Psychological Methods
2015, Vol. 20, No. 1, 63–83

1082-989X/15/$12.00 http://dx.doi.org/10.1037/met0000030

On Disaggregating Between-Person and Within-Person Effects With Longitudinal Data Using Multilevel Models

Lijuan (Peggy) Wang and Scott E. Maxwell
University of Notre Dame

This article extends current discussion of how to disaggregate between-person and within-person effects with longitudinal data using multilevel models. Our main focus is on the 2 issues of centering and detrending. Conceptual and analytical work demonstrates the similarities and differences among 3 centering approaches (no centering, grand-mean centering, and person-mean centering) and the relations and differences among various detrending approaches (no detrending, detrending *X* only, detrending *Y* only, and detrending both *X* and *Y*). Two real data analysis examples in psychology are provided to illustrate the differences in the results of using different centering and detrending methods for the disaggregation of between- and within-person effects. Simulation studies were conducted to further compare the various centering and detrending approaches under a wider span of conditions. Recommendations of how to perform centering, whether detrending is needed or not, and how to perform detrending if needed are made and discussed.

Keywords: within-person, between-person, longitudinal, centering, detrending

Understanding relations between variables is of fundamental importance in psychological research. As far back as Galton (1865), psychology has studied the extent to which people who are above average on one variable are also above average on another variable. This type of relation is referred to as a *between-person* relation and can be studied with cross-sectional data. However, cross-sectional data cannot answer questions related to within-person processes (Curran & Bauer, 2011). For example, consider the relation between mental health and telling white lies. The between-person relation reflects the extent to which individuals who tell more white lies differ in mental health from individuals who tell fewer white lies. In contrast, the within-person relation reflects the extent to which an individual is healthier when he or she tells fewer (or more) white lies than when he or she does not. Neither can generally be inferred from the other.

Methodologists have recently emphasized the importance of distinguishing within-person relations from between-person relations. For instance, Molenaar (2004) and Molenaar and Campbell (2009) theoretically and empirically showed that results from between-person analyses can be generalized to within-person findings only when certain strict assumptions (e.g., the ergodic conditions) are met. Curran and Bauer (2011), Hoffman and Stawski (2009), and Zhang and Wang (2014) have all pointed out that between-person relations are different from within-person relations conceptually and empirically. Not only could they have different magnitudes, but, in some cases, the two types of relations may even have different directions. Borrowing an example in Curran and Bauer, people who exercise more tend to have a lower risk of heart attack (negative between-person effect), whereas one is more likely to experience a heart attack while exercising (positive within-person effect). It is also possible that one of the effects exists whereas the other does not (e.g., see the first real data analysis example presented later in this article). As stated in Curran, Lee, Howard, Lane, and MacCallum (2012), “either confounding or mis-attributing” these two effects could lead to misleading results. Therefore, it is generally necessary to disaggregate between- and within-person effects. Fortunately, with longitudinal data, we can study both between- and within-person relations and try to disaggregate the two types of relations when both the outcome and the predictor are time varying.

Disaggregating between- and within-person relations requires both design and analysis considerations. Suppose a researcher wants to study the between- and within-person relations between mental health and reporting white lies. More specifically, the researcher is interested in the extent to which telling white lies predicts mental health. The design requires a longitudinal study in which both mental health and telling white lies are measured repeatedly over time for each individual in the study. The analysis requires special consideration of two methodological issues: centering and detrending. The centering issue is relevant for the disaggregation, even when neither variable exhibits any trend over time, whereas the detrending issue is relevant only when at least one of the variables exhibits some trend over time. We now briefly introduce each of these two issues and then discuss them in detail later in the article.

Lijuan (Peggy) Wang and Scott E. Maxwell, Department of Psychology, University of Notre Dame.

This research was made possible through the support of Grant 23014 (“The Science of Honesty: Can Being More Honest Improve Health?”) from the John Templeton Foundation to Lijuan (Peggy) Wang. The opinions expressed in this article are those of the authors and do not necessarily reflect the views of the John Templeton Foundation. We sincerely thank the principal investigator of the grant, Anita Kelly, for her work on proposing the research questions, designing the project, and collecting the data for the second empirical example presented in the article.

Correspondence concerning this article should be addressed to Lijuan (Peggy) Wang, 118 Haggar Hall, Department of Psychology, University of Notre Dame, Notre Dame, IN 46556. E-mail: lwang4@nd.edu

Centering a variable means that a constant or a vector of constants has been subtracted from every value of the variable. In the context of disaggregation with longitudinal data, centering refers to possible redefinitions of the 0 point of the predictor variable in the presumed relation. As we describe and compare in detail later in the article, three options are generally possible in this context. First, the predictor can be uncentered, which simply means that the raw scores of the predictor are used. Second, the predictor can be grand-mean centered, which means that the grand mean of predictor scores after averaging across all individuals and all time points is subtracted from the predictor for each person and time point. Third, the predictor can be person-mean centered (also called group-mean centered), which means that the person-specific mean of the predictor, averaging across time points of a person but not across persons, is subtracted from the predictor for each time point of the person.

Detrending is the statistical operation of removing the trend from a time series. In the context of examining the relation between a predictor and an outcome, detrending refers to controlling for the effect of time while examining the relation between the two variables. Because there are multiple possible detrending options—such as detrending neither variable, detrending only the predictor, detrending only the outcome, and detrending both variables—two questions need to be discussed. First, is detrending always needed when there is a trend over time? Second, if detrending is needed, how should it be conducted? For example, suppose both mental health and telling white lies are likely to change linearly over the course of the study. In such a case, any apparent within-person relation between mental health and telling white lies might be a result of time, which raises the question of whether the effect of time on the relation between these two variables should be controlled for. When the time-varying predictor exhibits a trend over time, Curran and Bauer (2011) suggested that we should *detrend the time-varying predictor* via a two-step procedure. After detrending, the between-person effect is studied as the *conditional* relation between levels of the time-varying predictor at a certain time point and average levels of the outcome variable pooling over individuals. The within-person effect is studied as the mean relation between a person's time-specific deviation in the predictor (relative to the *change curve*) and the individual's time-specific outcome. A question a reader might want to ask here is this: Do we have to detrend the time-varying predictor or remove the time effect from the time-varying predictor to study within-person effects whenever the predictor may be related to time itself? We address this question later in the article. Further, Curran et al. (2012) and Hoffman and Stawski (2009) discussed how to use multilevel models to disaggregate between- and within-person effects. In their models, time is included as a covariate in the first-level model. They stated that it is for the purpose of *detrending the time-varying outcome*, not for detrending the time-varying predictor. For instance, the within-person effect is described as a measure of how the person- and time-specific predictor predicts the outcome "above and beyond systematic growth in the outcome over time" (Curran et al., 2012, p. 219). Curran and Bauer suggested detrending the predictor, whereas both Curran et al. and Hoffman and Stawski suggested detrending the outcome. Therefore, if detrending is needed, how to conduct detrending requires more discussion. For example, does including time in the first-level model as a covariate detrend the outcome or the predictor? We also discuss this later in this article.

Disaggregating between- and within- person effects can be done either in the multilevel modeling framework (e.g., Curran & Bauer, 2011; Curran et al., 2012; Hoffman & Stawski, 2009) or in the structural equation modeling framework (e.g., Curran et al., 2012). Although the disaggregation of between- and within-person effects has been demonstrated and discussed in the multilevel modeling framework, several issues related to how to conceptualize and how to quantify the two effects, as discussed in the foregoing, still complicate the practical implementation of multilevel modeling. A small review of 36 recent substantive papers that studied between- and within-person effects using multilevel models was conducted. The substantive papers were found by using PsycINFO to search journal articles in English published after 2009 that (a) included "within-person effect" in the abstract, (b) included "longitudinal study" as the methodology, and (c) cited Hoffman and Stawski and/or Curran and Bauer. Journals in which the articles were published included *Developmental Psychology*, *Psychology and Aging*, *Journal of Family Psychology*, and *Journal of Applied Psychology*. The survey revealed the use of various centering and detrending methods. For example, although most studies used person-mean centering on the time-varying predictors, one study did not perform centering, two studies used grand-mean centering, and one study used person-mean centering on the outcome. In terms of detrending actions, eight studies did not conduct any detrending. Among those with detrending and with the detrending methods described in detail, 13 studies followed Hoffman and Stawski to include time as a Level-1 predictor, and some claimed that it was for the purpose of detrending the outcome, whereas eight studies followed Curran and Bauer to detrend the time-varying predictor by using the two-stage procedure. Furthermore, eight studies included the within-person effect as a fixed effect only, nine studies included it as both a fixed-effect and a random effect, and the others did not describe whether the within-person effects were fixed only or both fixed and random. The review results were not surprising, because answers to questions such as whether detrending is always necessary and how to perform detrending (e.g., detrending the predictor, detrending the outcome, or adding time as a Level-1 covariate) when detrending is needed are still unresolved in the literature.

Therefore, in this article, we adopt the multilevel modeling approach and extend the discussion on methods for disaggregating between- and within-person effects (Curran & Bauer, 2011; Curran et al., 2012; Hoffman & Stawski, 2009). We focus on the following two issues: centering (how to perform centering and relations among different centering approaches) and detrending (detrending or not detrending, how to perform detrending if needed, and relations among different detrending approaches). We conduct analytical, empirical, and simulation analyses to clarify the differences and similarities of various centering and detrending approaches. We also consider the implications of the within-person effect being either fixed or random, because most previous work on centering and detrending has been based on an assumption that the time-varying predictor has a fixed effect. Our goal is to make recommendations to researchers on which approaches are most appropriate under various conditions.

In the remainder of the article, we first review three centering approaches (no centering, grand-mean centering, and person-mean centering) conceptually and algebraically to clarify the differences among these approaches, with the analysis of a real data set for illustration. After that, we switch to the issue of detrending by reviewing two recommended detrending approaches in the literature and discussing the relations between them. We then discuss whether detrending is always necessary and analytically compare various detrending approaches (no detrending, detrending the time-varying covariate only, detrending the outcome only, and detrending both variables) to show the relations among them under certain conditions. Reliability of the intraindividual mean is also briefly discussed. Another real data analysis is included to demonstrate the effects of different detrending approaches on estimates and inference of between- and within-person effects. Then we show the results from two simulation studies with a wider span of conditions considered. More specifically, the two simulation studies focused on comparing results from different centering approaches and different detrending approaches respectively, allowing within-person effects to be random in both studies. We conclude with our recommendations on issues of centering and detrending for disaggregating between- and within-person effects using multilevel models.

A Basic Multilevel Model for Disaggregating Between- and Within-Person Effects

Suppose data on two time-varying variables (Y and X) are collected repeatedly from N individuals. For example, Y might represent self-concept and X might represent mood. The question of interest is whether change in mood over time predicts change in self-concept over time, intraindividually. Using the original scales of both Y and X, a multilevel model can be fitted to the data to study the *composite* effect of the time-varying predictor X on the outcome variable Y. The model (*Model M1*) is

$$\begin{aligned} y_{it} &= \gamma_{0i} + \gamma_{1i}x_{it} + e_{it} \\ \gamma_{0i} &= \gamma_{00} + u_{0i}, \\ \gamma_{1i} &= \gamma_{10} + u_{1i} \end{aligned} \tag{1}$$

where y_{it} and x_{it} are the observed scores of Y and X of individual i at time t. Some assumptions regularly made on a multilevel model include $E(u_{0i}) = 0$, $E(u_{1i}) = 0$, $E(e_{it}) = 0$, $cov(u_{0i}, e_{it}) = 0$, and $cov(u_{1i}, e_{it}) = 0$. In the model, γ_{10} is the average overall or composite effect of X on Y, and u_{1i} is the deviation of individual i's effect, γ_{1i}, from the average, γ_{10}. Note that the overall effect γ_{10} from Model M1 is a combination of both between- and within-person effects, and thus between- and within-person effects are not disaggregated. Although we do not recommend the use of this model, we include it for the purpose of illustrating the consequences of using original scales (i.e., no centering) of the time-varying predictor and the outcome variable.

When disaggregation of the two effects is considered, we need to include the respective predictors of between-person effects and within-person effects in a model. Using the notation in Curran and Bauer (2011), the model can be expressed as

$$\begin{aligned} y_{it} &= \gamma_{0i} + \gamma_{1i}xw_{it} + e_{it} \\ \gamma_{0i} &= \gamma_{00} + \gamma_{01}xb_i + u_{0i}. \\ \gamma_{1i} &= \gamma_{10} + u_{1i} \end{aligned} \tag{2}$$

In the model, xb_i and xw_{it}, respectively, are the between- and within-person characteristics of variable X. It is worth noting that XB is a time-invariant predictor, whereas XW is a time-varying predictor. The between-person effect, γ_{01} is included in the model as a fixed effect only, whereas the within-person effect, γ_{1i}, is included in the model as both a fixed effect, γ_{10}, and a random effect, u_{1i}. In some cases, when the number of time points is relatively small and/or the variance of u_{1i} is small, the model in Equation 2 may have a convergence problem. One way to resolve the problem is to include only the fixed-effect part in the model for the within-person effect (Singer & Willett, 2003) by constraining u_{1i} to equal 0. Then the model is reduced to

$$\begin{aligned} y_{it} &= \gamma_{0i} + \gamma_{1i}xw_{it} + e_{it} \\ \gamma_{0i} &= \gamma_{00} + \gamma_{01}xb_i + u_{0i}. \\ \gamma_{1i} &= \gamma_{10} \end{aligned} \tag{3}$$

The reduced model assumes that there is homogeneity in the within-person effects across different individuals, which may or may not be true in reality and can be tested by testing H_0: $var(u_{1i}) = 0$ when the model in Equation 2 can be successfully fitted to the data (e.g., Ke & Wang, 2014). The model in Equation 3 is also the main model discussed in Curran and Bauer. Similarly, a reduced form of the model in Equation 1, *Model M1r*, is

$$\begin{aligned} y_{it} &= \gamma_{0i} + \gamma_{1i}x_{it} + e_{it} \\ \gamma_{0i} &= \gamma_{00} + u_{0i}. \\ \gamma_{1i} &= \gamma_{10} \end{aligned} \tag{4}$$

Centering When There Are No Trends in Either *X* or *Y*

Methods and Analytical Comparisons

Obtaining valid values for XB and XW in Equations 2 and 3 is the key to appropriately disaggregating between- and within-person effects. Raudenbush and Bryk (2002) discussed how to center the Level-1 predictor variable for disaggregating between- and within-group effects using a hierarchical linear model in the organizational research context. It was reasonable that they did not discuss the effects of time on the disaggregation, because they were referring to cross-sectional multilevel data. Other researchers (e.g., Enders & Tofighi, 2007; Hofmann & Gavin, 1998; Kreft, de Leeuw, & Aiken, 1995) have also provided detailed discussion of grand-mean centering and group-mean centering in the context of two-level cross-sectional multilevel models.

