may serve as exchangeable indices of a latent Level-2 construct, in other settings differential Level-1 heterogeneity may be expected instead, and would not simply be "unreliability." Clearly this is an area that will benefit from further research. A related concern is the extent to which the person mean is an adequate representation of between-person variance. For example, would one obtain the same inferences when asking about "usual" levels of stress directly (*i.e.*, as a time-invariant predictor) as would be obtained from an aggregation over time-varying measures of stress, which is then supposed to represent "usual" levels of stress? However, this is fundamentally a question of measurement, not analysis, and the extent to which person mean aggregation can adequately represent stable individual differences is likely to depend on the particular question under study.

Finally, we note that though multilevel or structural equation models including time-varying predictors can be a useful way to examine within-person processes, there are other ways in which these models can be extended to address other types of longitudinal questions and data. For example, questions about the lead- lag relationships among a series of variables can be addressed by incorporating lagged effects (*e.g.*, McArdle, 2001). Questions regarding the timing and cycling of interdependent phenomena about a point of equilibrium can be addressed by dynamic systems models (*e.g.*, Boker & Laurenceau, 2006). Questions about individual differences in the regularity of a time series can be addressed by mixed models that include random effects of time series parameters (*e.g.*, autoregressive, moving average; Rovine & Walls, 2006). In any given modeling approach, however, the need to differentiate effects across levels of analysis is always relevant. To the extent that any model does not explicitly distinguish the contributions of stable individual differences from those of within-person variation, the ability to make inferences that reflect purely within-person associations may be compromised.

CONCLUSION

The goal of the current article was to address the importance of distinguishing between-person effects from within-person effects (*i.e.*, the modeling of persons as contexts), and to provide a detailed illustration of how to do so in the framework of multilevel modeling. Failure to explicitly consider separate between- and within-person sources of variation when modeling repeated measures (*e.g.*, daily diary) data can lead to biased results and potentially incorrect conclusions about within-person relationships over time. As our example has shown, the multilevel model is a useful and important tool with which to unpack the complex structure of variability inherent in multivariate longitudinal data. We contend that the benefits of longitudinal data for examining within-person processes can only be realized through the judicious use of models that formally distinguish the impact of between-person variation from that of within-person variation. Otherwise we are likely to be unable to quantify either between-person or within-person processes accurately. Although such formal modeling of persons as contexts is no doubt more complicated, it is ultimately a necessary and useful endeavor, and we hope our examples of how to do so using multilevel modeling can help guide researchers in this process

ACKNOWLEDGMENT

The authors wish to thank Marty Sliwinski (and NIA/NIH AG-124448) for the use of his data in the illustrative example.

REFERENCES

Almeida, D. M., Wethington, E., & Kessler, R. C. (2002). The Daily Inventory of Stressful Experiences (DISE): an interview-based approach for measuring daily stressors. *Assessment*, 9, 41-55.

Bauer, D. J. (2003). Estimating multilevel linear models as structural equation models. *Journal of Educational and Behavioral Statistics*, 28(2), 135-167.

Boker, S. M., & Laurenceau, J. P. (2006). Dynamical systems modeling: An application to the regulation of intimacy and disclosure in marriage. *In* T. A. Walls & J. L. Schafer (Eds.), *Models for Intensive Longitudinal Data* (pp. 195-218). New York: Oxford University Press.

Bolger, N., Davis, A., & Rafaeli, E. (2003). Diary methods: Capturing life as it is lived. *Annual Review of Psychology*, 54, 579-616.

Bronfenbrenner, U. (1979). *The Ecology of Human Development: Experiments by Nature and Design*. Cambridge. MA: Harvard University Press.

Curran, P. J. (2003). Have multilevel models been structural equation models all along? *Multivariate Behavioral Research*, 38(4), 529-569.

Curran, P. J., Lee, T.-H., MacCallum, R. A. (2007 October). Disaggregating within-person and between-person effects in latent curve models. Paper presented at the 2007 meeting of the Society of Multivariate Experimental Psychology, Chapel Hill, NC.

