Psychological Methods

ISSN: 1082-989X

https://doi.org/10.1037/met0000663

How Should We Model the Effect of "Change"—Or Should We?

Ethan M. McCormick[1] and Daniel J. Bauer[2]

[1] Methodology and Statistics Department, Institute of Psychology, Leiden University

[2] Department of Psychology and Neuroscience, University of North Carolina at Chapel Hill

OPEN DATA OPEN MATERIALS

Abstract

There have been long and bitter debates between those who advocate for the use of residualized change as the foundation of longitudinal models versus those who utilize difference scores. However, these debates have focused primarily on modeling change in the outcome variable. Here, we extend these same ideas to the covariate side of the change equation, finding similar issues arise when using lagged versus difference scores as covariates of interest in models of change. We derive a system of relationships that emerge across models differing in how time-varying covariates are represented, and then demonstrate how the set of logical transformations emerges in applied longitudinal settings. We conclude by considering the practical implications of a synthesized understanding of the effects of difference scores as both outcomes and predictors, with specific consequences for mediation analysis within multivariate longitudinal models. Our results suggest that there is reason for caution when using difference scores as time-varying covariates, given their propensity for inducing apparent inferential inversions within different analyses.

Translational Abstract

There have been long and bitter debates between those who advocate for the use of residualized change (regressing a variable on itself measured at some time lag prior) as the foundation of longitudinal models versus those who utilize difference scores (subtracting prior from current status). However, most of the methodological work on this topic has focused on the outcome variable in different models. Here, we extend these same issues to the covariates—or predictors—in longitudinal models of change and find that similar issues arise when using lagged versus difference score predictors. We show how apparently distinct models using different versions of time-varying covariates are, in fact, simply repackaged versions of the same predictive information and are related through a set of equations that we lay out. We then work through several applied examples across traditional and multilevel regression models. We conclude by considering the issues that arise where a time-varying variable acts as both outcome and predictor—with a specific focus on mediation analysis within multivariate longitudinal models. Our results suggest that users should exercise caution when using change scores as time-varying covariates—not because they are wrong per se, but because they can introduce apparent inferential inversions that can mislead researchers when drawing substantive conclusions.

Emilio Ferrer served as action editor.

Ethan M. McCormick https://orcid.org/0000-0002-7919-4340

Ethan M. McCormick was supported with funds from the NWO (Nederlandse Organisatie voor Wetenschappelijk Onderzoek), Domain Social Sciences and Humanities (SSH) Sector Plan: Resilience in Youth, and Jacobs Foundation Fellowship Program (2023-1510-00). The authors thank the authors who provided access to the empirical data for demonstration purposes. The authors have no conflicts of interest to disclose. Data used in the preparation of this article were partially obtained from the Adolescent Brain Cognitive Development (ABCD) Study (https://abcdstudy.org), held in the NIMH Data Archive (NDA). This is a multisite, longitudinal study designed to recruit more than 10,000 children aged 9–10 years and follow them over 10 years into early adulthood. The ABCD Study® is supported by the National Institutes of Health (NIH) and additional federal partners under Awards U01DA041048, U01DA050989, U01DA051016, U01DA041022, U01DA051018, U01DA051037, U01DA050987, U01DA041174, U01DA041106, U01DA041117, U01DA041028, U01DA041134, U01DA050988, U01DA051039, U01DA041156, U01DA041025, U01DA041120, U01DA051038, U01DA041148, U01DA041093, U01DA041089, U24DA041123, and U24DA041147. A full list of supporters is available at https://abcdstudy.org/federal-partners.html. A listing of participating sites and a complete listing of the study investigators can be found at https://abcdstudy.org/consortium_members/. ABCD consortium investigators designed and implemented the study and/or provided data but did not necessarily participate in the analysis or writing of this report. This article reflects the views of the authors and may not reflect the opinions or views of the NIH or ABCD consortium investigators. The code required to reproduce all analyses, as well as more- extensive derivations, can be found at https://osf.io/yc96v/ or in the associated online supplemental materials.

The data are available at https://doi.org/10.17605/OSF.IO/YC96V.

The experimental materials are available at https://doi.org/10.17605/OSF.IO/YC96V.

Correspondence concerning this article should be addressed to Ethan M. McCormick, Methodology and Statistics Department, Institute of Psychology, Leiden University, Wassenaarseweg 52, 2333 AK Leiden, The Netherlands. Email: e.m.mccormick@fsw.leidenuniv.nl

Keywords: time-varying covariates, raw score change, residualized change, multilevel modeling, latent change score model

Supplemental materials: https://doi.org/10.1037/met0000663.supp

One of the central goals of the psychological and behavioral sciences is to understand how processes unfold over time—within individuals, dyads, organizations, countries, or other units of interest.[1] Longitudinal data not only allows researchers to chart the course of change, but also to prospectively predict later outcomes using predictors observed at earlier time points (Curran & Hancock, 2021; Curran et al., 2010; McCormick et al., 2023; McNeish & Matta, 2020). Identifying these prospective relations is often of key interest to researchers, both in the context of panel data analysis (e.g., does physical activity predict stroke symptom recovery, Kollen et al., 2005) and when considering intensive longitudinal data (e.g., does positive affect early in the day predict drinking behavior later in the day?, Howard et al., 2015). The best way to model such effects, however, remains unclear. Indeed, time-varying covariates (TVCs) are often incorporated into longitudinal models with exclusively contemporaneous effects, where the value of the TVC at a given time point influences the value of the outcome at that same time point. While contemporaneous effects can be informative, they do not support inferences about prospective prediction.

In an effort to preserve temporal precedence, researchers have thus considered several alternative ways of embedding prospective effects of TVCs within their models. For a given TVC x_t, one common approach is to include the lag(1) version of the covariate (e.g., x_{t-1}, McNeish & Matta, 2020), where the goal is to assess the effect of prior covariate status on the current value of the outcome. Another strategy is to use a change score for the TVC ($\Delta x = x_t - x_{t-1}$) as a predictor (e.g., Grimm et al., 2012). With this approach, the idea is to see how the magnitude of change in the TVC between the prior and current time point predicts the outcome at the current time point. For both of these approaches, researchers might choose to control for the contemporaneous relationship between the TVC and outcome to isolate the prospective effects above-and-beyond concurrent associations. Still, a third option is to include the prior (rather than the contemporaneous) observation of the TVC with the change score (i.e., x_{t-1} with Δx) in an effort to control for the starting point when evaluating the effect of change. While the decision between these alternatives may seem to be a simple matter of addressing the specific research hypothesis at hand, there are some hidden relationships between these models that can lead to very different substantive interpretations depending on the option chosen. In this treatment, we outline these relationships and the complications they bring about, which harken back to long-standing debates on the relative merits of using residualized and raw change scores. While these debates historically focused on the definition of change in an outcome variable (Castro-Schilo & Grimm, 2018; Cronbach & Furby, 1970; Willett, 1997), here we show that many of the same principles also apply on the predictor side of the equation. At times, the choice of how to represent prospective effects within the model can even result in apparent inversions of effects. Although these principles can be illustrated through straightforward transformations, they do not appear to be widely known in the applied research community. Thus, our purpose is to bring greater clarity to the choice of models for capturing prospective effects of TVCs and to illustrate this within a variety of common longitudinal modeling approaches.