Let us start with a longitudinal scenario that most resembles the illustration in Chapter 5 of Raudenbush and Bryk (2002, pp. 134–142). That is, there are no linear or nonlinear trends over time in the time-varying variables of interest, X and Y. Then, the issue discussed in this section is, when there are no trends in either X or Y, how to appropriately perform centering to construct xb_i and xw_{it} such that γ_{01} and γ_{10} measure the between-person effect and the within-person effect defined as "the relation between average levels of X and average levels of Y pooling over individuals" and "the mean relation between a person's time-specific deviation in variable X (relative to the *overall level* of X for each person) and the individual's time-specific Y," respectively.

There are at least two approaches we can use to center X: grand-mean centering and person-mean centering (e.g., Curran & Bauer, 2011; Hoffman & Stawski, 2009; Raudenbush & Bryk, 2002). When the grand-mean centering approach is used, xb_i can be constructed by averaging across different time points of each person, such that $xb_i = \frac{\Sigma_t x_{it}}{T_i} = \bar{x}_{i.}$ and xw_{it} is calculated by the deviation of x_{it} from the grand mean, averaging across different individuals and different time points, such that $xw_{it} = x_{it} - \frac{\Sigma_i \Sigma_t x_{it}}{\Sigma T_i} = x_{it} - \bar{x}_{..}$. Therefore, the model in Equation 2 with the grand-mean centering approach, *Model M2*, can be rewritten as

$$\begin{aligned} y_{it} &= \gamma^g_{0i} + \gamma^g_{1i}(x_{it} - \bar{x}_{..}) + e^g_{it} \\ \gamma^g_{0i} &= \gamma^g_{00} + \gamma^g_{01}\bar{x}_{i.} + u^g_{0i}, \\ \gamma^g_{1i} &= \gamma^g_{10} + u^g_{1i} \end{aligned} \tag{5}$$

where the superscript g denotes grand-mean centered. Putting it in the composite form, we have

$$\begin{aligned} y_{it} = {} & \gamma^g_{00} - \gamma^g_{10}\bar{x}_{..} + \left(\gamma^g_{01} + \gamma^g_{10}\right)\bar{x}_{i.} + \gamma^g_{10}(x_{it} - \bar{x}_{i.}) \\ & + u^g_{0i} + u^g_{1i}(x_{it} - \bar{x}_{..}) + e^g_{it}. \end{aligned} \tag{6}$$

From Equation 6, we can see that the between-person effect is measured by $eff_{between} = \gamma^g_{01} + \gamma^g_{10}$, and the fixed-effects within-person effect is measured by $eff_{within} = \gamma^g_{10}$.

Alternatively, when the person-mean centering approach is used, xb_i is also constructed by the person mean of X such that $xb_i = \frac{\Sigma_t x_{it}}{T_i} = \bar{x}_{i.}$ and xw_{it} is calculated by the deviation of x_{it} from the person mean such that $xw_{it} = x_{it} - \bar{x}_{i.}$. Therefore the model in Equation 2 with the person-mean centering approach, *Model M3*, can be expressed as

$$\begin{aligned} y_{it} &= \gamma^p_{0i} + \gamma^p_{1i}(x_{it} - \bar{x}_{i.}) + e^p_{it} \\ \gamma^p_{0i} &= \gamma^p_{00} + \gamma^p_{01}\bar{x}_{i.} + u^p_{0i}, \\ \gamma^p_{1i} &= \gamma^p_{10} + u^p_{1i} \end{aligned} \tag{7}$$

where the superscript p denotes person-mean centered. Putting it in the composite form, we have

$$y_{it} = \gamma^p_{00} + \gamma^p_{01}\bar{x}_{i.} + \gamma^p_{10}(x_{it} - \bar{x}_{i.}) + u^p_{0i} + u^p_{1i}(x_{it} - \bar{x}_{i.}) + e^p_{it}. \tag{8}$$

From Equation 8, we can see that the between-person effect is measured by $eff_{between} = \gamma^p_{01}$, and the fixed-effects within-person effect is measured by $eff_{within} = \gamma^p_{10}$.

The models in Equations 5 and 7 can also be reduced to allow only the fixed-effects part of the within-person effects by removing u_{1i} from the models, and thus the reduced models, *Model M2r* and *Model M3r*, are

$$\begin{aligned} y_{it} &= \gamma^g_{0i} + \gamma^g_{1i}(x_{it} - \bar{x}_{..}) + e^g_{it} \\ \gamma^g_{0i} &= \gamma^g_{00} + \gamma^g_{01}\bar{x}_{i.} + u^g_{0i} \\ \gamma^g_{1i} &= \gamma^g_{10} \end{aligned} \tag{9}$$

for the grand-mean centering approach and

$$\begin{aligned} y_{it} &= \gamma^p_{0i} + \gamma^p_{1i}(x_{it} - \bar{x}_{i.}) + e^p_{it} \\ \gamma^p_{0i} &= \gamma^p_{00} + \gamma^p_{01}\bar{x}_{i.} + u^p_{0i} \\ \gamma^p_{1i} &= \gamma^p_{10} \end{aligned} \tag{10}$$

for the person-mean centering approach.

We summarize the mean structures and covariance structures of the models with different centering approaches in both full and reduced forms in Table 1. The mean structures of the models with the grand-mean and person-mean centering approaches always have the same specification after reparameterization, whereas that of the model without centering is generally different. The only exception is that when the fixed-effects between- and within-person effects share the same value, the mean structures of all three centering approaches are the same.

When the models are reduced, the covariance structures of the models with the grand-mean and person-mean centering approaches always have the same specification. Therefore, the reduced models with the grand-mean and person-mean centering approaches, Models M2r and M3r (Equations 9 and 10), are essentially equivalent models. This means that the two reduced models share the same model fit for every global model fit measure. In addition, the relations between the parameters from these two models are $eff_{between} = \gamma^g_{01} + \gamma^g_{10} = \gamma^p_{01}$, $eff_{within} = \gamma^g_{10} = \gamma^p_{10}$, $\gamma^g_{00} - \gamma^g_{10}\bar{x}_{..} = \gamma^p_{00}$, $var(e^g_{it}) = var(e^p_{it})$, and $var(u^g_{0i}) = var(u^p_{0i})$. When the models are in full form, the covariance structures of the models with the grand-mean and person-mean centering approaches are different unless $\bar{x}_{i.}$ are the same across individuals. However, (Laird & Ware, 1982, p. 966) have shown that when the mean structure is the same (i.e., the design matrix of the fixed-effects parameters is the same), the expectation of the maximum likelihood–based fixed-effects parameter estimates $\hat{\gamma}$ is γ, with any given known covariance structure $Cov(y_i)$. That is, $E(\hat{\gamma} \mid \text{any given known } Cov(y_i)) = \gamma$. Translated into our context, we will have $E(\hat{eff}_{between}) = E(\hat{\gamma}^g_{01} + \hat{\gamma}^g_{10}) = E(\hat{\gamma}^p_{01})$, $E(\hat{eff}_{within}) = E(\hat{\gamma}^g_{10}) = E(\hat{\gamma}^p_{10})$, and $E(\hat{\gamma}^g_{00} - \hat{\gamma}^g_{10}\bar{x}_{..}) = E(\hat{\gamma}^p_{00})$. Note that this does not mean that for a specific data set, the estimates of the fixed-effects between- and within-person effects are the same from the full models with grand-mean and person-mean centering. Instead, this means that the average fixed-effects estimates averaging across all possible samples are the same. Therefore, with a single real data set, we cannot observe this phenomenon. However, the equality that arises from averaging over multiple random samples can be shown via a simulation study. Furthermore, regardless of whether the model is full or reduced, the fixed-effects estimates from no centering are generally different from those with centering, unless the fixed-effects between- and within-person effects have the same value. Anyway, γ_{10} from the no centering model is conceptually different, which is neither the between-person effect nor the

Table 1
Mean Structures and Covariance Structures of Models With Different Centering Approaches

Variable	No centering	Grand-mean centering	Person-mean centering
Effect(s)	γ_{10}: Composite effect	$\gamma^g_{01} + \gamma^g_{10}$: Between-person effect γ^g_{10}: Within-person effect	γ^p_{01}: Between-person effect γ^p_{10}: Within-person effect
Mean structure: $E(y_i)$	$\gamma_{00} + \gamma_{10}x_{it}$	$\gamma^g_{00} - \gamma^g_{10}\bar{x}_{..} + (\gamma^g_{01} + \gamma^g_{10})\bar{x}_{i.} + \gamma^g_{10}(x_{it} - \bar{x}_{i.})$	$\gamma^p_{00} + \gamma^p_{01}\bar{x}_{i.} + \gamma^p_{10}(x_{it} - \bar{x}_{i.})$
Covariance structure (full): $Cov(y_i)$	$Cov(u_{0i} + u_{1i}x_{it} + e_{it})$	$Cov(u^g_{0i} - u^g_{1i}\bar{x}_{..} + u^g_{1i}x_{it} + e^g_{it})$	$Cov(u^p_{0i} + u^p_{1i}(x_{it} - \bar{x}_{i.}) + e^p_{it})$
Covariance structure (reduced): $Cov(y_i)$	$Cov(u_{0i} + e_{it})$	$Cov(u^g_{0i} + e^g_{it})$	$Cov(u^p_{0i} + e^p_{it})$

Table 2
Results From the Six Models Fitted to the Self-Concept Data

M1 (Equation 1): Composite full		M2 (Equation 5): G-center full		M3 (Equation 7): P-center full		M1r (Equation 4): Composite reduced		M2r (Equation 9): G-center reduced		M3r (Equation 10): P-center reduced	
Fixed-effects parameters											
b_{00}	2.448 (0.633)	γ_{00}^{g}	12.485 (1.771)	γ_{00}^{p}	7.028 (1.605)	b_{00}	2.446 (0.569)	γ_{00}^{g}	12.065 (1.736)	γ_{00}^{p}	6.778 (1.657)
b_{10}	0.421 (0.045)	γ_{01}^{g}	−0.436 (0.149)	γ_{01}^{p}	0.029 (0.136)	b_{10}	0.422 (0.042)	γ_{01}^{g}	−0.401 (0.147)	γ_{01}^{p}	0.051 (0.140)
		γ_{10}^{g}	0.450 (0.047)	γ_{10}^{p}	0.451 (0.049)			γ_{10}^{g}	0.452 (0.044)	γ_{10}^{p}	0.452 (0.044)
Random-effects variances											
$var(u_{0i})$	4.208 (3.646)	$var(u_{0i}^{g})$	1.608 (0.557)	$var(u_{0i}^{p})$	1.536 (0.529)	$var(u_{0i})$	1.975 (0.649)	$var(u_{0i}^{g})$	1.521 (0.529)	$var(u_{0i}^{p})$	1.521 (0.529)
$var(u_{1i})$	0.008 (0.020)	$var(u_{1i}^{g})$	0.007 (0.018)	$var(u_{1i}^{p})$	0.014 (0.020)						
$var(e_{it})$	5.856 (0.535)	$var(e_{it}^{g})$	5.838 (0.529)	$var(e_{it}^{p})$	5.792 (0.521)	$var(e_{it})$	5.957 (0.506)	$var(e_{it}^{g})$	5.947 (0.504)	$var(e_{it}^{p})$	5.947 (0.504)
Overall model fit											
$-2LL$	1,483.6	$-2LL$	1,476.0	$-2LL$	1,474.6	$-2LL$	1,484.2	$-2LL$	1,477.4	$-2LL$	1,477.4

Note. Values in the parentheses are standard error estimates. M = model; G-center = grand-mean centering approach; P-center = person-mean centering approach; $-2LL = -2$ log likelihood value.

within-person effect but the composite effect without disaggregation. As Raudenbush and Bryk (2002) noted, the composite effect is "generally an uninterpretable blend" of between- and within-person effects (p. 138). Therefore, centering is needed to achieve disaggregation.

Between the two approaches that implement centering, although the results are essentially the same from the reduced models, M2r and M3r, echoing the suggestions of Curran and Bauer (2011), we recommend using the person-mean centering approach (i.e., recommend M3r over M2r) because (a) it disaggregates between- and within- person effects and (b) it is easier to conduct statistical inference on both fixed-effects between- and within-person effects directly from modeling results (a single model parameter for each effect). When the models are full (i.e., allow between-person differences in within-person effects), the expected values of the fixed-effects between- and within-person effects are the same, respectively, from Models M2 and M3. For the variance component estimates, Raudenbush and Bryk (2002) showed that person-mean centering would generally yield more accurate variance estimates of within-person effects than grand-mean centering via "a heuristic illustration" (pp. 145–149) in the context of cross-sectional organizational research. More specifically, grand-mean centering would generally underestimate the variance of within-person effects because of the shrinkage in empirical Bayes estimates. We echo Raudenbush and Bryk and recommend person-mean centering over grand-mean centering (i.e., recommend M3 over M2). Because there are no analytic forms for the variance estimates, later in the article we compare the variance estimates from different centering approaches via a simulation study to provide further evidence to support the recommendation under the conditions of nonzero variances in within-person effects.

Real Data Analysis Illustration

In this subsection, we use data from an observational longitudinal study to empirically illustrate the differences and similarities in the results of applying different centering approaches when there are no linear or nonlinear trends in the variables. Using an experience-sampling design, working self-concepts and emotional states were repeatedly measured in 33 undergraduate students (23 female and 10 male) over a 3-week period. The numbers of time points ranged from one to 11 across the participants, and the median or mode was 10. The average interval between two time points was about 2 days. At each time point, participants reported their current mood and working self-concept. In this study, we use data from one of the mood indicators—the high-arousal positive mood (HAP)—and one of the self-concept variables—emotional stability—measured by the 16 Personality Factor Scales (Cattell, 1946) to study the relations between the two variables as an example for the illustration.[1]

No trends were detected (none of the change slopes were significant after multiple testing adjustment) on either of the two variables. This was expected as it was a short-term observational study. Therefore, six different models specified in Equations 1, 5, 7, 4, 9, and 10 were fitted to the data. Parameter estimates, standard error estimates, and model fit statistics from the six fitted models are displayed in Table 2. The reduced models did not fit significantly worse than the corresponding full models. Consistently, variances of u_{1i} were not significantly different from 0 in Models M1–M3, which implied that within-person effects did not differ significantly across individuals in this example. Therefore, in the following, the results from the reduced models are mainly discussed.

From Table 2, we can see that the effect of HAP on emotional stability estimated in Model M1r is a combination of both between- and within-person effects but not the simple sum of the two effects. M1r failed to disaggregate the two types of effects, and thus we do not have information on the sizes of either between-person or within-person effects from M1r. The estimated value (.422) of the combined effect from M1r was closer to the within-person effect estimates from Models M2r or M3r (.452) than the between-person effect estimates (.051). This is because X

[1] We are grateful to Dr. Jesse Pappas for his generous sharing of the data.

was more similar to *XW* than *XB* in this example (the correlation between *X* and *XW* was .87, whereas the correlation between *X* and *XB* was .51, pooling across all individuals and all time points).