Diggle, P., Heagerty, P., Liang, K.-Y., & Zeger, S. (2002). *Analysis of Longitudinal Data*. New York: Oxford University Press.

Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods*, 12(2), 121-138.

Hedeker, D., & Gibbons, R. D. (2006). *Longitudinal Data Analysis*. New York: Wiley.

Hofer, S., & Sliwinski, M. (2001). Understanding ageing: An evaluation of research designs for assessing the interdependence of age-related changes. *Gerontology*, 47, 341-352.

Hoffman, L. (2007). Multilevel models for examining individual differences in within-person variation and covariation over time. *Multivariate Behavioral Research*, 42(4), 609-629.

Hoffman, L., & Rovine, M. J. (2007). Multilevel models for experimental psychologists: Foundations and illustrative examples. *Behavior Research Methods*, 39(1), 101-117.

Hofmann, D. A., & Gavin, M. B. (1998). Centering decisions in hierarchical linear models: Implications for research in organizations. *Journal of Management*, 24(5), 623-641.

Larsen, R. J., & Kasimatis, M. (1991). Day-to-day physical symptoms: Individual differences in the occurrence, duration, and emotional concomitants of minor daily illnesses. *Journal of Personality*, 59, 387-423.

Lawton, M. P., Kleban, M., Dean, J., Rajagopal, D., & Parmelee, P. (1992). The factorial generality of brief positive and negative affect measures. *Journal of Gerontology: Psychological Sciences*, 47, 228-237.

Lüdtke, O., Marsh, H. W., Robitzsch, A., Trautwein, U., Asparouhov, T., & Muthén, B. (2008). The multilevel latent covariate model: A new, more reliable approach to group-level effects in contextual studies. *Psychological Methods*, 13(3), 203-229.

McArdle, J. J. (2001). A latent difference score approach to longitudinal dynamic structural analysis. *In* R. Cudeck & S. Du Toit (Eds.), *Structural Equation Modeling: Present and Future* (pp. 341-380). Lincolnwood, IL: Scientific Software International.

Nesselroade, J. R. (1991). The warp and woof of the developmental fabric. *In* R. Downs, L. Liben, & D. S. Palermo (Eds.), *Visions of Aesthetics, the Environment, and Development: The Legacy of Joachim F. Wohlwill* (pp. 213-240) Hillsdale, NJ: Erlbaum.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical Linear Models* (2nd ed.). Thousand Oaks, CA: Sage.

Rovine, M. J., & Walls, T. A. (2006) Multilevel autoregressive modeling of interindividual differences in the stability of a process. *In* T. A. Walls & J. L. Schafer (Eds.), *Models for Intensive Longitudinal Data* (pp. 124-147). New York: Oxford University Press.

Schaie, K. W. (1995). *Developmental Influences on Adult Intelligence: The Seattle Longitudinal Study*. New York: Oxford University Press.

Schwartz, J. E., & Stone, A. A. (1998). Strategies for analyzing ecological momentary assessment data. *Health Psychology*, 17(1), 6-16.

Selig, J. P., & Preacher, K. J. (2009). Mediation models for longitudinal data in developmental research. *Research in Human Development*, 6(2-3), 144-164.

Singer, J. D., & Willett, J. B. (2003). *Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence*. New York: Oxford University Press.

Sliwinski, M. J. (2008). Measurement-burst designs for social health research. *Social and Personality Psychology Compass*, 2, 245-261.

Sliwinski, M. J., Smyth, J. M., Hofer, S. M., & Stawski, R. S. (2006). Intraindividual coupling of daily stress and cognition. *Psychology and Aging*, 21, 545-557.

Snijders, T. A. B., & Bosker, R. (1999). *Multilevel Analysis*. Thousand Oaks, CA: Sage.

Stawski, R. S., Sliwinski, M. J., Almeida, D. M., & Smyth, J. M. (2008). Reported exposure and emotional reactivity to daily stressors: The roles of adult-age and global perceived stress. *Psychology and Aging*, 23, 52-61.