Time-Varying Covariates (TICs)

We can begin by drawing a conceptual and statistical distinction between time-invariant and TVCs. TICs are predictors whose values remain constant over time, either representing unchanging characteristics of the person or variables that were only measured once, typically at the outset of the longitudinal study (e.g., baseline measures). In contrast, TVCs are repeatedly measured predictors which can take on different values from one point to the next. Variation on the TVC over time is thought to be predictive of variation in the outcome over time.

In modeling TVCs, it is often important to distinguish within-person variability versus between-person variability (Curran & Bauer, 2011). For example, in predicting heart rate from exercise, one would expect to observe both a between-person relationship—those who exercise more on average have lower average heart rates—as well as a within-person relationship—a person's heart rate increases at times when they are exercising. How within-person versus between-person effects of TVCs are distinguished differs between modeling frameworks (McCormick et al., 2023; McNeish & Matta, 2018, 2020), but the goals are similar across techniques. There have been many treatments of how to properly include TVCs in models of change over time (Curran & Bauer, 2011; Gottfredson, 2019; Hoffman & Stawski, 2009; McCormick, 2021; McNeish & Matta, 2020; Wang & Maxwell, 2015), and how this differs from multivariate growth modeling (see Curran & Hancock, 2021; McCormick et al., 2023), so we do not expand on these topics here. Our concern, instead, is with modeling prospective effects of TVCs.

For TICs, the modeling of prospective effects is relatively straightforward and primarily is facilitated through study design. If the TIC was measured in advance of the repeated measures, then the effect is considered to be a prospective one. Even if the TIC was measured contemporaneously with the first observation of the outcome, effects of the TIC on aspects of subsequent change in the repeated measures are typically still interpreted as prospective. For TVCs, by contrast, prospective prediction is made more challenging by the fact that TVCs are usually collected at the same time points as the repeated measures of the outcome. How best to tease out concurrent associations versus prospective effects when both predictors and outcomes are measured contemporaneously and repeatedly remains uncertain.

Residualized and Raw Change

A long-standing debate in longitudinal modeling concerns the use of residualized versus raw change scores, with sometimes

[1] For our purposes here, we will assume the units of analysis are individuals, although all the conclusions we draw generalize to these other kinds of units.

acrimonious exchanges stretching back over decades (Cronbach & Furby, 1970; Lord, 1956; Willett, 1997). At issue is the best way to measure change. An intuitively appealing measure of change is the simple difference score, or raw change score, defined as $\Delta y_{t,t-1} = y_t - y_{t-1}$. Detractors of difference scores, however, have argued that they are inherently unreliable, combining uncertainty in the measurement of both y_t and y_{t-1} (Cronbach & Furby, 1970). Moreover, one must assume invariance of measurement for y, as any recalibration of responses across time (e.g., sensitization to what is being measured) would be conflated with true change (Bereiter, 1963). An alternative is residualized change, in which change is measured via the residual of a regression equation, (i.e., $y_t - \hat{y}_t$), where $\hat{y}_t$ is the predicted value obtained from regressing current status (y_t) on prior status (y_{t-1}). In this approach, change is redefined to be the difference between current status and predicted current status based on prior status. Residualized change too has been critiqued, with the chief argument against it being that the relative rather than absolute change captured by this approach lacks intuitive interpretations (Willett, 1997). Many myths associated with the debate between raw and residualized change have since been debunked. For instance, Rogosa and Willett (1983, 1985) demonstrated that under many realistic conditions, the difference score has a higher reliability than anticipated by earlier research. Additionally, the difference score forms the basis for many models for assessing longitudinal change over time, including paired-samples *t* tests, repeated measures analysis of variance, and growth models (Rogosa & Willett, 1985).

Mathematically, it can also be shown that the difference score is a special case of the residualized change model in which the slope (autoregressive [AR] effect of y_t on y_{t-1}) is set to 1. We can see this below (see Castro-Schilo & Grimm, 2018, for a more thorough treatment), beginning with the regression equation:

$$y_t = \beta_0 + \beta_1 y_{t-1} + \varepsilon_t \tag{1}$$

can be rearranged so that:

$$y_t - \beta_1 y_{t-1} = \beta_0 + \varepsilon_t, \tag{2}$$

when $\beta_1 = 1$, we have:

$$y_t - \left(1 * y_{t-1}\right) = y_t - y_{t-1} = \Delta y_{t,t-1} = \beta_0 + \varepsilon_t. \tag{3}$$

which is the difference score model.[2] One can argue the consequences of this observation either way—the difference scores is just a constrained residualized change score so the distinction is not as stark as it first appears, or that the constraint of $\beta_1 = 1$ is a (typically) untested assumption of raw score change that one should not necessarily expect to comport with the observed data. While these relationships have not resolved debates surrounding the use of residualized and difference change scores—especially in the context of whether or not to control for baseline status in the assessment of experimental effects—they have demystified the superficially incongruent forms the different models take.

Note that the debate outlined above centered on residualized versus change scores as outcomes in longitudinal models, whereas the TVC is a repeatedly measured predictor. Nevertheless, we show that many of the same considerations encountered on the y-side also emerge on the x-side of the equation, where choice of residualized versus raw change can induce seemingly discordant results in prospective predictions. Below we outline the general analytic relationships which underlie these differences, and then demonstrate the implications for real-data in the context of standard multiple regression and multilevel models (MLMs). Finally, we will combine what we learn here about using change scores as predictors with prior work on change scores as outcomes to better understand how these parameter transformations influence mediation analysis within the latent change score (LCS) modeling framework.

Equivalencies Between TVC Models

We can consider three scenarios for including lagged TVC relationships in models of change. Here we draw out the analytic relationships that exist between these three scenarios, independent of the specific model that is being estimated. In our subsequent empirical demonstrations, we will highlight how the following derivations emerge specifically in different modeling frameworks. Our three putative scenarios are as follows. Consider some outcome y_t, measured at time t, that is a perfect linear combination of any two of the following: a TVC measured at the same time point (x_t), the TVC at the prior time point (x_{t-1}), and the raw-score change in the TVC between the prior and current time point ($\Delta x_{t,t-1}$). Three possible arrangements exist. First, we could include both contemporaneous and lag(1) effects of x_t:

$$y_t = a * x_t + b * x_{t-1}, \tag{4}$$

where a and b represent the weights of the predictors within the linear combination. Second, we could include the contemporaneous and change effects:

$$y_t = c * x_t + d * \Delta x_{t,t-1}, \tag{5}$$

where c and d are again weights. Third, we could include the lag(1) and change effect:

$$y_t = e * x_{t-1} + f * \Delta x_{t,t-1}, \tag{6}$$

where e and f are the weights.