From both Model M2r and Model M3r results, we can see that the average within-person effect was different from the between-person effect in magnitude (.452 vs. .051) and significance (significant effect vs. nonsignificant effect). In addition, the difference test of average within-person effect versus between-person effect was significant, $t(df = 277) = 2.73$, $p = .007$. Periods of more positive mood are significantly associated with periods of higher emotional stability within people, but there is no evidence that more positive individuals have higher emotional stability. The results demonstrate that between- and within-person effects are two distinct concepts and bring evidence to the need of modeling both of them. The relations in parameter estimates and fit indexes between Models M2r and M3r described earlier are also verified. For example, $\hat{eff}_{between} = \hat{\gamma}^{g}_{01} + \hat{\gamma}^{g}_{10} = -.401 + .452 = \hat{\gamma}^{p}_{01} = .051$, and $\hat{eff}_{within} = \hat{\gamma}^{g}_{10} = \hat{\gamma}^{p}_{10} = .452$. Also notice that the estimated between- and within-person effects from full Models M2 and M3 are slightly different, such that for $\hat{eff}_{between}$, $\hat{\gamma}^{g}_{01} + \hat{\gamma}^{g}_{10} = -.436 + .450 \neq \hat{\gamma}^{p}_{01} = .029$, and for $\hat{eff}_{within}$, $\hat{\gamma}^{g}_{10} = .450 \neq \hat{\gamma}^{p}_{10} = .451$. Although the exact relations in parameter estimates do not hold for Models M2 and M3, the estimates were not very different because the expected values are the same as discussed earlier.

For variance components, Models M2r and M3r yielded exactly the same estimates. Models M2 and M3, however, had different results. For example, the variance estimate of u_{1i} from grand-mean centering (Model M2) was .007, half of that from person-mean centering (.014 [Model M3]). This is consistent with the "homogenization under grand-mean centering" observation by Raudenbush and Bryk (2002, p. 146).

In terms of fit indexes, Models M2r and M3r had exactly the same model fit ($-2LL$ [-2 log likelihood value] = 1,477.4) as predicted in the Methods and Analytical Comparisons section, whereas Model M1r had a slightly different model fit ($-2LL$ = 1,484.2) from those of Models M2r and M3r. This result, however, is not consistent with a statement[2] of equal model fit from three scalings (original scale, grand-mean centering, and person-mean centering) of the time-varying covariate in Curran et al. (2012). As discussed earlier, Model M1r has a different mean structure from Models M2r and M3r, such that M1r has x_{it} as its only predictor, whereas M2r and M3r have both x_{it} and $\bar{x}_{i.}$ as predictors. As long as $\bar{x}_{i.}$ explains a part of the variance in the outcome variable, the model fit from Model M1r should be different from the model fit for Models M2r and M3r. In this real data example, because $\bar{x}_{i.}$ explained a part of variance in the outcome variable, we observed a slight difference in the model fits between Model M1r and Models M2r and M3r. When the within-person effects had random effects—as in Models M1, M2, and M3—all three scalings of the TVC resulted in different model fits in this example, as shown in Table 2.

To summarize this section, we have illustrated that when there are no trends in either the time-varying predictor or the time-varying outcome, (a) centering of the time-varying predictor is necessary to disaggregate the between- and within-person effects, and (b) person-mean centering on the time-varying predictor is recommended.

Detrending When There Are Trends in Either *X* or *Y*

In this section, we consider situations with trends in either or both of the time-varying variables of interest: *X* and *Y*. For example, in the next real data analysis example, *X* is the number of white lies, and *Y* is self-reported mental health complaints. Both variables are time-varying and have trends over time. The issues now are whether and when detrending (i.e., controlling for the time effect) is necessary for studying between- and within-person effects and how to appropriately perform detrending to estimate the relations between *X* and *Y* net the time effect.

Two Proposed Detrending Models/Methods in the Literature

Curran and Bauer (2011) argued that the person-mean centering approach can be used under the condition of no growth in the time-varying predictor *X*. When there is a trend in *X*, they suggested a two-step approach for modeling between- and within-person effects. In the first step, regressions of the time-varying predictor *X* on grand-mean centered time are fitted to each individual. In the second step, the estimated intercepts and residuals from the case-based regressions are used as observed data for xb_i and xw_{it} in Equation 3. For example, use ordinary least squares (OLS) estimates $\hat{a}_{0i}$ and $\hat{r}_{xit}$ from $x_{it} = a_{0i} + a_{1i}(time_{it} - \bar{time}_{..}) + r_{xit}$ for the estimates of xb_i and xw_{it}. This idea is essentially the idea of detrending *X* before modeling. More specifically, $\hat{a}_{0i}$ is the estimated level of individual *i* at the grand-mean time $\bar{time}_{..}$, and $\hat{r}_{xit}$ is the detrended score of *X* for individual *i* at time *t*. Therefore, the two-stage model proposed by Curran and Bauer with linear detrending, *Model M4r*, can be expressed as

$$\begin{cases} x_{it} = a_{0i} + a_{1i}(time_{it} - \bar{time}_{..}) + r_{xit} \\ y_{it} = \gamma_{0i} + \gamma_{1i}[x_{it} - \hat{a}_{0i} - \hat{a}_{1i}(time_{it} - \bar{time}_{..})] + e_{it} \\ \gamma_{0i} = \gamma_{00} + \gamma_{01}\hat{a}_{0i} + u_{0i} \\ \gamma_{1i} = \gamma_{10} \end{cases}, \quad (11)$$

where γ_{01} measures the between-person effect and γ_{10} measures the within-person effect. The first equation in Equation 11 is the detrending step, and the last three equations form the multilevel model in the second step. Notice that Model M4r is in the reduced form (i.e., the within-person effect is included into the model only as a fixed effect, not a random effect). Via simulations, Curran and Bauer found that the between- and within-person effects were recovered with near-perfect accuracy in the balanced case but with only modest bias in the unbalanced case. When all the individuals have their average times equal to the grand mean of time, we have $\bar{time}_{i.} = \bar{time}_{..}$ (e.g., when the study is balanced and no missing data exist) and, thus, $\hat{a}_{0i} = \bar{x}_{i.}$. Then Model M4r can be rewritten as *Model M4r-a*:

[2] The statement is this: "Importantly, all three scalings of the TVC [time varying covariant] will result in precisely the same model fit (i.e., in terms of deviance and all deviance-based measures)" (Curran et al., 2012, p. 224).

$$\begin{cases} y_{it} = \gamma_{0i} + \gamma_{1i}(x_{it} - \bar{x}_{i.}) - \gamma_{1i}\hat{a}_{1i}(time_{it} - \bar{time}_{..}) + e_{it} \\ \gamma_{0i} = \gamma_{00} + \gamma_{01}\bar{x}_{i.} + u_{0i} \\ \gamma_{1i} = \gamma_{10} \end{cases}, \tag{12}$$

or

$$y_{it} = \gamma_{00} + \gamma_{01}\bar{x}_{i.} + \gamma_{10}(x_{it} - \bar{x}_{i.}) - \gamma_{10}\hat{a}_{1i}(time_{it} - \bar{time}_{..}) + u_{0i} + e_{it}. \tag{13}$$

Comparing Equations 12 and 10 reveals that the difference between the person-mean centering approach (Model M3r) and the two-step detrending X approach (Model M4r-a) when $\bar{time}_{i.} = \bar{time}_{..}$ is the inclusion or exclusion of the time effect of the time-varying predictor. Notice that the correlation of $\bar{x}_{i.}$ with $\hat{a}_{1i}(time_{it} - \bar{time}_{..})$ is 0 in this case. Therefore, point estimates of the between-person effect, γ_{01}, should be identical from models M3r and M4r-a regardless of whether the time effect is included. Similarly, when $\bar{time}_{i.} = \bar{time}_{..}$, and thus $\hat{a}_{0i} = \bar{x}_{i.}$, the estimates of the intercept, γ_{00}, should also be identical from the two models because both intercepts are intercepts at $\bar{time}_{..}$.

With regard to the estimate of the within-person effect, the estimates of γ_{10} are different from the two models when at least one $\hat{a}_{1i}$ is nonzero. The estimate of the within-person effect from the two-step detrending approach (Model M4r-a) could be smaller or larger than the estimate from the person-mean centering approach (Model M3r). Let $cx_{it} = (x_{it} - \bar{x}_{i.})$ and $dx_{it} = x_{it} - \bar{x}_{i.} - \hat{a}_{1i}(time_{it} - \bar{time}_{..})$. When $\frac{r_{y.cx}}{s_{cx}} > \frac{r_{y.dx}}{s_{dx}}$ ($r_{y.cx}$ and $r_{y.dx}$ are the correlations of Y with cX and of Y with dX, and s_{cx} and s_{dx} are sample standard deviations of cX and dX across all time points and all individuals, respectively), the estimate of the fixed-effects within-person effect from the person-mean approach is larger than the estimate from the two-step detrending approach, or vice versa.

Although Curran and Bauer (2011) proposed the two-stage model in Equation 11, from Equation 12 we can see that we should be able to implement their suggested approach in a single multilevel model by including the time variable as a Level-1 covariate with both fixed and random effects into the multilevel model under the conditions of $\bar{time}_{i.} = \bar{time}_{..}$ and linear detrending. The one-step implementation is beneficial in that we do not bring measurement errors of estimates of $\hat{a}_{1i}$ (Willett, 1989) into the modeling. Instead, we treat a_{1i} as a random effect in the multilevel model, and thus a_{1i}s is measurement free. Note that $\hat{a}_{0i} = \bar{x}_{i.}$ is still not measurement-error free, which is discussed in a following subsection.

Curran and Bauer (2011) did not discuss how a trend in Y influences the disaggregation and whether we should detrend Y for the study of between- and within-person effects when Y has a trend. In an earlier article, Hoffman and Stawski (2009) suggested including time as a Level-1 covariate into the multilevel model for the purpose of controlling for reactivity in the *outcome* variable in the context of observational studies. They stated the rationale of detrending the outcome variable or removing unexpected trend from the outcome variable: "There is no inherent reason why physical symptoms would change over the courses of 2 weeks in individuals who are otherwise healthy" (Hoffman & Stawski, 2009, p. 113), where physical symptoms is the outcome of interest in the study. Curran et al. (2012) made a similar suggestion of including time as a Level-1 covariate into a multilevel model for obtaining within- and between-person effects that are "net the contribution of the linear effect of time" (p. 224). The idea is to "capture within-person effects by directly predicting the repeated measures above and beyond systematic growth in the outcome over time" from time-varying covariates (Curran et al., 2012, p. 219). A reduced form of the models fitted in Curran et al. or Hoffman and Stawski, *Model M5r*, can be written as

$$\begin{aligned} y_{it} &= \gamma_{0i} + \gamma_{1i}(x_{it} - \bar{x}_{i.}) + \gamma_{2i}(time_{it} - \bar{time}_{..}) + e_{it} \\ \gamma_{0i} &= \gamma_{00} + \gamma_{01}\bar{x}_{i.} + u_{0i} \\ \gamma_{1i} &= \gamma_{10} \\ \gamma_{2i} &= \gamma_{20} + u_{2i} \end{aligned} \tag{14}$$

and the composite form is

$$y_{it} = \gamma_{00} + \gamma_{01}\bar{x}_{i.} + \gamma_{10}(x_{it} - \bar{x}_{i.}) + (\gamma_{20} + u_{2i})(time_{it} - \bar{time}_{..}) + u_{0i} + e_{it}. \tag{15}$$

where γ_{01} measures the conditional between-person effect at $\bar{time}_{..}$ and γ_{10} measures the *net* fixed-effects within-person effect.

The models in Equations 12 and 14, Models M4r-a and M5r, appear similar to each other. However, results from these two models may be different because (a) $\hat{a}_{1i}$s are included in Equation 12 as fixed values, whereas γ_{2i} or u_{2i} are random effects in Equation 14, and (b) $cov(u_{0i}, \gamma_{10}\hat{a}_{1i})$ in Equation 12 is 0, whereas $cov(u_{0i}, u_{2i})$ in Equation 14 is freely estimated, and thus its estimate can be nonzero. When an additional constraint, $cov(u_{0i}, u_{2i}) = 0$, is added to the model in Equation 14, the two models become more similar to each other, and the difference between the two models only lies in whether the linear slopes of time are treated as fixed effects or random effects. More specifically, for the two-step procedure, the linear slopes are treated as fixed effects, which contain measurement errors. For the one-step procedure using the model in Equation 14, the linear slopes are treated as random effects and, thus, are measurement-error free. Therefore, the model in Equation 14, Model M5r, will give us more accurate and efficient parameter estimates than the model in Equation 12, Model M4r-a. In addition, when the constraint $cov(u_{0i}, u_{2i}) = 0$ is included into the model in Equation 14, the estimates of the between-person effect γ_{01} and the intercept γ_{00} should also be identical to the estimates from the person-mean centering approach and the two-step detrending approach under the condition of $\bar{time}_{i.} = \bar{time}_{..}$ because, in this case, between- and within-person effects are orthogonal. Note that Model M5r itself does not require $\bar{time}_{i.} = \bar{time}_{..}$, so it can be used without this assumption. We make this assumption here simply for the purpose of easier comparison of different detrending models/methods.

Detrending or Not Detrending

Our current literature (e.g., Curran & Bauer, 2011; Hoffman & Stawski, 2009) suggests we should detrend X or Y when X or Y have trends over time. In this subsection, we discuss the issue of whether and when we need to detrend X or Y or both for the disaggregation of between- and within-person effects when trends in either X or Y or both exist.

We start our presentation with a simple hypothetical data set. Suppose we have data from a balanced longitudinal design study.

In the data set, both the outcome variable Y and the time-varying predictor X have linear changes over time such that

$$\begin{aligned} Y_{it} &= a_{y0i} + a_{y1i} time_t \\ X_{it} &= a_{x0i} + a_{x1i} time_t. \end{aligned} \tag{16}$$

Again, previewing our next real data analysis example, X is the number of white lies, and Y is the number of mental health complaints. Notice that in Equation 16, both Y and X change over time perfectly linearly. Although the scenario might not be realistic in the social sciences, this example with no measurement errors and no prediction errors simplifies the illustration, as we explain in the following.

An artificial data set with three individuals and five time points for each individual, generated from the model in Equation 16, is displayed in Table 3. For Y, we have $a_{y01} = a_{y02} = a_{y03} = 6$ and $a_{y11} = a_{y12} = a_{y13} = -1$. For X, we have $a_{x01} = 6$, $a_{x02} = 13$, $a_{x03} = 21$, $a_{x11} = -1$, $a_{x12} = -2$, and $a_{x13} = -3$. Both the outcome and predictor change over time for all three individuals. All the individuals have the same intercepts and slopes for the outcome variable Y, whereas for the time-varying predictor X, different individuals differ in both intercepts and linear slopes.

Because the manipulated data are so simple, from the raw data we can directly see that the average levels of Y ($\bar{y}_{i.}$) do not differ at all across different individuals. Thus, the average levels of the time-varying predictor X ($\bar{x}_{i.}$) do not relate to the average levels of outcome Y ($\bar{y}_{i.}$), pooling over individuals. Therefore, conceptually, there is no between-person effect in the data.