Knowing that $\Delta x = x_t - x_{t-1}$, we can rewrite Equations 5 and 6 in the following forms:

$$\begin{aligned} y_t &= c * x_t + d * (x_t - x_{t-1}), \\ y_t &= e * x_{t-1} + f * (x_t - x_{t-1}). \end{aligned} \tag{7}$$

Finally, we can re-arrange like terms to give the following additive expressions:

$$\begin{aligned} y_t &= (c + d) * x_t + (-d) * x_{t-1}, \\ y_t &= f * x_t + (e - f) * x_{t-1}. \end{aligned} \tag{8}$$

This algebraic reformulation illustrates three things. First, despite the fact that the three different linear combinations reflect different theoretical conceptualizations of how to capture prospective effects, they are all mathematically equivalent in the sense that the weights for one linear combination can be expressed as a direct function of

[2] This formulation of the difference score manifests directly in the latent change score model framework (e.g., Grimm et al., 2012; McArdle & Nesselroade, 1994) where the AR path between observations is set to 1 to define the latent difference factor ($\Delta\eta_{t,t-1}$).

the weights of any another (here demonstrated in the form of Equation 4). Specifically, the weights (a–f) can be related as follows:

$$a = c + d = f\,,\quad b = -d = e - f\,. \tag{9}$$

Thus, any model that expresses an outcome as a linear combination of two of the three representations of the TVC—x_t, x_{t-1}, and $\Delta x_{t,t-1}$—will be equivalent to any other. This algebra also reveals why adding all three representations of the TVC simultaneously would be ill-advised. Given their redundancies, it would be impossible to obtain unique weights for all three forms at once. Finally, it is apparent that this equivalence breaks down if any of the linear combinations is restricted to only one representation of the TVC. For instance, a linear combination consisting solely of the change predictor ($\Delta x_{t,t-1}$) would be equivalent to Equation 4 only if the a and b weights were equal in magnitude but opposite in sign. Such circumstances seem highly implausible, suggesting that change scores for TVCs should never be the sole representation of the TVC.

By contrast, there are contexts where researchers might have sound theoretical reasons to include only the contemporaneous ($x_{i,t}$) or only the lagged ($x_{i,t-1}$) effects of the TVC. For example, lagged paths might decay to zero between measurements separated by long periods of time, or on the other extreme, sampling in time-series analysis (e.g., physiological recording) might exceed biological constraints to transmit contemporaneous effects, leaving only lagged relationships. In these cases, researchers might include only one of the forms of the TVC to match the theoretical characteristics of the data, which would still allow those TVCs to be interpreted in isolation. However, in the psychological and behavioral sciences, repeated measures data have often fallen between these two extremes, and the inclusion of contemporaneous and lagged relationships is common in both panel and intensive longitudinal settings (e.g., Arizmendi et al., 2021; Asparouhov et al., 2018; Curran & Hancock, 2021; Epskamp et al., 2018; Grimm et al., 2012; McCormick et al., 2023; McNeish & Matta, 2020). Therefore, omission of these pathways should be chosen with care to avoid bias arising from misspecification.

TVCs in the General Linear Model

We can first demonstrate the relevant parameter transformations within the multiple regression framework. Here we can write out the model expressions of y as linear combinations of different forms of the TVC, mimicking Equations 4–6 but with the addition of a regression intercept and person-specific residuals. The first, corresponding to Equation 4, models the contemporaneous and lagged effect of $x_{i,t}$:

$$y_{i,t} = \beta_0 + \beta_1 x_{i,t} + \beta_2 x_{i,t-1} + \varepsilon_{i,t}, \tag{10}$$

the second, corresponding to Equation 5, includes the contemporaneous and change effect:

$$y_{i,t} = \beta_0 + \beta_3 x_{i,t} + \beta_4 \Delta x_i + \varepsilon_{i,t}, \tag{11}$$

and the third, corresponding to Equation 6, includes the lagged and change effect:

$$y_{i,t} = \beta_0 + \beta_5 x_{i,t-1} + \beta_6 \Delta x_i + \varepsilon_{i,t}. \tag{12}$$

The expectation of y in these equations is as follows:

$$\begin{aligned}\mathbb{E}[y_{i,t}] &= \mathbb{E}[\beta_0 + \beta_1 x_{i,t} + \beta_2 x_{i,t-1}]\\ &= \mathbb{E}[\beta_0 + \beta_3 x_{i,t} + \beta_4 \Delta x_i]\\ &= \mathbb{E}[\beta_0 + \beta_5 x_{i,t-1} + \beta_6 \Delta x_i].\end{aligned} \tag{13}$$

The algebraic relationships explored in the prior section show that

$$\beta_1 x_{i,t} + \beta_2 x_{i,t-1} = \beta_3 x_{i,t} + \beta_4 \Delta x_i = \beta_5 x_{i,t-1} + \beta_6 \Delta x_i. \tag{14}$$

Given these equalities, β_0 obtains the same value in all three expressions. Furthermore, similar to the weights first derived, we can re-express and re-arrange the regression coefficients in the same fashion (see Equations 7 and 8) to give the following relationships:

$$\beta_1 = \beta_3 + \beta_4 = \beta_5,\quad \beta_2 = -\beta_4 = \beta_5 - \beta_6. \tag{15}$$

While we have framed these transformations in terms of multiple regression, this model subsumes many special cases of the general linear model such as analysis of covariance, and the same system of relationships would emerge in generalized linear models with linear prediction components (e.g., logits in logistic regression). These parameter relations hold even when we alter ancillary parts of the model, such as when expanding the model to also include TICs or control variables, such as the AR effect of y ($y_{i,t-1}$). In effect, as long as the linear combinations from above remain unaltered within the regression models, their equivalence will continue to hold. We turn to an empirical demonstration to illustrate the relevant points.

Empirical Example (Gray Matter and Cognitive Performance)

To facilitate our example, we drew two-wave data from the Adolescent Brain and Cognition Development (ABCD Study, Casey et al., 2018), including a measure of cognitive performance (verbal intellect and language, Luciana et al., 2018) and cortical surface area from the prefrontal cortex (for a description of the relevant measures in this sample, see Michel et al., 2023). Here, we will use the cognitive measure as the outcome (y_t) and cortical surface area as the time-varying predictor (x_t).[3] We can fit three versions of the model corresponding to Equation 13, which are displayed in Table 1 (Models 1–3). Lined up side-by-side, the equivalencies jump off the page, where both the estimates and standard errors behave as expected. For instance, the effect of x_{t-1} in Model 1 ($B = 1.612$, $SE = 0.338$) shows the $b = -d$ relationship with the effect of Δx_i in Model 2 ($B = -1.612$, $SE = 0.338$). Less obviously, we can see that the effect of x_{t-1} from Model 1 ($B = 1.612$, $SE = 0.338$) is the same as the effect of x_{t-1} ($B = 2.124$, $SE = 0.101$) minus the effect of $\Delta x_{t,t-1}$ ($B = 0.513$, $SE = 0.335$) from Model 3 (i.e., $b = e - f$).

In these initial models, we did not include any additional predictors in the model, however, we could do so to without influencing the equivalencies across models. To demonstrate this, we ran the same set of models but controlling for a TIC, the age at the first age of assessment (Table 1; Models 4–6). Although the values of the

[3] Code to replicate all analyses is available here: https://osf.io/yc96v/.

Table 1
Equivalent TVC Models in the General Linear Model

	TVC only			TVC + TIC		
Predictor	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
x_t	0.513 (0.335)	2.124*** (0.101)		0.853* (0.331)	2.108*** (0.099)	
x_{t-1}	1.612*** (0.338)		2.124*** (0.101)	1.256*** (0.334)		2.108*** (0.099)
$\Delta x_{t,t-1}$		−1.612*** (0.338)	0.513 (0.335)		−1.256*** (0.334)	0.853* (0.331)
TIC				0.185*** (0.013)	0.185*** (0.013)	0.185*** (0.013)
R^2	0.057	0.057	0.057	0.083	0.083	0.083

Note. R^2 is the proportion variance explained; the intercept and time-specific residual variance are omitted for brevity, but were equal in value across the three models. TVC = time-varying covariate; TIC = time-invariant covariate.
* $p < .05$. *** $p < .001$.

coefficients are different between the two sets of models (1–3 vs. 4–6), the relationships between the coefficients remain identical (from Equation 15). The same holds for models where we include an AR effect of the outcome $y_{i,t-1}$ as a predictor (see Section 1.2 in the online supplemental materials code).

Thus far we have demonstrated the equivalencies between TVC forms without considering the implications for interpretation and inference in these various models. We can return to Models 1–3 to center our discussion. First, let us consider the contemporaneous effect of x_t on y_t. In Model 1, which includes the lagged effect, the contemporaneous effect is small and nonsignificant. By contrast, in Model 2, which includes the change score rather than the lagged effect, the contemporaneous effect is larger and statistically significant. That is, in Model 1, we would conclude there is no within-time effect of the covariate on the outcome, whereas in Model 2, we would conclude that such an effect exists. Second, consider Model 3, for which the lagged effect is equal to the contemporaneous effect in Model 2, and the change score effect is equal to the contemporaneous effect from Model 1. Yet a researcher fitting Model 3 would interpret these effects very differently than if they had fit either Model 1 or 2. Third, and perhaps even more confounding, the lagged effect in Model 1 is equal in magnitude but opposite to the change score effect in Model 2, despite both effects being intended to convey prospective prediction while controlling for the contemporaneous value of the TVC. The effect in Model 1 implies that those with higher prior levels of cortical surface area show higher levels of cognitive performance a year later, while the latter suggests that those who show increases in surface area will show lower cognitive performance, in each case while controlling for concurrent associations between cortical surface area and cognitive performance. When phrased carefully, we can see that these are distinct questions, although the results provide an ambiguous picture of how surface area is prospectively linked to cognitive performance, and unsuspecting substantive researchers could easily draw opposing conclusions from the two sets of results. We will return to recommendations for how to avoid these misinterpretations in a later section.

One might speculate that this inversion reflects a strong negative relationship between prior status and the magnitude of change, as we might expect if there were strong floor or ceiling effects of change in the TVC. Boundary effects would limit those already high or low at the prior timepoint from further change toward those boundaries. In our sample data, however, the initial levels of prefrontal cortical surface area are only weakly negatively correlated with the magnitude of change ($r = -0.125$). Instead, the mathematical relationship between the two models reveals why this counter-intuitive sign inversion occurs—namely, when controlling for x_t, the effect of $\Delta x_{t,t-1}$ will always be equal in magnitude but opposite in sign from x_{t-1}, regardless of the correlation between x_t and $\Delta x_{t,t-1}$. As such, this is, a property of how the $\Delta x_{t,t-1}$ score is computed, rather than the characteristics of a given data set (e.g., boundary effects on the outcome or predictor).

One final point to highlight is something mentioned at the end of the first derivations of the weights, which is the result of only including the change predictor in the model. We mentioned that this would be equivalent to including both $x_{i,t}$ and $x_{i,t-1}$ in the model but constraining their parameter values to be equal in magnitude, but opposite in sign. We demonstrate this result in the empirical data, comparing the results of the regression model, where Δx_i is the only predictor with a structural equation modeling approach to the regression analysis, which allows us to include $x_{i,t}$ and $x_{i,t-1}$ but apply the relevant model constraint during estimation (see the online supplemental materials code for full model results).[4] Table 2 contains the relevant parameter estimates, which confirm this effect. Given how unlikely this constraint is to conform to reality in most substantive applications, we reiterate that using Δx_i alone as a predictor seems inadvisable.

TVCs in the MLM

Another framework within which TVCs are commonly modeled is the MLM. Although here we will focus on the multilevel instantiation of these TVC models (see Curran & Bauer, 2011 for an overview), similar results could be obtained within the latent curve modeling (LCM) framework. Within the MLM, we can model the effect of the TVC across repeated observations, either pooling the effects across time or uniquely estimating each time-specific effect. The pooling approach is standard in the MLM (which we will see below), while the time-specific approach is the default in the LCM, although constraining the effects to be equal is a common simplification in LCMs (for more in-depth explication of the differences, see McNeish & Matta, 2020).

The three forms of the MLM with the different TVC options resemble the models we saw in the regression context, but are now extended to include random effects (u terms). We can see the first model with

[4] Note that constraints could alternatively be implemented within a multilevel regression model if desired, however, this feature is not universally available in all software. We use a structural equation model approach here for convenience.

Table 2
Prediction With Only the TVC Change Score

Predictor	Regression on TVC change score	Equivalent constrained TVC lag model[a]
$\Delta x_{t,t-1}$	−0.362 (0.342)	
x_t		−0.362 (0.342)
x_{t-1}		0.362 (0.342)
-2ℓ	52,863.7	52,863.7

Note. TVC = time-varying covariate; SEM = structural equation modeling. [a] SEM coefficients are constrained to be equal in magnitude but opposite in sign; -2ℓ is the −2 log-likelihood.

the contemporaneous and lagged effect of the TVC below:

$$y_{ti} = \underbrace{\gamma_{00} + \gamma_{10}x_{ti} + \gamma_{20}x_{t-1,i}}_{\text{fixed effects}} + \underbrace{u_{0i} + u_{1i}x_{ti} + u_{2i}x_{t-1,i}}_{\text{random effects}} + r_{ti}, \quad (16)$$

where the repeated measure (y_{ti}) is modeled as a function of the fixed (or average) linear effect of the contemporaneous and lagged predictors, as well as individual deviations from that fixed effect (i.e., the random effect). While here we show both the random intercept (u_{0i}) and the random slopes (u_{1i} and u_{2i}) for completeness, we need not model all of these effects at once. Indeed, we will start with models that only include a random intercept and build up to models with random slopes as we work through our example analyses. As we will show, the equivalencies between parameter estimates across models will hold regardless of whether random slopes are included, as they are all linear expressions of equivalent form. The other two models include the version with the contemporaneous and change effect:

$$y_{ti} = \gamma_{00} + \gamma_{30}x_{ti} + \gamma_{40}\Delta x_i + u_{0i} + u_{3i}x_{ti} + u_{4i}\Delta x_i + r_{ti}, \quad (17)$$

and the version with the lagged and change effect

$$y_{ti} = \gamma_{00} + \gamma_{50}x_{t-1,i} + \gamma_{60}\Delta x_i + u_{0i} + u_{5i}x_{t-1,i} + u_{6i}\Delta x_i + r_{ti}. \quad (18)$$

The MLM expectation resembles the regression model we have seen before Equation 13 but with γ's to represent the fixed effects:

$$\begin{aligned}\mathbb{E}[y_{ti}] &= \gamma_{00} + \gamma_{10}x_{ti} + \gamma_{20}x_{t-1,i} \\ &= \gamma_{00} + \gamma_{30}x_{ti} + \gamma_{40}\Delta x_i \\ &= \gamma_{00} + \gamma_{50}x_{t-1,i} + \gamma_{60}\Delta x_i.\end{aligned} \quad (19)$$

We could also include additional time-varying or time-invariant predictors into the model, but we will leave these aside here for simplicity. Given the similarity in the form of this expectation to the one for the standard regression model Equation 13, we can expect that the fixed effects will behave along the same principles we have seen so far. Namely,

$$\begin{aligned}\gamma_{10} &= \gamma_{30} + \gamma_{40} = \gamma_{50}, \\ \gamma_{20} &= -\gamma_{40} = \gamma_{50} - \gamma_{60}.\end{aligned} \quad (20)$$

However, one potentially complicating feature we want to consider is the inclusion of random effects of the various TVCs. That is, can we expect that the equivalencies in the fixed effects across models hold when we allow individual variation to exist around these effects? For this assessment, $a - f$ are now treated as random variables, and we must use the quadratic form for computing the variances of a linear combination of random variables to establish the relationships between their variances.