Within each individual (for a given i), we have the following observations: (a) X (x_{it}) is perfectly related to Y (y_{it}) because Y in our case is manipulated the same as the time variable, and X has a perfect linear relation with the time variable; (b) time-specific deviation in X relative to the individual overall level of X ($cx_{it} = x_{it} - \bar{x}_{i.}$) is perfectly related to Y (y_{it}); (c) time-specific deviation in X relative to the individual overall level of X ($cx_{it} = x_{it} - \bar{x}_{i.}$) is perfectly related to the time-specific deviation in Y relative to the individual overall level of Y ($y_{it} - \bar{y}_{i.}$); (d) time-specific deviation in X relative to the individual linear curve of X ($dx_{it} = x_{it} - \hat{a}_{x0i} - \hat{a}_{x1i} time_t$) is not related to Y (y_{it}), time-specific deviation in Y relative to the individual overall level of Y ($y_{it} - \bar{y}_{i.}$), or time-specific deviation in Y relative to the individual linear curve of Y ($dy_{it} = y_{it} - \hat{a}_{y1i} time_t$), because dx_{it} is 0; and (e) time-specific deviation in Y relative to the linear curve of Y (dy_{it}) is not related to any versions of X, because $dy_{it} = 0$. Because the between-person effect is 0 and there is no between-person variation in the within-person effect, the "composite" effect is equivalent to the within-person effect in this case. Therefore, the calculated within-person effects from different methods could change dramatically. For instance, there is a perfect within-person effect with the person-mean centering approach without detrending, whereas no within-person effect is detected using Curran and Bauer's (2011) two-step detrending approach or the one-step multilevel detrending approach described by Curran et al. (2012) and Hoffman and Stawski (2009).

The hypothetical example shows that detrending or not detrending could greatly influence the results on between- and within-person effects. Whether and when detrending is needed depends on how we define between-and within-person effects conceptually or what kinds of between- and within-person effects we intend to measure. For example, when we are interested in the within-person effect after controlling for the time effect, detrending is necessary. Sometimes, when the study is an experimental study, trends in both variables X and Y are caused by the experimental manipulation. In other words, without the experimental intervention, there are no trends in X and Y. If we detrended one of the variables or both, we would remove the purposefully designed experimental manipulation, which may not be consistent with the original intention of the researchers. For instance, if the trends in the hypothetical example come from the experimental manipulation (e.g., our next real data example is an experimental design), detrending either X or Y or both would artificially diminish the within-person effects from a perfect within-person relation to no within-person relation. In this case, detrending may not be necessary

Table 3
A Scenario With Growth In Both Y and X

			Raw data		P-centered X		Detrended X	Detrended Y
ID	Time	G-time	Y	X	$\bar{x}_{i.}$	$x_{it} - \bar{x}_{i.}$	dx_{it}	dy_{it}
1	1	−2	5	5	3	2	0	0
1	2	−1	4	4	3	1	0	0
1	3	0	3	3	3	0	0	0
1	4	1	2	2	3	−1	0	0
1	5	2	1	1	3	−2	0	0
2	1	−2	5	11	7	4	0	0
2	2	−1	4	9	7	2	0	0
2	3	0	3	7	7	0	0	0
2	4	1	2	5	7	−2	0	0
2	5	2	1	3	7	−4	0	0
3	1	−2	5	18	12	6	0	0
3	2	−1	4	15	12	3	0	0
3	3	0	3	12	12	0	0	0
3	4	1	2	9	12	−3	0	0
3	5	2	1	6	12	−6	0	0

Note. ID = participant number; G-time = the grand-mean centered time variable, $time - \overline{time}_{..}$; P-centered = the person-mean centering approach; $dx_{it} = x_{it} - \hat{a}_{x0i} - \hat{a}_{x1i} time_t$; $dy_{it} = y_{it} - \hat{a}_{y1i} time_t$.

to keep the natural feature of the study when studying relations between two variables.

In sum, detrending is not always as necessary as previous research has suggested. The decision depends on whether the researchers want to control for the effect of time when looking into the within-person relation between two variables. When the time effect is purposefully introduced by the study, one should preserve the time effects in the variables when studying the relations between the variables. In this case, detrending is not needed, and we recommend person-mean centering of the predictor variable for disaggregating between- and within-person effects. In contrast, when the time effect is a result of factors that are not relevant to the study design or are not of research interest, one may want to control for the time effect via a detrending approach. Basically, only when it is necessary to control for the time effect is detrending needed. In other cases, detrending may bring misleading results.

How to Detrend

When detrending is needed, how should one perform detrending to better understand the relation between *X* and *Y* net the contribution of the linear effect of time? Curran and Bauer (2011) suggested detrending *X* via a two-step approach, whereas Curran et al. (2012) and Hoffman and Stawski (2009) included time as a Level-1 covariate with multilevel modeling and stated that it is for detrending *Y*. It seems that the literature is still not clear on whether we should detrend *X* or detrend *Y* or maybe both. Therefore, in this subsection, we examine how different ways of doing detrending affect the estimates of within-person effects.

Detrending a variable essentially is removing or controlling for the linear or nonlinear time effect of the variable. For studying the relations between two variables *X* and *Y*, there could be four possible actions in terms of detrending: detrend neither, detrend *X* only, detrend *Y* only, or detrend both. Diagrams of different detrending actions for studying within-person effects are presented in Figure 1. There are two notes for appropriately reading Figure 1: (a) Figure 1 assumes that there is only a linear trend of time in *X* and *Y* and that *X* has been appropriately centered (i.e., person-mean centered) to simplify the illustrations. (b) The diagrams are not path diagrams in structural equation modeling (e.g., Bollen, & Curran, 2006; McArdle & McDonald, 1984) and are only illustrations for visually demonstrating our ideas. Figure 1a shows the no detrending approach (Action a), in which *Y* in the original scale and the person-mean centered *X* (*cX*) are included in the data analysis. Figure 1b shows the two-stage detrending *X* approach (Action b), in which detrended *X* (*dX*) is first obtained by using the residuals of regressing *X* on grand-mean centered time and then *Y* and *dX* are used in the second-stage analysis. Figure 1c shows the adding time as a covariate approach (Action c), in which *Y* is directly regressed on time and *cX*. Figure 1d shows the two-stage detrending both *X* and *Y* approach (Action d), in which detrended *X* (*dX*) and detrended *Y* (*dY*) are first obtained by using the residuals of regressing *X* or *Y* on time, respectively, and then *dY* and *dX* are used in the second-stage analysis. In contrast, Figure 1e

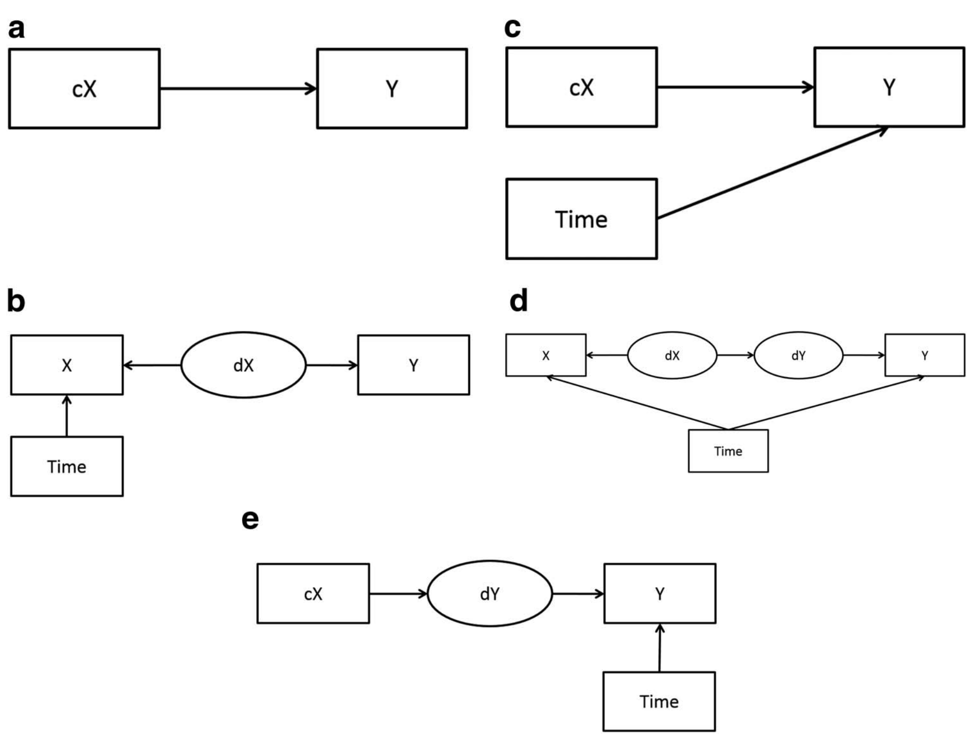

Figure 1. Illustrations of different detrending actions: (a) no detrending, or the person-mean centering approach ($cx_{it} = x_{it} - \bar{x}_{i.}$); (b) detrend *X* only by the two-stage approach in Curran and Bauer (2011); (c) add time as a covariate; (d) detrend both *X* and *Y*; and (e) detrend *Y* only. *cX* = the person-mean centered *X*; *dX* = the detrended *X*; *dY* = the detrended *Y*.

shows the two-stage detrending Y approach (Action e), in which detrended Y (dY) is first obtained by using the residuals of regressing Y on time, and then dY and cX are used in the second-stage analysis.

Appendix A gives the point estimates of within-person effects for individual i from different detrending actions depicted in Figure 1. From the estimates displayed in Equations A2–A4, we can see that Actions b–d—detrending X by the two-step approach, adding time as a covariate, and detrending both X and Y, respectively—would yield exactly the same point estimates for individual i. In other words, Actions b–d are functionally equivalent at the individual level. Appendix A also shows that results from no detrending or detrending only Y might produce different point estimates of individual within-person effects from detrending X only or detrending both X and Y. Therefore, adding time as a covariate, in this case, is not detrending Y but detrending X or detrending both instead.

More specifically, when $r^{(i)}_{x.time} = 0$ or there is no systematic linear trend in X for individual i, the point estimates are the same from all five actions included in Figure 1, regardless of whether $r^{(i)}_{y.time} = 0$. When $r^{(i)}_{y.time} = 0$ but $r^{(i)}_{x.time} \neq 0$ (there is no systematic linear trend in Y, but there is systematic change in X), the point estimates from no detrending and detrending Y only are the same but are smaller than those of detrending X only, adding time as a covariate, or detrending both X and Y, because $\frac{1}{1-(r^{(i)}_{x.time})^2} > 1$ with nonzero $r^{(i)}_{x.time}$.

When both $r^{(i)}_{y.time} \neq 0$ and $r^{(i)}_{x.time} \neq 0$, rank orders in the magnitudes of the point estimates of individual within-person effects from different detrending actions depend on the sizes of $r^{(i)}_{y.x}$, $r^{(i)}_{x.time}$, and $r^{(i)}_{y.time}$. When the product of $r^{(i)}_{x.time}$ and $r^{(i)}_{y.time}$ is less than 0 (one variable increases over time and the other decreases over time), the point estimates from detrending X only, adding time as a covariate, or detrending both X and Y are larger than those from detrending Y only. These, in turn, are larger than those from no detrending. When both $r^{(i)}_{x.time}$ and $r^{(i)}_{y.time}$ are positive and $r^{(i)}_{y.x}r^{(i)}_{x.time} > r^{(i)}_{y.time}$, or when both $r^{(i)}_{x.time}$ and $r^{(i)}_{y.time}$ are negative and $r^{(i)}_{y.x}r^{(i)}_{x.time} < r^{(i)}_{y.time}$, the point estimates from detrending X only, adding time as a covariate, or detrending both X and Y are larger than those from no detrending. These, in turn, are larger than those from detrending Y only. In contrast, when both $r^{(i)}_{x.time}$ and $r^{(i)}_{y.time}$ are positive and $r^{(i)}_{y.x}r^{(i)}_{x.time} < r^{(i)}_{y.time}$, or when both $r^{(i)}_{x.time}$ and $r^{(i)}_{y.time}$ are negative and $r^{(i)}_{y.x}r^{(i)}_{x.time} > r^{(i)}_{y.time}$, the point estimates from no detrending are larger than those from detrending X only, adding time as a covariate, or detrending both, which, in turn, are larger than those from detrending Y only.

In sum, depending on individual sample statistics including $r^{(i)}_{y.x}$, $r^{(i)}_{x.time}$, and $r^{(i)}_{y.time}$, detrending X only, adding time as a covariate, detrending both X and Y, or detrending Y only would yield the same individual within-person effect estimates or larger or smaller individual within-person effect estimates than no detrending (or person-mean centering approach) would. One implication is that the assertion of "applying the standard methods of person-mean centering to data in which the TVC varies as a function of time results in a within-person effect that drastically underestimates the known population value," from a hypothetical example in Curran and Bauer (2011, p. 607) may hold only for the considered scenario. Specifically, in this scenario there is no linear trend in Y, but there is linear change in X (i.e., $r^{(i)}_{y.time} = 0$ and $r^{(i)}_{x.time} \neq 0$).

These analytical results can be applied to compare estimates of within-person effects from different detrending actions for any single individual. With repeated data from multiple individuals, we can aggregate individual within-person effects via multilevel models. The point estimates of pooled (fixed-effects) within-person effects from different detrending actions depicted in Figure 1 can be calculated by averaging the individually obtained OLS within-person effect estimates shown in Appendix A with appropriate weights. The weights depend on the sampling variance of individual OLS estimates, the design matrix (time-varying covariates values), and the variance of within-person effects (Raudenbush & Bryk, 2002). It is difficult to show the exact weights in simple analytical forms with time-varying covariates. In this article, we apply OLS to the pooled long-format data with all individuals and time points to obtain a simple form for the point estimates of the fixed-effects within-person effects from some multilevel models under certain constraints, as shown in Appendix B. From Appendix B, we can see that the point estimates of the aggregated fixed-effects within-person effects could be different across all five detrending actions, although Actions b–d yield identical point estimates at the individual level.

Conceptually, detrending X only (Action b), adding time as a covariate (Action c), and detrending both X and Y (Action d) could be used to answer the research question of how X and Y are related net the time effect. Statistically, they may perform differently. More specifically, although Actions b–d can produce exactly the same point estimates of within-person effects for each individual separately under certain conditions, the standard errors could be different. As discussed earlier in this article and in the literature (e.g., Curran & Bauer, 2011), the two-stage detrending X only method does not consider uncertainty in the linear time effect for X and, thus, may produce incorrect standard errors. Similarly, the uncertainty is also ignored using the detrending both X and Y approach. Therefore, among Actions b, c, and d, we recommend the one-step multilevel approach with time as a Level-1 covariate (Action c) for better precision in the estimates when studying within-person relations between two variables net the time effect. Later in this article, to empirically evaluate our recommendation under a broader span of conditions, we compare the results on the average/expected estimates and the standard error estimates of the aggregated fixed-effects within-person effects using different detrending actions from a simulation study.

It is worth noting that both Curran et al. (2012) and Hoffman and Stawski (2009) described the adding time as a covariate approach as an approach for detrending outcome Y or controlling for the time effect in Y. From our analytical analysis, however, we found that this approach is actually equivalent to detrending X only or detrending both X and Y at the individual level. Basically, all three approaches (Actions b–d) can be used to control for the time effect in X or control for the time effect in the relation when studying the relation between X and Y. In other words, all three approaches yield a coefficient that estimates the relation between X and Y conditionally on time (assuming that the model correctly captures the functional form of the time trend). Another reason for recommending Action c is that it is simple and straightforward in terms of implementations, just one step instead of two.

Results from detrending Y only, however, conceptually do not answer the research question of how X and Y are related net the time effect. Instead, the results can only be interpreted as the relations between detrended Y and X or the relations between X and Y above and beyond systematic growth in Y over time, which may be less interesting to psychologists.

Furthermore, our analytical analysis showed that results from no detrending (Action a, person-mean centering) can be very different from those from approaches with detrending (Actions b–d). Again, we reiterate the importance of thinking about the need to detrend rather than automatically doing it.

In sum, to understand the relation between X and Y net the contribution of the linear effect of time, we recommend the one-step approach with time as a Level-1 covariate (Action c) in the multilevel modeling framework. In particular, as long as one is interested in the relation between two variables net the time effect, we recommend this approach regardless of whether only X, only Y, or both X and Y display trends over time.

Reliability of $\bar{x}_{i.}$

For many of the models discussed in the foregoing, $\bar{x}_{i.}$ is included as a predictor for the between-person effect. $\bar{x}_{i.}$ is not measurement-error free and, thus, may cause biased/attenuated estimates of between-person effects. Curran et al. (2012) highlighted that "the between-person effect is attenuated due to the omission of within-person variability around the person-specific mean" (p. 228). The magnitude of the attenuation is closely related to the reliability of $\bar{x}_{i.}$. When $\bar{x}_{i.}$ has relatively low reliability, the attenuation would be a large concern. Thus, in this subsection, we briefly review and discuss the reliability issue of $\bar{x}_{i.}$ with longitudinal data.

From the test theory perspective, the Spearman–Brown (prophecy) formula (Brown, 1910; Spearman, 1910) gives the reliability of r parallel items (equal factor loadings and unique variances; see also McDonald, 1999, p. 124). That is

$$\rho_r = \frac{r\rho_1}{(r-1)\rho_1 + 1}, \tag{17}$$

where r is the number of items, and ρ_1 is the reliability of any one item.

The Spearman–Brown formula shows that ρ_r is a monotone increasing function of r with a given ρ_1, and thus ρ_r is greater than ρ_1 when $r > 1$. Assuming that repeated measurements at different time points have equal reliabilities, we can apply the Spearman–Brown formula to study the reliability of $\bar{X}$ with information about the reliability of X_t. Therefore, the reliability of $\bar{X}$ is greater than the reliability of X_t at a given time point t when the total number of time points is two or larger. In well-designed and -conducted psychological studies, the reliability of X_t is often larger than .70, and thus the reliability of $\bar{X}$ is larger than .88 or .95 with three or eight time points, respectively, on the basis of Equation 17.

Taking within-person variability into account in addition to the measurement errors considered in the foregoing, Estabrook Grimm, and Bowles (2012) conducted a series of Monte Carlo simulations based on real results from previously published psychological longitudinal studies to study the reliabilities of intraindividual means. Their results showed "intraindividual means to be reliable in virtually all simulation conditions" (Estabrook et al., 2012, p. 568). Wang and Grimm (2012) analytically derived the reliability formula of $\bar{X}$. They found that the reliability of $\bar{X}$ depends on the between-person variance in true level scores, the expectation of true within-person variability, the measurement error variance, and the number of assessments. Wang and Grimm also included two real data examples, negative affect for the first example and perceptual–motor performance for the second example. In the examples, negative affect had its internal consistency reliability measured by Cronbach's coefficient alpha at .88, whereas perceptual–motor performance had its internal consistency reliability at .70. On the basis of the real data, the reliabilities of $\bar{X}$ were calculated using the formula that Wang and Grimm derived. The values were above .80, even with only three measurement occasions for both examples.

However, X_t could have lower reliabilities. For example, the reliability of X_t could be .45 or even lower in some psychological studies. In this case, the reliability of $\bar{X}$ can be .90 only when the number of time points is at least as large as 11, according to Equation 17. Further, Wang and Grimm (2012) also showed that the reliability of $\bar{X}$ can be relatively low—for example, when between-person variance is relatively small and/or within-person variance is relatively large. Therefore, we echo Curran et al. (2012)'s observation and suggest that researchers check the reliability of $\bar{X}$ using the Spearman–Brown formula or the formula in Wang and Grimm with their real longitudinal data before directly including $\bar{x}_{i.}$ into a model. When the reliability of $\bar{X}$ is low, the latent variable approach discussed in Curran et al. is recommended.

Real Data Analysis Illustration

In this subsection, we use a subset of data from an experimental longitudinal study to illustrate the effects of different detrending approaches on the disaggregation of between- and within-person effects when there are indeed trends in the variables of interest. The data come from 51 participants in a treatment group who were instructed to stop lying for 10 weeks. During each lab visit, the participants completed health measures and took a polygraph test assessing the number of white lies they had told in the past week. The substantive research question was whether dropping lies would link to better health. For the purpose of illustration, we use data on two variables: the number of white lies (X) and self-reported mental health complaints (Y) measured by the Brief Symptom Inventory (Derogatis, 1992), collected at Lab Visits 2–7. For the subset of data, there are no missing data, and all 51 participants were measured at the same measurement occasions. Sample standard deviations of variables of interest and their correlations are displayed in Table 4. Note that the data used for producing Table 4 are the long-format data, pooled from all individuals and time points.

Because the participants were instructed to stop lying and the substantive hypothesis was that dropping lies would link to better mental health, we would expect decreases in both the number of white lies and the amount of mental health complaints for the sample. Therefore, unconditional linear growth curve models were fitted to the data for each variable. As predicted, significant linear changes in both variables were detected, and the results are displayed in Table 5. From the growth curve modeling (GCM) results in Table 5 (Y and GCM-X columns), participants on average

Table 4
Descriptive Statistics of the Long-Format Data Pooled From All Individuals and Time Points

Variable	*SD*	*Y*	*dY*	*cX*	*dX*	*TIME*
Y: Mental health complaints	19.436					
dY: Detrended mental health complaints	17.980	.925				
cX: Person-mean centered lying frequency	2.041	.122	.031			
dX: Detrended lying frequency	1.416	.042	.045	.694		
TIME: Grand-mean centered time	1.711	−.210	.000	−.372	0	
$\overline{X}_{i.}$: Person mean in lying frequency	1.507	.132	.142	.000	.000	.000

Note. Values other than standard deviations are Pearson's correlation coefficients.

dropped .44 white lies per week, and their mental health complaints dropped by 2.39 points per week.

The number of white lies (X) was assessed by a polygraph test to maximize accuracy and reliability of the estimates. Assuming the polygraph test worked and, thus, the reliability of X at a given time point is as high as .80, then the reliability of $\overline{X}$ could be as high as .96 with six time points from Equation 17. Therefore, we may be able to directly use intraindividual means of lying frequency in the following models.

Different detrending actions depicted in Figure 1 were applied to analyze the data at the individual level. OLS estimates of individual within-person effects for five selected individuals are listed in Table 6. Consistent with the results in Appendix A, detrending X using the two-step approach (Action b), adding time as a covariate (Action c), and detrending both X and Y (Action d) yielded exactly the same point estimates for any single individual. In contrast, the estimates from the person-mean centering approach without detrending and the detrending Y only approach were different from each other and also different from the other three identical values for a given individual. Interestingly, the rank orders of the estimates from different methods were different across the five individuals because of individually different descriptive statistics, including $r^{(i)}_{y.x}$, $r^{(i)}_{x.time}$, and $r^{(i)}_{y.time}$.

A series of multilevel models were then fitted to the longitudinal data to study both between- and within-person effects of lying on health. First, the model in Equation 7, Model M3, the person-mean centering approach model in the full form without detrending either variable, was fitted. Convergence criteria, however, were not met, even with the "nobound" option in SAS PROC MIXED.[3] The reduced form with the assumption of homogeneity in the within-person effects (Model M3r in Table 5; Equation 10) was then applied to analyze the data. Results from the person-mean centering approach (Model M3r) showed that the between-person effect was not significantly different from 0 ($\hat{\gamma}_{01} = 1.22$, $SE = 1.10$). In other words, it is plausible that individuals who tell more white lies do not differ in health complaints from individuals who tell fewer white lies. However, the fixed-effects within-person effect was significantly positive ($\hat{\gamma}_{10} = r_{y.cx}\frac{s_y}{s_{cx}} = .122 \times \frac{19.436}{2.041} = 1.16$, $SE = .30$), such that dropping lies linked to better health over time intraindividually. In other words, in a week when a person reported more white lies, the same individual reported more mental health complaints than when he or she reported fewer white lies in another week.

For comparison, we fitted the model (Model M4r in Table 5; Equation 12) by the two-step detrending approach to the data. Because the study design is a balanced study and there are no missing data in this data set, and thus $\bar{time}_{i.} = \bar{time}_{..}$, as predicted, estimates of both the intercept (γ_{00}) and the between-person effect (γ_{01}) were identical to the estimates from the person-mean centering approach. Also, consistent with the theoretical analysis, the estimate of the fixed-effects within-person effect from Model M4r became smaller ($\hat{\gamma}_{10} = r_{y.dx}\frac{s_y}{s_{dx}} = .042 \times \frac{19.436}{1.416} = .57$) than the estimate from Model M3r because of $\frac{r_{y.dx}}{s_{dx}} < \frac{r_{y.cx}}{s_{cx}}$. In addition, the fixed-effects within-person effect was no longer statistically significant from Model M4r.

After that, we fitted two versions of the model (M5r and M5r-a) in Equations 14 or 15 to the data, where M5r is the model in Equation 14 without any constraints and M5r-a is the model in Equation 14 or Equation 15 with a 0 constraint on the covariance between u_{0i} and u_{1i}. Results from models M5r and M5r-a were consistent with our expectations such that the estimates of the intercept and the between-person effect from M5r-a were identical to the corresponding estimates from the person-mean centering approach. In addition, the estimate of the within-person effect was more efficient from M5r and M5r-a (the standard error of the within-person effect estimate was .30) than from M4r (the standard error was .44). This was because of the measurement-error free feature in rates of change in Models M5r and M5r-a.

Detrending of both variables was also applied by fitting the model (Model M6r in Table 5 [see Equation B4]) in multiple steps to the data. The point estimates and standard error estimates of the intercept and the between-person effect from M6r were again identical to the corresponding estimates from the person-mean centering approach. The within-person effect estimate, $\hat{\gamma}_{10} = r_{dy.dx}\frac{s_{dy}}{s_{dx}} = .045 \times \frac{17.980}{1.416} = .57$, was also smaller than the estimate from person-mean centering approach and close to the estimates from the detrending X only and adding time as a covariate approaches (i.e., Models M4r, M5r, and M5r-a).

Finally, we explored applying the two-step detrending approach to the outcome variable instead of the predictor variable. We fitted a regression to each individual's data by regressing the outcome variable on the grand-mean centered time variable. Then the linear trend was removed from the response data by subtracting only the fitted linear term from the response data while keeping the fitted intercept term for estimating the between-person effect. Then we applied the person-mean centering approach to the predictor variable and fit the model to the detrended outcome data. The model

[3] The "nobound" option in SAS PROC MIXED requests the removal of boundary constraints on variance components. That is, it allows variance estimates to be negative.

Table 5
Parameter Estimates and Standard Error Estimates From a Series of Multilevel Models

Parameter	GCM-*Y*: Growth in *Y*	GCM-*X*: Growth in *X*	M3r: P-center (Action a)	M4r: Detrend *X* (Action b)	M5r: Time as covariate (Action c)	M5r-a: Time as covariate (Action c)	M6r: Detrend both (Action d)	M7r: Detrend *Y* (Action e)
				Fixed-effects parameters				
Initial level	29.56 (3.06)	2.62 (0.48)	—		—			—
Linear rate	−2.39 (0.50)	−0.44 (0.10)	—		−2.15 (0.52)	−2.15(0.52)		—
γ_{00}	—	—	21.76 (2.83)	21.76 (2.83)	21.46 (2.75)	21.76 (2.83)	21.76 (2.83)	21.76 (2.83)
γ_{01}	—	—	1.22 (1.10)	1.22 (1.10)	1.42 (1.01)	1.22 (1.10)	1.22 (1.10)	1.22 (1.10)
γ_{10}	—	—	1.16 (0.30)	0.57 (0.44)	0.52 (0.30)	0.53 (0.30)	0.57 (0.30)	0.27 (0.21)
				Random-effects variances				
var(initial level)	441.05 (94.67)	9.94 (2.29)	—	—	—			—
var(linear rate)	8.93 (2.59)	0.37 (0.11)	—		8.91 (2.58)	8.91 (2.58)		—
var(u_{0i})	—	—	249.98 (53.31)	249.98 (53.31)	257.72 (53.36)	257.54 (53.29)	259.84 (53.29)	259.78 (53.29)
var(e_{it})	70.01 (6.93)	3.00 (0.30)	114.43 (10.13)	120.39 (1.66)	69.05 (6.84)	69.04 (6.84)	55.23 (4.89)	55.64 (4.93)
				Overall model fit				
−2*LL*	2,380.3	1,359.7	2,453.8	2,466.8	2,375.7	2,385.2	2,268.0	2,269.9

Note. Action labels correspond to the labels in Figure 1. Dashes in cells indicate not applicable; GCM = growth curve model; *Y* = Brief Symptom Inventory mental health complaints; *X* = number of white lies; P-center = person-mean centering approach; M = model; γ_{00} = intercept; γ_{01} = between-person effect; γ_{10} = within-person effect; *var* = variance; *var*(u_{0i}) = variance of the intercept term; *var*(e_{it}) = variance of the Level-1 residual variable; −2*LL* = −2 log likelihood value.

is denoted as Model M7r (see Equation B6) in Table 5. The estimate of the between-person effect from Model M7r was equivalent to that from model M3r, whereas the estimate of the within-person effect, $\hat{\gamma}_{10} = r_{dy.cx}\frac{s_{dy}}{s_{cx}} = .031 \times \frac{17.980}{2.041} = .27$, was shrunk toward 0 for our data.