The variance relationships for the first set of equivalencies are outlined below:

$$\text{Var}(a) = \text{Var}(c) + 2\,\text{Cov}(c, d) + \text{Var}(d) = \text{Var}(f), \quad (21)$$

and for the second set, a similar approach yields the following equations:

$$\text{Var}(b) = \text{Var}(d) = \text{Var}(e) - 2\,\text{Cov}(e, f) + \text{Var}(f). \quad (22)$$

Note that the quadratic form prevents negative variance values despite the inverse $b = -d$ relationship, or subtraction of point estimates in $b = e - f$. For an alternative matrix-based approach to obtaining the full covariance matrix transformations simultaneously, interested readers can refer to the Appendix. We can turn to our empirical data examples below to highlight these various transformations in practice.

Empirical Examples

To demonstrate the ubiquity of these model equivalencies, and their robustness to different model specifications, we highlight two empirical examples. In the first example (detailed more fully by Wright & Simms, 2016), 94 participants recorded their daily positive and negative affect across ∼ 100 days (Mdn = 92.5; range = 59–101 days). Building on this data, Arizmendi et al. (2021) drew historical weather data from the National Weather Service and linked it with the window of observation, and so included daily temperature recordings in addition to the affect data. Here, we tested the link between daily average temperature and individuals' self-reported negative affect. This is an attractive example since weather is a purely exogenous TVC, where we do not need to be concerned about reciprocal links from the outcome of interest over time (with the plausible assumption that none of our participants govern the current or future weather via their emotional state). Our second data example draws on ecological momentary assessment data of emotional experiences during the COVID-19 pandemic (Fried et al., 2022), where 79 subjects were pinged across 14 days ($\text{Mdn}_{\text{obs}} = 53$; range = 12–56 observations) during March of 2020 (the initial lockdown period in the Netherlands). Here, we modeled the effect of feelings of loneliness ("I felt like I lack companionship, or that I am not close to people") on feelings of anhedonia ("I couldn't seem to experience any positive feeling at all") over the two-week period. While the assumption of exogeneity is weaker than in the weather data, we nevertheless included loneliness as a TVC to reflect common practice. Our goal in doing so was to illustrate the same TVC equivalencies as before, without concern for causal attributions. The code, data, and full output associated with these analyses are available in the online supplemental material code (https://osf.io/yc96v/; McCormick, 2023b).

Fixed Slopes Model (Weather and Negative Affect)

With the weather and negative affect data, we fit MLMs with a random intercept and fixed effect of the various TVCs, represented in each possible combination. To adjust the scale of the variables, we standardized both measures before fitting the models. As anticipated, the addition of the random intercept did not influence the pattern of effects seen in the various TVC models, as given in Table 3.

Table 3
Equivalent TVC Models in the Multilevel Model With Fixed Effects

Predictor	Model 1	Model 2	Model 3
$x_{i,t}$	0.013 (0.009)	−0.014*** (0.004)	
$x_{i,t-1}$	−0.027** (0.009)		−0.014*** (0.004)
Δx_i		0.027** (0.009)	0.013 (0.009)
R^2 marginal	0.117	0.117	0.117
R^2 conditional	0.504	0.504	0.504

Note. R^2 is the proportion variance explained (Nakagawa et al., 2017); the intercept and time-specific residual variance are omitted for brevity, but were equal in value across the three models. All regressioncoefficients presented are standardized due to the rescaling of the data prior to fitting the model. TVC = time-varying covariate.
** $p < .01$. *** $p < .001$.

Indeed here we see the same pattern of both equivalencies and changes in significance that we saw in the multiple regression models, as given in Table 1. Additionally, we would make different substantive conclusions about the effect of same-day temperature ($x_{i,t}$) depending on which other form of the TVC we include in the model (nonsignificant in Model 1 but significant and negative in Model 2; Table 1). This straightforward extension of the single-level regression analysis conforms perfectly to expectations and highlights the need to be concerned about these apparent inversions within a multilevel modeling context.

Random Slopes (Loneliness and Depression During COVID)

For the loneliness and anhedonia data, we fit the same three models—with a random intercept and exclusively fixed effects for the TVCs—and a second triplet of models that also included random effects for each of the TVCs. As we have seen throughout, the same equivalencies derived initially hold for the fixed effects even in the relatively complex random-effects multilevel model. Additionally, the equivalency rules that we outlined above for the random effects hold in these models (Table 4; Models 4–6). Namely, that the variance estimates of the contemporaneous effect in Model 4 and the change effect in Model 6 are equal, as are the variance estimates for the lagged effect in Model 4 and change effect in Model 5. The variance estimates for the contemporaneous effect in Model 5 and the lagged effect in Model 6 follow the expressions in Equations 21 and 22 exactly. The correlation among random effects differs across models (following the relationships we outline in Equation A1), with an especially high correlation between the random effect of $x_{i,t-1}$ and Δx_i in Model 6. Note that the correlation between the random effects of $x_{i,t}$ and Δx_i has the opposite sign of the other two correlations (see Equation A2 for details).

Here, if our substantive question relates to how changes in loneliness relate to levels of depression during the lockdown period, we would make opposite theoretical conclusions about the direction of effect, depending on whether we controlled for contemporaneous levels of loneliness (Models 2 and 5) where there is a negative effect of Δx_i, or prior levels (Models 3 and 6) where there is a strong positive effect, as given in Table 4. To stress, neither effect is wrong, but reconciling the apparent discrepancy across models requires a more nuanced understanding of what the effect of Δx_i means conditioned on the presence of the other form of the TVC in the model. Without such a careful understanding, applied research may interpret these as contradictory, leading to confusion in the literature. As such, it may be a useful default approach to avoid using Δx_i as a TVC unless there are strong theoretical reasons for its inclusion (we will discuss this more thoroughly in the Recommendations for Applied Research section).

Table 4
Equivalent TVC Models in the Multilevel Model With Fixed and Random Effects

Predictor	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Fixed effects						
$x_{i,t}$	0.343*** (0.017)	0.392*** (0.019)		0.393*** (0.043)	0.466*** (0.055)	
$x_{i,t-1}$	0.049** (0.017)		0.392*** (0.019)	0.073** (0.023)		0.466*** (0.055)
Δx_i		−0.049** (0.017)	0.343*** (0.017)		−0.073** (0.023)	0.393*** (0.043)
Random effect variances						
$x_{i,t}$				0.086	0.141	
$x_{i,t-1}$				0.010		0.141
Δx_i					0.010	0.086
Random effect correlations						
$x_{i,t}$ with $x_{i,t-1}$				0.740		
$x_{i,t}$ with Δx_i					−0.850	
$x_{i,t-1}$ with Δx_i						0.983
R^2 marginal	0.205	0.205	0.205	0.226	0.226	0.226
R^2 conditional	0.369	0.369	0.369	0.538	0.538	0.538

Note. R^2 is the proportion variance explained (Nakagawa et al., 2017); the intercept and time-specific residual variance are omitted for brevity, but were equal in value across the three models. TVC = time-varying covariate.
** $p < .01$. *** $p < .001$.