From Table 5, we can see that the two-step approach with detrending *X* only (Model M4r), the multilevel models with adding time as a covariate (Models M5r and M5r-a), and the model with detrending both *X* and *Y* yielded slightly different estimates of fixed-effects within-person effects, as expected.[4] Not surprisingly from our analytical analysis, the estimates of within-person effects from those models were very different from the estimates from models without detrending (Model M3r) and with detrending *Y* only (Model M7r). For this example, because the systematic changes in both *X* and *Y* came from the experimental manipulation and there is no need to remove the experimental manipulation from the time-varying data, it is not appropriate to control for the time effect for studying the within-person relation between the two variables. Actually, detrending one or both variables would violate the natural feature of the experiment. Therefore, to answer the substantive research question of whether dropping lies would link to better health, we recommend detrending neither variable and using Model M3r results.

Simulation

In the second empirical data example, we did not include the within-person effect as a random effect in the models because of convergence problems. The analytical derivation on detrending in Appendix B was also based on the assumption that there was no between-person variation in the within-person effects. To further compare the results from different centering and detrending methods under a wider range of scenarios, we conducted the following simulations. In both simulations, the parameter values used in the population models were adopted and modified from the real data analysis results in Table 5. For example, the parameter values used for generating the time-varying predictor were adopted and modified from the GCM-X model results in Table 5. The number of replications for each condition is 500.

Simulation Study on Comparing Results From Different Centering Approaches

The goal of this simulation study was to compare the results from different centering approaches (i.e., no centering, person-mean centering of *X*, and grand-mean centering of *X*) when there is between-person variation in the within-person effects.[5] To fulfill the goal, we simulated data based on Model M3 in Equation 7 (the full model with person-mean centering), in which we assume that neither *X* nor *Y* have trends over time. x_{it} was generated from a no change model with the average intercept at 2.5, the Level-1 residual variance at 3, and the variance of intercept at 10. The manipulated factors in the simulation were (a) sample size (N = 50, 100, or 200); (b) number of time points (T = 5 or 10); (c) size of the fixed-effects between-person effect (γ_{01} = 0 or 0.5); (d) size of the fixed-effects within-person effect (γ_{10} = 0 or 0.5); (e) size of the Level-1 residual variance (σ_e^2 = 1 or 64); and (f) Level-2 covariance structure, $\Sigma[u_{0i}, u_{1i}] = \begin{pmatrix} \sigma_{00} & \sigma_{01} \\ \sigma_{10} & \sigma_{11} \end{pmatrix}$. For Σ, we considered a baseline scenario with $\Sigma = \begin{pmatrix} 225 & -15 \\ -15 & 4 \end{pmatrix}$, under which the standard deviation of the within-person effects is 2. We varied σ_{10} (σ_{10} = 15, 0, or −15) such that the correlation between u_{0i} and

[4] Even though all the participants have the same number of time points and there are no missing data in the data set, the values of the time-varying covariate are different across individuals, which leads to different sets of weights for aggregating the individual within-person effect estimates from different methods.

[5] When there is no between-person variation in the within-person effects, the models with person-mean or grand-mean centering are equivalent.

Table 6
Ordinary Least Squares Estimates of Individual Within-Person Effects for Five Selected Individuals

ID	P-centering (Action a)	Detrend X, two step (Action b)	Add time as covariate (Action c)	Detrend both X and Y (Action d)	Detrend Y (Action e)
10	1.250	0.750	0.750	0.750	0.686
11	1.587	−0.981	−0.981	−0.981	−0.257
14	1.098	−0.573	−0.573	−0.573	−0.425
110	0.781	0.432	0.432	0.432	0.497
116	1.297	0.720	0.720	0.720	0.666

Note. Action labels correspond to the labels in Figure 1. ID = participant number; P-centering = person-mean centering approach.

u_{1i} was .5, 0, or −.5. Thus, we had three levels for Σ in total. In total, there were $3 \times 2 \times 2 \times 2 \times 2 \times 3 = 144$ conditions.

For each generated data set, the three centering approaches were implemented to analyze the data. Given the large number of results, we focus on four parameters. They are the fixed-effect between- and within-person effects γ_{01} and γ_{10}, the Level-1 residual variance σ_e^2, and the variance of within-person effects σ_{11}. The results from conditions with $\sigma_e^2 = 64$ and the baseline Σ are presented in Tables 7 and 8. With nonzero between-person variation in within-person effects, we have the following observations from the tables: (a) When $\gamma_{10} \neq \gamma_{01}$, the estimated composite effect captures neither the between-person effect nor the within-person effect, even when $\gamma_{10} = 0$ or $\gamma_{01} = 0$. (b) As expected, the average fixed-effects within-person and between-person estimates ($\hat{\gamma}_{10}$ and $\hat{\gamma}_{01}$) from both person-mean centering and grand-mean centering, respectively, were close to the true values as a result of the same mean structure. (c) The average standard error estimates of $\hat{\gamma}_{10}$ from person-mean centering were slightly larger than those from grand-mean centering, whereas those of $\hat{\gamma}_{01}$ from person-mean centering were slightly smaller than the counterpart; however, the standard error estimates of $\hat{\gamma}_{10}$ from grand-mean centering were slightly underestimated by comparing the average standard error estimates to the standard deviations of the estimates (see the standard error estimates comparison table at http://www3.nd.edu/~lwang4/withinbetween/; the highest relative

Table 7
Estimates of the Fixed-Effects Within-Person Effects ($\hat{\gamma}_{10}$) and Between-Person Effects ($\hat{\gamma}_{01}$) and Their Standard Errors Across Conditions From Different Centering Approaches

				Composite effect		Within-person effect estimate: $\hat{\gamma}_{10}$				Between-person effect estimate: $\hat{\gamma}_{01}$			
				No centering		Grand-mean centering		Person-mean centering		Grand-mean centering		Person-mean centering	
γ_{10}	γ_{01}	Nper	Ntime	Est.	*SE*	Est.	*SE*	Est.	*SE*	Est.	*SE*	Est.	*SE*
0.00	0.00	50	5	−0.01	0.39	−0.02	0.42	−0.02	0.45	0.03	0.71	0.03	0.64
0.00	0.00	50	10	−0.01	0.33	0.00	0.34	0.01	0.36	−0.03	0.71	−0.02	0.63
0.00	0.00	100	5	0.00	0.27	0.00	0.30	0.01	0.32	0.00	0.50	0.01	0.45
0.00	0.00	100	10	0.01	0.24	0.01	0.24	0.00	0.26	0.01	0.05	0.01	0.44
0.00	0.00	200	5	0.00	0.19	0.00	0.21	−0.01	0.23	0.00	0.36	−0.01	0.32
0.00	0.00	200	10	0.01	0.17	0.01	0.17	0.01	0.18	−0.01	0.36	−0.01	0.31
0.50	0.00	50	5	0.41	0.39	0.50	0.42	0.49	0.45	0.02	0.71	0.00	0.65
0.50	0.00	50	10	0.47	0.33	0.53	0.34	0.50	0.36	−0.03	0.71	0.00	0.63
0.50	0.00	100	5	0.40	0.27	0.50	0.30	0.50	0.32	0.01	0.50	0.01	0.45
0.50	0.00	100	10	0.45	0.24	0.51	0.24	0.51	0.26	−0.04	0.51	0.00	0.44
0.50	0.00	200	5	0.40	0.19	0.51	0.21	0.50	0.23	0.00	0.36	0.00	0.32
0.50	0.00	200	10	0.45	0.17	0.50	0.17	0.51	0.18	−0.01	0.36	0.00	0.31
0.00	0.50	50	5	0.10	0.39	−0.01	0.42	0.01	0.45	0.54	0.70	0.51	0.64
0.00	0.50	50	10	0.06	0.34	0.01	0.34	−0.01	0.36	0.50	0.71	0.50	0.63
0.00	0.50	100	5	0.09	0.27	−0.03	0.30	−0.02	0.32	0.52	0.50	0.50	0.45
0.00	0.50	100	10	0.05	0.24	0.00	0.24	0.01	0.26	0.44	0.50	0.47	0.44
0.00	0.50	200	5	0.11	0.19	0.00	0.21	0.00	0.23	0.50	0.36	0.50	0.32
0.00	0.50	200	10	0.05	0.17	0.00	0.17	0.00	0.18	0.51	0.35	0.50	0.31
0.50	0.50	50	5	0.50	0.39	0.50	0.42	0.50	0.45	0.49	0.72	0.48	0.64
0.50	0.50	50	10	0.50	0.33	0.50	0.34	0.50	0.36	0.56	0.71	0.52	0.63
0.50	0.50	100	5	0.51	0.27	0.52	0.30	0.51	0.32	0.53	0.50	0.51	0.45
0.50	0.50	100	10	0.48	0.24	0.48	0.24	0.49	0.26	0.52	0.51	0.51	0.44
0.50	0.50	200	5	0.50	0.19	0.51	0.21	0.51	0.23	0.47	0.36	0.48	0.32
0.50	0.50	200	10	0.50	0.17	0.50	0.17	0.50	0.18	0.51	0.36	0.50	0.31

Note. Nper = number of individuals; Ntime = number of time points; Est. = average estimate; *SE* = average standard error estimate.

Table 8
Estimates of Variance Components and Their Standard Error Estimates Across Conditions From Different Centering Approaches

True value	Nper	Ntime	No centering Est.	No centering *SE*	Grand-mean centering Est.	Grand-mean centering *SE*	Person-mean centering Est.	Person-mean centering *SE*
Level-1 residual variance (σ_e^2)								
64	50	5	66.10	7.36	66.04	7.35	64.02	7.21
64	50	10	64.82	4.59	64.79	4.59	64.01	4.51
64	100	5	66.10	5.20	66.07	5.20	63.97	5.09
64	100	10	64.81	3.24	64.79	3.24	64.03	3.19
64	200	5	66.15	3.68	66.12	3.68	63.98	3.60
64	200	10	64.89	2.30	64.88	2.30	64.05	2.26
Variance of within-person effects (σ_{11})								
4.00	50	5	2.79	1.54	2.81	1.55	4.03	2.02
4.00	50	10	3.27	1.14	3.28	1.15	4.00	1.34
4.00	100	5	2.68	1.06	2.69	1.06	4.03	1.42
4.00	100	10	3.27	0.80	3.27	0.80	4.01	0.94
4.00	200	5	2.65	0.74	2.65	0.74	4.02	0.99
4.00	200	10	3.24	0.56	3.24	0.56	3.99	0.66

Note. Nper = number of individuals; Ntime = number of time points; Est. = average estimate; *SE* = average standard error estimate.

bias was −14%). (d) The average variance estimates of σ_e^2 and σ_{11} from no centering and grand-mean centering were very close to each other. And (e) the variance estimates from person-mean centering were very close to the true values, whereas those from no centering and grand-mean centering were inaccurate. For Point e, more specifically, σ_{11}s were underestimated from grand-mean centering, which is consistent with what Raudenbush and Bryk (2002) showed via the heuristic illustration. In addition, an ad hoc simulation analysis with $T = 20$, 50, or 150 showed that with more time points, the estimates of σ_{11} from grand-mean centering were closer to the true value, which is consistent with the shrinkage problem of empirical Bayes estimates (e.g., Carlin & Louis, 2000; Raudenbush & Bryk, 2002).

For the conditions with $\sigma_e^2 = 1$ or $\sigma_{10} = 0$ or 15, the observations just described remain, except the average standard error estimates of $\hat{\gamma}_{10}$ from both grand-mean centering and person-mean centering were closer to each other under the conditions with $\sigma_e^2 = 1$.

Therefore, consistent with the recommendation we made earlier, our simulation results supported the person-mean centering approach for disaggregating between- and within-person effects. In particular, person-mean centering yielded more precise fixed-effects estimates and more accurate variance component estimates under the conditions of nonzero variances in the within-person effects.

Simulation Study on Comparing Results From Different Detrending Approaches

The goal of this simulation study was to compare the results from different detrending approaches (i.e., Actions a–e in Figure 1), especially when there was between-person variation in the within-person effects. In this simulation study, we simulated data on the basis of a full-model generalization of M5r in Equation 14, where the full model relaxes the restriction of no individual differences in within-person effects. Thus, the full model, M5, becomes

$$\begin{aligned} y_{it} &= \gamma_{0i} + \gamma_{1i}(x_{it} - \bar{x}_{i.}) + \gamma_{2i}(time_{it} - \bar{time}_{..}) + e_{it} \\ \gamma_{0i} &= \gamma_{00} + \gamma_{01}\bar{x}_{i.} + u_{0i} \\ \gamma_{1i} &= \gamma_{10} + u_{1i} \\ \gamma_{2i} &= \gamma_{20} + u_{2i} \end{aligned} \quad , \qquad (18)$$

where we could allow individual differences in within-person effects, γ_{1i}, as shown by the inclusion of u_{1i}. x_{it} was generated from a linear growth curve model with the average intercept at 2.5, the average slope at −.5, the Level-1 residual variance at 3, and the Level-2 covariance matrix $\begin{pmatrix} 10 & -1.6 \\ -1.6 & 0.36 \end{pmatrix}$. The manipulated factors in this simulation were the same as in the simulation for centering, except we manipulated the size of fixed-effects time effect ($r_{20} = 0$ or -2) and manipulated the Level-2 covariance structure, $\sum[u_{0i}, u_{1i}, u_{2i}] = \begin{pmatrix} \sigma_{00} & \sigma_{01} & \sigma_{02} \\ \sigma_{10} & \sigma_{11} & \sigma_{12} \\ \sigma_{20} & \sigma_{21} & \sigma_{22} \end{pmatrix}$, differently. For Σ, we considered a baseline scenario with $\sum = \begin{pmatrix} 225 & -15 & -10 \\ -15 & 4 & 2 \\ -10 & 2 & 9 \end{pmatrix}$, under which the standard deviation of the within-person effects was 2. We also considered 11 other scenarios with variants of this baseline matrix by varying the signs of σ_{10}, σ_{20}, and σ_{21} to be positive, negative, or 0 and σ_{11} and σ_{22} to be 0 or nonzero. (The other considered matrices are listed on our paper Web page: http://www3.nd.edu/~lwang4/withinbetween/.) In total, there were $3 \times 2 \times 2 \times 2 \times 2 \times 2 \times 12$ conditions.

For each generated data set, the five actions in Figure 1 were implemented to analyze the data. Again, we focus on comparing the results of four parameters in Model M5 across different approaches: γ_{10}, γ_{01}, σ_e^2, and σ_{11}. We also mainly present the results (Tables 9 and 10) from conditions with $\sigma_e^2 = 64$ and the baseline Σ, because between-person variation in the within-person effects is allowed in this case.

Because the model for generating the data is the time as covariate model, the results from the time as covariate model in Tables

9 and 10 showed that the fixed effects and the variance components can be recovered well across all the conditions, as expected. In addition, the average standard error estimates were close to the standard deviation of the parameter estimates with the highest relative bias at 4% (see the standard error estimates comparison table on our paper Web page: http://www3.nd.edu/~lwang4/withinbetween/). Thus, the standard error estimates were also accurate. Therefore, we use the results from the time as covariate model as baseline and compare the results from other approaches with those from the time as covariate model.

Detrending versus no detrending. When the average time effect is 0, $\gamma_{20} = 0$, the fixed-effects average within-person or between-person estimates from the no detrending approach were close to those from the time as covariate model. When $\gamma_{20} = -2$, the fixed-effects within-person effect estimates were larger than those from the time as covariate model. For variance components estimates, regardless of the value for γ_{20}, the estimates of both the Level-1 residual variance and the variance of the within-person effects from the no detrending approach were larger than those from the time as covariate model. Therefore, our simulation results showed that when there is variation in the within-person effects ($\sigma_{11} = 4$), the estimates from detrending versus no detrending approaches could be different, especially with nonzero γ_{20}. For our analytical and real data analyses conducted under zero variance of the within-person effects, different results were also found from detrending and no detrending approaches. Therefore, before conducting the data analysis, a researcher needs to carefully consider whether it is necessary to control for the time effect when looking into the relationship between two time-varying variables. If there is no motivation to control for the time effect in the relation between the two variables, do not detrend.

How to detrend. When it is necessary to control for the time effect, our simulation results from various detrending approaches showed that (a) the average estimates of the fixed-effects between-person effects from the detrend *X* only, time as covariate, and detrend both *X* and *Y* approaches were close to each other; (b) the average estimates of the fixed-effects within-person effects from the detrend *X* only, time as covariate, and detrend both *X* and *Y* approaches were also close to each other, whereas those from the detrend *Y* only approach were substantially different (smaller) when the fixed-effects within-person effect was nonzero (i.e., .5); (c) the average standard errors of the fixed-effects within-person estimates from the detrend *X* only and the detrend both *X* and *Y* approaches were larger than those from the time as covariate approach, especially when both the sample size and the number of time points were small; and (d) both the average Level-1 residual variance estimates and the average variance estimates of the within-person effects from the detrend *X* only, detrend both *X* and

Table 9
Estimates of the Fixed-Effects Within-Person Effects ($\hat{\gamma}_{10}$) and Between-Person Effects ($\hat{\gamma}_{01}$) and Their Standard Errors Across Conditions From Different Detrending Approaches

			No detrending											
			$r_{20} = 0$		$r_{20} = -2$		Detrend *X*		Time as covariate		Detrend both		Detrend *Y*	
True effect	Nper	Ntime	Est.	*SE*	Est.	*SE*	Est.	*SE*	Est.	*SE*	Est.	*SE*	Est.	*SE*
Fixed-effects within-person effect (γ_{10})														
0.00	50	5	0.00	0.42	0.54	0.44	0.00	0.49	0.00	0.44	0.00	0.50	0.00	0.26
0.00	50	10	−0.01	0.38	0.95	0.42	−0.01	0.38	−0.01	0.36	−0.01	0.37	−0.01	0.14
0.00	100	5	−0.01	0.30	0.53	0.31	0.00	0.35	−0.01	0.31	0.00	0.36	0.00	0.18
0.00	100	10	−0.01	0.27	0.96	0.30	0.01	0.27	0.01	0.26	0.01	0.26	0.00	0.09
0.00	200	5	0.00	0.21	0.53	0.22	0.00	0.25	0.00	0.22	0.00	0.25	0.00	0.13
0.00	200	10	0.00	0.19	0.98	0.21	0.00	0.19	0.00	0.18	0.00	0.19	0.00	0.06
0.50	50	5	0.51	0.42	1.01	0.44	0.49	0.50	0.49	0.44	0.49	0.50	0.26	0.27
0.50	50	10	0.47	0.38	1.45	0.42	0.50	0.38	0.49	0.36	0.50	0.37	0.18	0.14
0.50	100	5	0.49	0.30	1.02	0.31	0.50	0.35	0.50	0.31	0.51	0.35	0.25	0.18
0.50	100	10	0.52	0.27	1.47	0.30	0.49	0.27	0.49	0.26	0.49	0.26	0.16	0.09
0.50	200	5	0.51	0.21	1.03	0.22	0.49	0.25	0.49	0.22	0.50	0.25	0.24	0.13
0.50	200	10	0.50	0.19	1.47	0.21	0.50	0.19	0.50	0.18	0.51	0.19	0.16	0.06
Fixed-effects between-person effect (γ_{01})														
0.00	50	5	0.04	0.75	0.05	0.76	0.01	0.75	0.02	0.73	0.01	0.75	0.01	0.75
0.00	50	10	0.00	0.72	−0.15	0.72	0.01	0.66	0.00	0.66	0.01	0.65	0.00	0.69
0.00	100	5	−0.03	0.53	0.01	0.53	0.01	0.53	0.01	0.52	0.01	0.53	0.01	0.53
0.00	100	10	−0.02	0.51	−0.13	0.51	−0.01	0.46	0.00	0.47	0.01	0.46	0.08	0.49
0.00	200	5	0.00	0.37	0.02	0.38	0.00	0.37	0.00	0.37	0.00	0.37	0.00	0.37
0.00	200	10	−0.02	0.36	−0.12	0.36	0.02	0.33	0.02	0.33	0.03	0.32	0.06	0.35
0.50	50	5	0.51	0.75	0.54	0.76	0.48	0.76	0.50	0.73	0.49	0.76	0.48	0.76
0.50	50	10	0.48	0.72	0.37	0.72	0.47	0.65	0.50	0.66	0.49	0.65	0.48	0.69
0.50	100	5	0.48	0.53	0.57	0.53	0.49	0.53	0.50	0.52	0.50	0.53	0.49	0.53
0.50	100	10	0.54	0.51	0.36	0.51	0.49	0.46	0.51	0.47	0.50	0.46	0.55	0.49
0.50	200	5	0.50	0.38	0.54	0.38	0.50	0.37	0.50	0.37	0.50	0.37	0.50	0.37
0.50	200	10	0.49	0.36	0.38	0.36	0.49	0.33	0.52	0.33	0.50	0.32	0.55	0.34

Note. Detrend both = detrend both *X* and *Y*; Nper = number of individuals; Ntime = number of time points; Est. = average estimate; *SE* = average standard error estimate.

Table 10
Estimates of Variance Components and Their Standard Error Estimates Across Conditions From Different Detrending Approaches

			No detrending											
			$r_{20} = 0$		$r_{20} = -2$		Detrend *X*		Time as covariate		Detrend both		Detrend *Y*	
True value	Nper	Ntime	Est.	*SE*	Est.	*SE*	Est.	*SE*	Est.	*SE*	Est.	*SE*	Est.	*SE*
Level-1 residual variance (σ_e^2)														
64	50	5	76.94	8.69	84.36	9.53	94.94	10.30	63.81	8.22	43.57	4.98	55.31	6.17
64	50	10	96.46	6.82	116.60	8.25	156.70	10.87	64.00	4.78	56.04	3.95	67.78	4.79
64	100	5	77.04	6.14	84.76	6.76	94.94	7.25	63.89	5.80	43.60	3.52	55.45	4.36
64	100	10	96.07	4.80	117.05	5.86	157.38	7.70	63.96	3.38	56.03	2.79	68.03	3.39
64	200	5	76.72	4.32	84.60	4.77	95.07	5.12	64.01	4.10	43.71	2.49	55.67	3.09
64	200	10	96.30	3.40	116.97	4.14	157.40	5.44	63.99	2.39	56.03	1.98	68.22	2.38
Within-person effect variance														
4.00	50	5	3.79	1.79	4.01	1.94	1.78	2.22	4.09	2.01	6.30	2.61	0.54	0.75
4.00	50	10	5.51	1.53	6.56	1.84	0.92	1.35	4.03	1.33	4.34	1.41	0.18	0.26
4.00	100	5	3.69	1.24	3.83	1.33	1.72	1.50	4.00	1.38	6.23	1.82	0.41	0.50
4.00	100	10	5.52	1.07	6.60	1.30	0.89	0.93	4.00	0.93	4.33	0.99	0.07	0.15
4.00	200	5	3.70	0.87	3.84	0.93	1.73	1.05	3.99	0.97	6.18	1.27	0.34	0.34
4.00	200	10	5.49	0.75	6.61	0.92	0.93	0.66	4.00	0.65	4.34	0.70	0.02	0.09

Note. Detrend both = detrend both *X* and *Y*; Nper = number of individuals; Ntime = number of time points; Est. = average estimate; *SE* = average standard error estimate.

Y, and detrend *Y* only approaches were very different from those from the time as covariate approach. Note that the results here were obtained when the variance of the within-person effects was 4. The overall pattern is consistent with that from our analytical analysis and real data analysis with the variance of the within-person effects fixed to be 0.

For simulation conditions with $\sigma_e^2 = 1$, the results were similar to those from the conditions with $\sigma_e^2 = 64$, except the average standard errors of the fixed-effects between- and within-person effect estimates from the detrend both approach with $\sigma_e^2 = 1$ were closer to those from the time as covariate approach than were those with $\sigma_e^2 = 64$. For simulation conditions with other Level-2 covariance matrices, the findings were similar to those found from the conditions with the baseline Σ. Thus, we do not list them here for the sake of saving space.

On the basis of the simulation results from various detrending approaches, among the detrending approaches, we recommend the time as covariate approach for controlling for the time effect when looking into the relation between two time-varying variables, because it will give more accurate and efficient fixed-effects within- and between-person effects and variance component estimates.

The analyses in the simulation were conducted via SAS PROC MIXED with the nobound option. For illustrating the procedures and results of different centering and detrending methods, we provide example SAS code and an example data set on our paper Web page: http://www3.nd.edu/~lwang4/withinbetween/.

Discussion and Recommendation

This article extends previous discussions of disaggregating between-person and within-person effects in the multilevel modeling framework (e.g., Curran & Bauer, 2011; Curran et al., 2012; Hoffman & Stawski, 2009) with a focus on two issues: centering and detrending.

For centering, there are three approaches researchers currently use: no centering, grand-mean centering, and person-mean centering. Our analytical, empirical, and simulation results showed that centering the time-varying predictor is needed for appropriately disaggregating between- and within-person effects when there are no trends in the two time-varying variables. Although the grand-mean and person-mean centering approaches are equivalent when the within-person effects are included in the model as fixed effects only, the person-mean centering approach is still preferred. The reason is that it is easier to conduct statistical inference on both between- and within-person effects with the person-mean centering approach, because the between-person effect using the grand-mean centering approach is a combination of two parameters, whereas both between- and within-person effects are single parameters from the person-mean centering approach. When the within-person effects are included as random effects, our simulation results supported the person-mean centering approach for its accuracy and precision in both fixed-effects within-person effect and variance component estimates.

For detrending, our conceptual illustration, real data analysis examples, and simulation study all showed that different detrending approaches could lead to very different conclusions on between- and within-person effects, especially within-person effects. Further, detrending might not be necessary for studying between- and within-person effects even when time-varying variables *X* and/or *Y* have trends over time. Before adopting a detrending approach, researchers need to decide whether controlling for the time effect when investigating the relations between two variables is necessary. For instance, in experimental longitudinal studies, participants in an intervention group receive the intervention, and thus trends in either *X* or *Y* or both over time for those participants may be expected. The experimental intervention may also be deliberately designed to cause the within-person effects of *X* on *Y* for those participants. Further, compared with the participants in the control group, those in the intervention group may have stronger between- and within-person effects as a result of the intervention. In this case, if we control for the time effect or detrend either X or Y or both, we may distort the natural feature of the experimental study design and, thus, may prevent the discovery of the between- and within- person effects of interest. In contrast, in short-term observational longitudinal studies, no manipulation or treatment is carried out on any participants in the sample; therefore, it may be reasonable to expect no trends in either X or Y before collecting data. In real data, however, researchers may observe trends in one of the variables or both for some participants. These trends may be attributable to some natural or everyday life events that are not related to the study design. In such cases, the researchers may want to control for the time effect by removing trends unrelated to either the study or the substantive research questions. Therefore, to detrend or not to detrend is a question researchers need to consider. The decision depends on whether controlling for the time effect (or detrending) is theoretically sound or helps answer the substantive research questions about between- and within-person effects. More specifically, when the researchers are interested in the relations between two time-varying variables *net the time effect*, detrending is needed.

When detrending is needed, various detrending approaches have been discussed in the current literature. For example, Curran and Bauer (2011) recommended detrending X by a two-step approach. Curran et al. (2012) and Hoffman and Stawski (2009) suggested controlling for the time effect in Y by including time as a Level-1 covariate via multilevel modeling. Our analytical and empirical results showed that at the individual level, detrending X by the two-step approach, adding time as a covariate, and detrending both X and Y would produce identical estimates of within-person effects under the conditions of linear time effects in X and/or Y. Therefore, the adding time as a Level-1 covariate approach is for controlling the time effect in X or controlling for the time effect in the relations between X and Y instead of for controlling the time effect in Y. No detrending and detrending Y only would produce different individual estimates, and the rank orders of the estimated within-person effects for a given individual from different detrending methods depend on the correlations of X with Y, of X with time, and of Y with time for the individual. At the population level, different detrending approaches may produce all different estimates of fixed-effects within-person effects aggregated across individuals. Our empirical data analysis and our simulation results showed that when the data are balanced, detrending X by the two-step approach, adding time as a covariate, and detrending both X and Y produced slightly different but quite close estimates of fixed-effects within-person effects (aggregated across individuals), as compared with estimates from the no detrending or detrending Y only approaches. Essentially, these three detrending actions can be used to study the relations between X and Y net the contribution of the effect of time. Among them, we recommend the one-step multilevel approach of including time as a Level-1 covariate for its easy implementation and better accuracy and efficiency. When the detrending Y only approach is implemented, we should explain the resultant within-person relations as the relations between detrended Y and X, not the relations between Y and X net the time effect from our analysis. In other words, detrending Y only does not answer the research question about within-person relations after controlling for the time effect. Further, looking into the relations between detrended Y and X may be less interesting to psychologists than investigating the relations between Y and X net the time effect. Therefore, the detrending Y only approach may rarely be appropriate in psychological studies.

As with previous literature (e.g., Curran & Bauer, 2011; Curran et al., 2012), we first mainly discussed models that include within-person effects as fixed effects only. The discussion of detrending and the derivations in Appendix B are based on this assumption. Our data analysis results from both empirical examples coincidentally did not support the inclusion of random effects in within-person effects due to various reasons. The variance of within-person effects was not significantly different from 0 in the first example, and the convergence criteria were not met with random effects in within-person effects in the second example. On the one hand, our experiences were similar to those of modeling time-varying covariates from other researchers (e.g., Singer & Willett, 2003). On the other hand, our results do not imply or suggest that random effects are not needed at all when studying within-person effects in psychology. Therefore, we conducted two simulation studies in which the within-person effect variance was allowed to be nonzero. The simulation results from nonzero variance conditions supported our recommendations on both centering and detrending issues.