Combining Predictors and Outcomes—Mediation With Difference Scores

To complement the extensive literature on residualized change versus difference scores on the outcome, our focus has been on how similar challenges emerge when using lagged and change variables as predictors. However, there are cases where a variable might plausibly play both roles as part of a larger path or graph model. We have seen hints of this in the literature on LCSs (Grimm et al., 2013; McArdle, 2009), where the AR parameter in a latent AR model is equivalent to the proportional parameter in a dual change score model minus 1 (see Castro-Schilo & Grimm, 2018, Equation 5 for a regression expression that highlights this point). In the traditional specification of the LCSmodel, latent difference factors ($\Delta\eta$) are treated only as outcomes (see Grimm et al., 2012, Figure 4 for an example), where the relevant issues have been well-articulated and addressed in prior research. However, when we use a difference score (latent or otherwise) as a mediator (Goldsmith et al., 2018; O'Laughlin et al., 2018; Selig & Preacher, 2009; Valente et al., 2021; Valente & MacKinnon, 2017)—that is, as both an outcome and a predictor simultaneously—we need to take care to recognize the equivalent relationships noted above and how they may influence interpretations and inferences. In the following sections, we give a brief overview of the LCS model, and then outline how the issues we raised in the univariate linear model (e.g., generalized linear model and mixed linear model) generalize to multivariate methods which involve contemporaneous, lagged, and change variables as predictors.

Time-Varying Measures in the LCS Model

For symmetry with prior sections, we will lay out the expectations for parameter estimates associated with the time-varying predictions within the LCS model. However, to fully appreciate how the model equivalencies play out, we will first sketch out the general model specification of the LCS model with two time points for simplicity. While there are many ways to parameterize a LCS, several of which involve specifying latent "phantom" variables for each repeated measure (e.g., Grimm et al., 2012), we will retain the simplest version as all of the repeated measures we will deal with here are observed (Kievit et al., 2018), and we are not embedding the latent difference within a larger path or growth model.

With respect to model equivalencies, we will begin by restating prior work (Castro-Schilo & Grimm, 2018) that addresses how parameter estimates will change when the target outcome is the contemporaneous variable (i.e., residualized change model) versus the change score (i.e., difference score model). We can first start with the residualized change model for the variable $x_{i,t}$, which takes the following form within the LCS model[5]:

$$x_{i,t} = \theta_{AR} x_{i,t-1} + \varepsilon_{i,t}. \tag{23}$$

Here, θ_{AR} is the autoregressive effect of $x_{i,t-1}$ on $x_{i,t}$—we will use θ as our general way to refer to regression parameters in these models to distinguish them from other models. To convert this equation into the difference score model, we subtract $x_{i,t-1}$ from both sides of Equation 23 and simplify to produce

$$\Delta x_i = (\theta_{AR} - 1) x_{i,t-1} + \varepsilon_{i,t}. \tag{24}$$

Note that while Equation 24 is expressed in terms of the observed Δx_i, these expressions apply equally to the latent difference, as shown in Figure 1. We can see by the expression in Equation 24 that when the lagged TVC $x_{i,t-1}$ predicts the change score (Δx_i), which in the LCS framework is typically referred to as the proportional parameter and denoted as β, this parameter equals the corresponding AR effect -1.

Next, we outline the equivalencies we have become familiar with in prior sections that hold when using different forms of a TVC. Here, we will continue using x as our variable of interest as it acts as both outcome and predictor within the LCS model. We can exclude intercept terms in our equations without any loss of generality in the expected results. We outline equations for all three versions of the model below, corresponding to Equations 4–6, respectively:

$$\begin{aligned} y_{i,t} &= \theta_1 x_{i,t} + \theta_2 x_{i,t-1} + \varepsilon_{i,t} \\ &= \theta_3 x_{i,t} + \theta_4 \Delta x_i + \varepsilon_{i,t} \\ &= \theta_5 x_{i,t-1} + \theta_6 \Delta x_i + \varepsilon_{i,t}. \end{aligned} \tag{25}$$

By the process of substitution for Δx_i, we can determine that the following parameter equivalencies hold across versions of the LCS model:

$$\theta_1 = \theta_3 + \theta_4 = \theta_5, \quad \theta_2 = -\theta_4 = \theta_5 - \theta_6. \tag{26}$$

However, unlike prior examples of different TVC models, the usual specification of the LCS model limits the use of the model where we use both the contemporaneous and change forms of the predictor (θ_3 and θ_4; Equation 5) because the variance of $x_{i,t}$ is constrained to be zero to identify the LCS, as shown in Figure 1. As such we will be primarily concerned with the equivalencies of the other two models:

$$\theta_1 = \theta_5, \quad \theta_2 = \theta_5 - \theta_6. \tag{27}$$

Empirical Examples

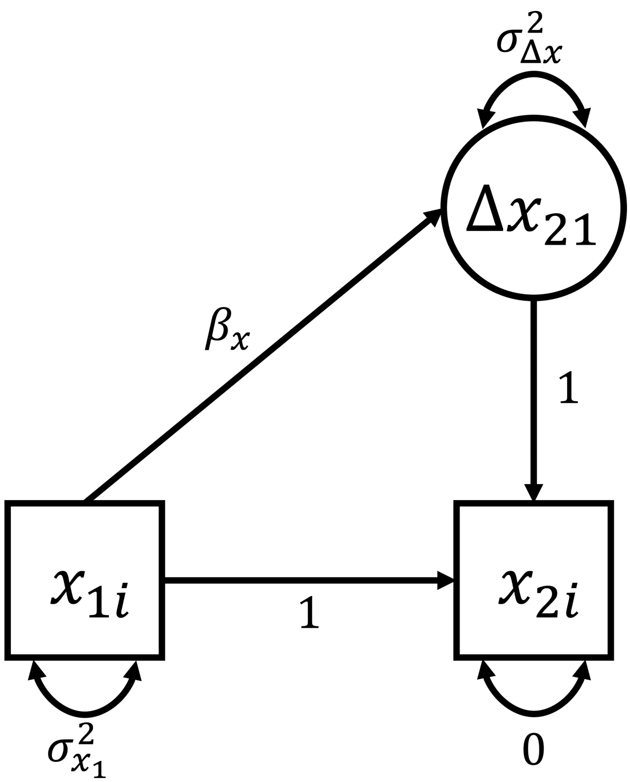
We will consider two empirical examples which exemplify the complexities of using change scores as mediators in longitudinal models. In the first simplified example, we can consider the effect of a TVC x measured at time t and $t - 1$ on an outcome measured only at time t. This example will help us to highlight the relevant parameter equivalencies. We then expand this into a more complex three-variable model, where all variables are measured repeatedly.

Simple Change Score Mediation Model (White Matter and Reading Comprehension)

For the first example, we can return to the two-wave data from the ABCD Study (Casey et al., 2018), but this time draw a measure of reading comprehension (Luciana et al., 2018) and mean diffusivity of the forceps minor white matter tract (Michel et al., 2023). Here, we will use reading comprehension as the outcome (y_t) and mean diffusivity as the time-varying predictor (x_t). We can outline

[5] Note that intercepts may or may not be estimated in the LCS model depending on the goals of the analysis, so we will leave them aside here for simplicity—including them changes nothing of what we will discuss in this section.

Figure 1
Simple Latent Change Score



Note. Here, we present the path diagram of a two-time point latent change score model with observed repeated measures. The latent difference (Δx_{21}) is parameterized by setting the autoregressive path and factor loading for x_{2i} to 1, and constraining the residual variance of x_{2i} to 0. In this form of the model, we do not include latent status "phantom" variables at each time point, and we estimate the variance of the latent difference (i.e., $\sigma^2_{\Delta x}$). Here, we include the proportional regression effect (β_x), which we will build on when we move into mediation models.

simplified versions of the two covariate models we will consider using SEM path diagrams, as shown in Figure 2, one where the LCS (Δx_{21}) serves as the mediator (A) and the other where the contemporaneous measure of the covariate (x_2) does instead (B).

We can then fit the two mediation models we laid out in Figure 2. First, examine the pathways where our time-varying x variable is an outcome rather than a covariate (A path in both diagrams; Table 5). Here, we can clearly see that the proportional path ($x_1 \rightarrow \Delta\eta$) in Figure 2a is the autoregressive path ($x_1 \rightarrow x_2$) minus 1 ($0.441 - 1 = -0.559$). Note that for autoregressive effects <0.5, this subtraction increases the magnitude of the estimated effect (for an autoregressive effect of 0, the proportional effect is -1)[6] —this will have implications for our inferences that we will explore in greater detail in our second example. In contrast, the B path is identical across models.

If we turn our attention toward the indirect effect—often the primary target of mediation analysis—we can see how this relationship between the proportional and AR effects can present a challenge for our inferences. In the LCS model, we have a significant negative indirect effect (-0.129, $SE = 0.058$, $p = .025$), while in the AR model, the indirect effect is significant and positive (0.102, $SE = 0.046$, $p = .026$). This change in sign and magnitude is the result of the proportional versus AR path being used in computing the indirect effect. Given that the effect of subtracting 1 will be quite substantial in most applications, this means that we can expect that shifting between the different model specifications is likely to lead to these sorts of inversions with some regularity. Without the benefit of side-by-side model comparisons (and indeed even with the benefit if we are not careful) we could imagine unsuspecting researchers proceeding with either of these model estimates. However, these models give inferentially opposite results, especially, if we focus on the indirect effect as the primary estimate target. Additionally, the simplified nature of this initial example makes these changes easier to spot. We can see how these issues, and the TVC equivalencies we have discussed throughout, present in more complex mediation models involving difference scores through a second empirical example.

Extended Change Score Mediation Model (Gratitude and Social Media Use)

Thus far, we have seen a simplified example that highlighted how the equivalences manifest between a model where the contemporaneous measurement of the predictor (x_2) versus the LCS ($\Delta\eta$) is used as a mediator. Our focus has been on highlighting the relationship between the AR path ($x_1 \xrightarrow{AR_x} x_2$) and the proportional path ($x_1 \xrightarrow{\beta_x} \Delta x$), where $\beta_x = AR_x - 1$, and how that will often cause the indirect effect to change in magnitude and sign. However, we have yet to see how all of the equivalencies explored thus far present in a more realistic model of empirical data. We turn to this here.

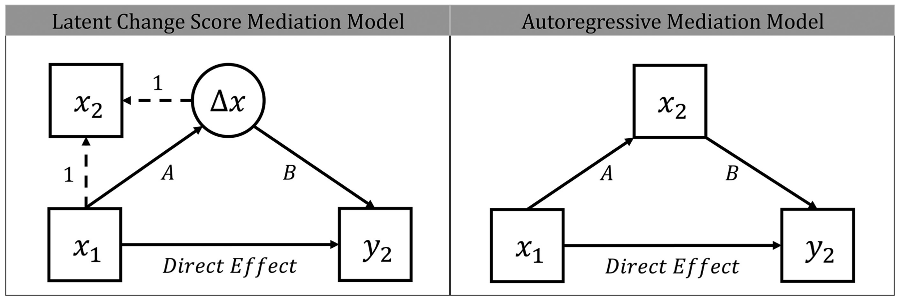
Before fitting the models to our empirical example, it is useful to expand on the two alternative formulations of a mediation model with time-varying measures (in these multivariate outcomes, the line between predictors and outcomes is murkier, so we will refer to them generally). We can see these formulations in Figure 3, where we can see a LCS mediation model (A) and an equivalent autoregressive model (B) for three variables measured twice each across four waves. We use the word "equivalent" because these models have the exact same number of parameters and fit to the data, similar to all of the models we have seen thus far. Indeed, as we strip away the apparent complexity of these models, we will see that the two models are multivariate versions of Equation 6 (lagged and change predictors) and Equation 4 (contemporaneous and lagged predictors), respectively.

We have highlighted equivalent paths between the two models which represent the same predictive pathways ($\theta_1 - \theta_{12}$) with the AR and proportional pathways labeled separately. We can focus here on the relationships between the measurements of x and y_2 (θ_1 and θ_2) to illustrate our expectations for the model results. The LCS version of the model Figure 3A corresponds to the e and f weights in Equation 6 denoting the effects of the lagged (e.g., x_1) and change (e.g., Δx_{21}) predictors, respectively. By contrast, the parameters of the AR version of the model correspond to the contemporaneous a and lagged b weights in Equation 4. Given the equivalencies between model parameters Equation 9, we can expect

[6] When dealing with time-series analyses, this means that a stationary process will have proportional paths between -2 and 0, corresponding to autoregressive paths between -1 and 1.

Figure 2

Latent Change Score and Autoregressive Mediation Models



Note. We specified two alternative mediation models using a time-varying covariate, using the latent change score (left) and contemporaneous observed measure of x (right) as the mediators. The indirect (A and B) and direct effect paths are highlighted. Variances/residuals and intercepts/means are omitted from the diagram for visual clarity.

that θ_2 should be identical between the two versions of the model (i.e., $a = f$). The $\theta_{1,AR}$ parameter ($x_1 \rightarrow y_2$) from the AR mediation model (B) should be the difference between the two parameters ($\theta_{1,LCS} - \theta_{2,LCS}$) from the LCS model (A). Finally, the proportional path (β_x) will be the AR path (AR_x) − 1.

To illustrate these points, we drew longitudinal data from a four-wave study of gratitude and social media use (Maheux et al., 2021), using the covariance and mean vector for the repeated measures provided in the article. To fit the models in Figure 3, we used a measure of social media use (measured as amount of time spent) from waves 1 and 2 (x), a measure of the subjective importance of social media use from waves 2 and 3 (y), and a measure of subjective feelings of gratitude from waves 3 and 4 (z). The principles we will highlight using this data would generalize further to a full longitudinal model with four repeated measures on each construct, however, this simplification allows us to see the main point without excessive redundancy.