It is worth also noting that the derivations in Appendix B are based on an assumption of $\overline{time}_{i.} = \overline{time}_{..}$ so that the estimates can be in simple forms. In addition, the two-step approach proposed by Curran and Bauer (2011) is similar to the one-step multilevel approach with time added as a Level-1 covariate (Curran et al., 2012; Hoffman & Stawski, 2009), also when $\overline{time}_{i.} = \overline{time}_{..}$. Therefore, when the assumption of $\overline{time}_{i.} = \overline{time}_{..}$ cannot be met, the results from the two-step approach proposed by Curran and Bauer may be more discrepant than those from the one-step multilevel approach. Future research should look into the relations among different detrending approaches when this assumption does not hold.

Our Recommendations

Centering is needed when disaggregating between- and within-person effects, and we recommend the person-mean centering approach on the time-varying predictor. Whether to perform detrending depends on study context and the substantive research question of interest, and thus detrending may or may not be needed when the time-varying predictor and/or outcome exhibit trends. If one is interested in the relation after controlling for the time effect, detrending is needed. If one believes the changes in the time-varying variables are attributable to study design or are of research interest, detrending may not be needed. When detrending is not needed, again, we recommend the person-mean centering approach. When detrending is needed, we suggest using one-step multilevel modeling with time included as a Level-1 covariate for the purpose of detrending the time-varying predictor or studying the relation between two time-varying variables net the time effect. The one-step approach is easy to implement and yields more accurate and precise results than detrending the predictor via two steps. We hope that our recommendations can help researchers better use multilevel models for disaggregating between- and within-person effects.

References

Bollen, K., & Curran, P. (2006). *Latent curve models: A structural equation perspective*. Hoboken, NJ: Wiley.

Brown, W. (1910). Some experimental results in the correlation of mental abilities. *British Journal of Psychology, 3,* 296–322.

Carlin, B. P., & Louis, T. A. (2000). *Bayes and empirical Bayes methods for data analysis* (2nd ed.). Boca Raton, FL: Chapman & Hall/CRC.

Cattell, R. B. (1946). *The description and measurement of personality*. New York: World Book.

Curran, P. J., & Bauer, D. J. (2011). The disaggregation of within-person and between-person effects in longitudinal models of change. *Annual Review of Psychology, 62,* 583–619.

Curran, P. J., Lee, T., Howard, A. L., Lane, S., & MacCallum, R. A. (2012). Disaggregating within-person and between-person effects in multilevel and structural equation growth models. In J. R. Harring & G. R. Hancock (Eds.), *Advances in longitudinal methods in the social and behavioral sciences* (pp. 217–253). Charlotte, NC: Information Age.

Derogatis, L. R. (1992). *BSI: Administration, scoring and procedures manual—II*. Towson, MD: Clinical Psychometric Research.

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

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

Galton, F. (1865). Hereditary talent and character. *MacMillan's Magazine, 12,* 157–166, 318–327.

Hoffman, L., & Stawski, R. S. (2009). Persons as contexts: Evaluating between-person and within-person effects in longiutdinal analysis. *Research in Human Development, 6,* 97–120.

Hofmann, D. A., & Gavin, M. B. (1998). Centering decisions in hierarchical linear models: Theoretical and methodological implications for organizational science. *Journal of Management, 24,* 623–641.

Ke, Z., & Wang, L. (2014). Detecting individual differences in change: Methods and comparisons. *Structural Equation Modeling*. http://dx.doi.org/10.1080/10705511.2014.936096

Kreft, I. G. G., de Leeuw, J., & Aiken, L. S. (1995). The effect of different forms of centering in hierarchical linear models. *Multivariate Behavioral Research, 30,* 1–21.

Laird, N. M., & Ware, J. H. (1982) Random-effects models for longitudinal data. *Biometrics, 38,* 963–974.

McArdle, J. J., & McDonald, R. P. (1984). Some algebraic properties of the reticular action model for moment structures. *British Journal of Mathematical and Statistical Psychology, 27,* 234–251.

McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: Erlbaum.

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

Molenaar, P. C. M., & Campbell, C. G. (2009). The new person-specific paradigm in psychology. *Current Directions in Psychological Science, 18,* 112–117.

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

Singer, J., & Willett, J. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York: Oxford University Press.

Spearman, C. (1910). Correlation calculated from faulty data. *British Journal of Psychology, 3,* 271–295.

Wang, L., & Grimm, K. (2012). Investigating reliabilities of intra-individual variability indicators. *Multivariate Behavior Research, 47,* 1–31.

Willett, J. B. (1989). Some results on reliability for the longitudinal measurement of change: Implications for the design of studies of individual growth. *Educational & Psychological Measurement, 49,* 587–602.

Zhang, Q., & Wang, L. (2014). Aggregating and testing intra-individual correlations: Methods and comparisons. *Multivariate Behavioral Research, 49,* 130–148.

(*Appendices follow*)

Appendix A

Point Estimates of Individual Within-Person Effects

For individual i, the estimated within-person effect from ordinary least squares, $\hat{\gamma}_{10}^{(i)}$, with different detrending actions in Figure 1 can be derived as follows.

No Detrending, or the Person-Mean Centering Approach

Using the person-mean centering approach, we have $cx_{it} = (x_{it} - \bar{x}_{i.})$ for the cX variable in Figure 1a. For individual i, from the simple regression of regressing Y_i on cX_i, we have the estimated within-person effect as

$$\hat{\gamma}_{10}^{(i)} = r_{y.x}^{(i)} \frac{s_y^{(i)}}{s_x^{(i)}}, \tag{A1}$$

where $\hat{\gamma}_{10}^{(i)}$ is the individual i's estimated within-person effect, $r_{y.x}^{(i)}$ is the correlation between Y_i and X_i for individual i, and $s_y^{(i)}$ and $s_x^{(i)}$ are sample standard deviations of Y_i and X_i, respectively. Note that within each individual, $\bar{x}_{i.}$ is a constant. Therefore, we have $r_{y.cx}^{(i)} = r_{y.x}^{(i)}$ and $s_x^{(i)} = s_{cx}^{(i)}$.

Detrend X Only by the Two-Stage Method in Curran and Bauer (2011)

From the two-stage regression, first predicting X_i from time and then predicting Y_i from dX_i (detrended X_i; residuals of the regression for predicting X_i from $TIME_i$, $dx_{it} = x_{it} - \hat{a}_{x0i} - \hat{a}_{x1i}time_{it}$), we have the estimated within-person effect for individual i as

$$\begin{aligned}\hat{\gamma}_{10}^{(i)} &= r_{y.dx}^{(i)} \frac{s_y^{(i)}}{s_{dx}^{(i)}} \\ &= \left(r_{y.x}^{(i)} s_y^{(i)} s_x^{(i)} - r_{y.time}^{(i)} s_y^{(i)} s_{time}^{(i)} r_{x.time}^{(i)} \frac{s_x^{(i)}}{s_{time}^{(i)}} \right) \frac{1}{(1-(r_{x.time}^{(i)})^2)(s_x^{(i)})^2} \\ &= \frac{\left(r_{y.x}^{(i)} - r_{y.time}^{(i)} r_{x.time}^{(i)}\right)}{\left(1-\left(r_{x.time}^{(i)}\right)^2\right)} \left(\frac{s_y^{(i)}}{s_x^{(i)}}\right),\end{aligned} \tag{A2}$$

where $r_{x.time}^{(i)}$ and $r_{y.time}^{(i)}$ are the correlations of X_i with $TIME_i$ and Y_i with $TIME_i$ for individual i.

Add Time as a Covariate

From the multiple regression with two independent variables cX_i and $TIME_i$, we have the estimated within-person effect for individual i as

$$\hat{\gamma}_{10}^{i} = \frac{\left(r_{y.x}^{(i)} - r_{y.time}^{(i)} r_{x.time}^{(i)}\right)}{\left(1-\left(r_{x.time}^{(i)}\right)^2\right)} \left(\frac{s_y^{(i)}}{s_x^{(i)}}\right). \tag{A3}$$

Detrend Both X and Y

From the regression of predicting dY_i (detrended Y_i, $dy_{it} = y_{it} - \hat{a}_{y1i}time_{it}$) from dX_i (detrended X_i; $dx_{it} = x_{it} - \hat{a}_{x0i} - \hat{a}_{x1i}time_{it}$), we have the estimated within-person effect for individual i as

$$\begin{aligned}\hat{r}_{10}^{(i)} &= r_{dy.dx}^{(i)} \frac{s_{dy}^{(i)}}{s_{dx}^{(i)}} \\ &= \frac{r_{y.x}^{(i)} s_y^{(i)} s_x^{(i)} - r_{y.time}^{(i)} s_y^{(i)} r_{x.time}^{(i)} s_x^{(i)}}{\left(s_{dx}^{(i)}\right)^2} \\ &= \frac{\left(r_{y.x}^{(i)} - r_{y.time}^{(i)} r_{x.time}^{(i)}\right)}{\left(1-\left(r_{x.time}^{(i)}\right)^2\right)} \left(\frac{s_y^{(i)}}{s_x^{(i)}}\right).\end{aligned} \tag{A4}$$

Detrend Y Only

From the regression of predicting dY_i (detrended Y_i, $dy_{it} = y_{it} - \hat{a}_{y1i}time_t$) from cX_i, we have the estimated within-person effect for individual i as

$$\begin{aligned}\hat{\gamma}_{10}^{(i)} &= r_{dy.x}^{(i)} \frac{s_{dy}^{(i)}}{s_x^{(i)}} \\ &= \frac{\left(r_{y.x}^{(i)} s_y^{(i)} s_x^{(i)} - r_{x.time}^{(i)} s_x^{(i)} s_{time}^{(i)} r_{y.time}^{(i)} \frac{s_y^{(i)}}{s_{time}^{(i)}} \right)}{\left(s_x^{(i)}\right)^2} \\ &= \left(r_{y.x}^{(i)} - r_{x.time}^{(i)} r_{y.time}^{(i)}\right) \frac{s_y^{(i)}}{s_x^{(i)}}.\end{aligned} \tag{A5}$$

(*Appendices continue*)

Appendix B

Point Estimates of Fixed-Effects Within-Person Effects Aggregated Across Individuals

With repeated data from multiple individuals, we can aggregate individual within-person effects via multilevel models. In the multilevel models, when $\bar{time}_{i.} = \bar{time}_{..}$, the correlation of $\bar{x}_{i.}$ with $\hat{a}_{1i}(time_{it} - \bar{time}_{..})$ or $x_{it} - \bar{x}_i$ is 0, which leads to orthogonal between- and within-person effects. From those multilevel models, point estimates of the fixed-effects within-person effects aggregated across individual can be obtained by ordinary least squares (OLS). Note that when the within-person effects are included only as fixed effects (or only intercepts are included as random effects) in the multilevel models, those OLS point estimates for fixed-effects parameters are equal to maximum likelihood point estimates.

For different detrending actions in Figure 1, the OLS estimates for the aggregated fixed-effects within-person effects from the reduced multilevel models, $\hat{\gamma}_{10}$s, can be derived as follows.

No Detrending, or the Person-Mean Centering Approach (Model M3r [See Also Equation 10])

$$\hat{\gamma}_{10} = r_{y.cx}\frac{s_y}{s_{cx}}, \tag{B1}$$

where $\hat{\gamma}_{10}$ is the estimated within-person effect aggregated across individuals, $r_{y.cx}$ is the correlation between Y and cX, and s_y and s_{cx} are sample standard deviations of Y and cX across all time points and all individuals, respectively. Note that $\bar{x}_{i.}$ may be different across individuals, and thus $r_{y.cx}$ may be different from $r_{y.x}$.

Detrend *X* Only by the Two-Stage Method in Curran and Bauer (2011; Model M4r-a [See Also Equations 12 or 13)

$$\hat{\gamma}_{10} = r_{y.dx}\frac{s_y}{s_{dx}}, \tag{B2}$$

where $r_{y.dx}$ is the correlation between Y, and dX and s_y and s_{dx} are sample standard deviations of Y and dX across all time points and all individuals, respectively.

Add Time as a Covariate (See Also Equations 14 or 15 With Constraints)

$$\hat{\gamma}_{10} = \frac{(r_{y.cx} - r_{y.time}r_{cx.time})}{(1 - r^2_{cx.time})}\left(\frac{s_y}{s_{cx}}\right), \tag{B3}$$

where $r_{y.cx}$, $r_{y.time}$, and $r_{cx.time}$ are the correlations of Y with cX, of Y with grand-mean centered time, and of cX with grand-mean centered time, and s_y and s_{cx} are sample standard deviations of Y and cX across all time points and all individuals, respectively. Note that the estimates shown in Equation B3 are valid for the multilevel models in Equations 14 or 15 with only intercepts included as random effects. We name this *Model M5r-b*.

Detrend Both *X* and *Y*

The fitted two-step model, *Model M6r*, is

$$\begin{cases} y_{it} = a_{y0i} + a_{y1i}(time_{it} - \bar{time}_{..}) + r_{yit} \\ x_{it} = a_{x0i} + a_{x1i}(time_{it} - \bar{time}_{..}) + r_{xit} \\ dy_{it} = \gamma_{0i} + \gamma_{1i}\hat{r}_{xit} + e_{it} \\ \gamma_{0i} = \gamma_{00} + \gamma_{01}\hat{a}_{0i} + u_{0i} \\ \gamma_{1i} = \gamma_{10} \end{cases}, \tag{B4}$$

where $dy_{it} = y_{it} - \hat{a}_{y1i}(time_{it} - \bar{time}_{..})$, $\hat{r}_{xit} = x_{it} - \hat{a}_{x0i} - \hat{a}_{x1i}(time_{it} - \bar{time}_{..}) = dx_{it}$, and $\hat{a}_{0i} = \bar{x}_{i.}$. The first two equations are for detrending Y and X for the first step, and the last three equations form the multilevel model of studying the relations between detrended Y and detrended X for the second step. The estimate is

$$\hat{\gamma}_{10} = r_{dy.dx}\frac{s_{dy}}{s_{dx}}, \tag{B5}$$

where $r_{dy.dx}$ is the correlations of dY with dX, and s_{dy} and s_{dx} are sample standard deviations of dY and dX across all time points and all individuals, respectively.

Detrend *Y* Only

The fitted two-step model, *Model M7r*, is

$$\begin{cases} y_{it} = a_{y0i} + a_{y1i}(time_{it} - \bar{time}_{..}) + r_{yit} \\ dy_{it} = \gamma_{0i} + \gamma_{1i}(x_{it} - \bar{x}_{i.}) + e_{it} \\ \gamma_{0i} = \gamma_{00} + \gamma_{01}\bar{x}_{i.} + u_{0i} \\ \gamma_{1i} = \gamma_{10} \end{cases}, \tag{B6}$$

where $dy_{it} = y_{it} - \hat{a}_{y1i}(time_{it} - \bar{time}_{..})$; and the estimate is

$$\hat{\gamma}_{10} = r_{dy.cx}\frac{s_{dy}}{s_{cx}}, \tag{B7}$$

where $r_{dy.cx}$ is the correlations of dY with cX, and s_{dy} and s_{cx} are sample standard deviations of dY and cX across all time points and all individuals, respectively.

Received September 2, 2013
Revision received October 29, 2014
Accepted November 14, 2014 ■