We can examine coefficients for the direct effects between variables across the two models to demonstrate that the equivalencies we expect from prior models appear again when using different versions of the time-varying measures Table 6. We can first compare the proportional paths (β) with the AR paths (AR) for corresponding variables. We can see the expected $\beta = AR - 1$ relationship holds for all variables and clearly demonstrates that for variables with weaker AR stability, the proportional pathway predicting the LCS increases commensurately when using the LCS version of the model, and vice versa.

Some parameters (θ_2, θ_4, etc.) are exactly equal between models, reflecting the $a = f$ equivalency across versions of the time-varying predictors. The other parameters (θ_1, θ_3, etc.) are not equal across models, but instead, the parameter in the AR version of the model

Table 5

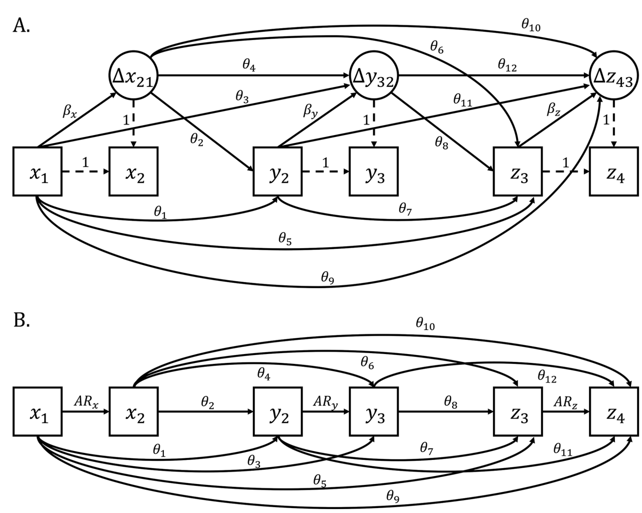
Parameter Estimates From Simple Latent Change and Autoregressive Mediation Models

Parameter (path)	Latent difference mediator	Contemporaneous mediator
$x_1 \rightarrow \Delta x$ (A path)	−0.559*** (0.009)	
$x_1 \rightarrow x_2$ (A path)	1.000[a]	0.441*** (0.009)
$\Delta x \rightarrow y$ (B path)	0.230* (0.103)	
$x_2 \rightarrow y$ (B path)		0.230* (0.103)
$x_1 \rightarrow y$ (direct effect)	0.303** (0.097)	0.073 (0.090)
Indirect effect	−0.129* (0.058)	0.102* (0.046)
-2ℓ	61,366.7	61,366.7

Note. -2ℓ is the −2 log-likelihood.
[a] Parameter is fixed rather than estimated.
* $p < .05$. ** $p < .01$. *** $p < .001$.

Figure 3

Extended Multivariate Time-Varying Mediation Models



Note. We specified two likelihood-equivalent forms of a multivariate mediation model with time-varying measures. (A) A latent change score model with lagged and change score predictors of x, y, and z, and (B) an autoregressive model with lagged and contemporaneous variables. Parameters capturing the same relationship share notation across models (e.g., $\theta_1 - \theta_{12}$), and show the equivalent relationships that we have outlined. The proportional paths in the latent change score model (A; β_x, β_y, and β_z) and the autoregressive paths in the AR model (B; AR_x, AR_y, AR_z) are related by the equation $\beta = AR - 1$. Variances/residuals and intercepts/means are omitted from the diagram for visual clarity. Note that while some paths are curved to avoid overlapping with variables, all paths are single-headed regression paths. AR = autoregressive.

Table 6
Parameter Estimates From Latent Change and Autoregressive Mediation Models

Parameter	Latent change score and lagged mediation model	Contemporaneous and lagged mediation model
$\beta_x: x_1 \rightarrow \Delta x_{21}$	−0.445*** (0.031)	
$AR_x: x_1 \rightarrow x_2$	1.000*** (0.000)	0.555*** (0.031)
$\beta_y: y_2 \rightarrow \Delta y_{32}$	−0.648*** (0.034)	
$AR_y: y_2 \rightarrow y_3$	1.000*** (0.000)	0.352*** (0.034)
$\beta_z: z_3 \rightarrow \Delta z_{43}$	−0.312*** (0.029)	
$AR_z: z_3 \rightarrow z_4$	1.000*** (0.000)	0.688*** (0.029)
$\theta_1: x_1 \rightarrow y_2$	0.096*** (0.021)	0.037 (0.022)
$\theta_2: \Delta x_{21} \vee x_2 \rightarrow y_2$	0.060** (0.023)	0.060** (0.023)
$\theta_3: x_1 \rightarrow \Delta y_{32} \vee y_3$	0.003 (0.019)	−0.005 (0.020)
$\theta_4: \Delta x_{21} \vee x_2 \rightarrow \Delta y_{32} \vee y_3$	0.008 (0.020)	0.008 (0.020)
$\theta_5: x_1 \rightarrow z_3$	−0.060*** (0.018)	−0.018 (0.019)
$\theta_6: \Delta x_{21} \vee x_2 \rightarrow z_3$	−0.042* (0.019)	−0.042* (0.019)
$\theta_7: y_2 \rightarrow z_3$	0.190*** (0.039)	−0.023 (0.034)
$\theta_8: \Delta y_{32} \vee y_3 \rightarrow z_3$	0.212*** (0.035)	0.212*** (0.035)
$\theta_9: x_1 \rightarrow \Delta z_{43} \vee z_4$	−0.009 (0.014)	−0.026+ (0.014)
$\theta_{10}: \Delta x_{21} \vee x_2 \rightarrow \Delta z_{43} \vee z_4$	0.016 (0.014)	0.016 (0.014)
$\theta_{11}: y_2 \rightarrow \Delta z_{43} \vee z_4$	0.039 (0.030)	−0.028 (0.026)
$\theta_{12}: \Delta y_{32} \vee y_3 \rightarrow \Delta z_{43} \vee z_4$	0.067* (0.027)	0.067* (0.027)
-2ℓ	13,492.1	13,492.1

Note. -2ℓ is the −2 log-likelihood. AR = autoregressive.
* $p < .05$. ** $p < .01$. *** $p < .001$.

is the difference between two parameters from the LCS version. For instance, $\theta_{1,AR}$ from the AR model is the difference between $\theta_{1,LCS}$ and $\theta_{2,LCS}$ from the LCS model (Table 6; small inconsistencies are due to rounding). This reflects the $b = e - f$ equivalency. Each pair of parameters (e.g., θ_7 and θ_8, θ_{11} and θ_{12}) recapitulates these equivalencies, demonstrating how the equations we derived in the first model-free derivations radiate throughout the multivariate system. Furthermore, like before, these models have identical fit to the data ($-2\ell = 13,492.1$; Table 6), highlighting their equivalence further. Despite the greater complexity of these models, our derivations continue to allow us insight into the relationships between the two versions of the model.

While the direct effects are interesting for us in terms of extending the equivalencies outlined in univariate models to their multivariate counterparts, of substantive interest for most researchers would be the indirect effects that can be estimated within the mediation models. In Table 7, we outline the estimates for the indirect effects constructed by multiplying pairs of parameters from Table 6. Because we did not have access to the raw data, we could not generate bootstrapped confidence intervals through resampling, so confidence intervals were estimated using the traditional Delta (or "Sobel") method (Sobel, 1982).

We can begin by examining the mediation pathways which include the proportional and AR pathways. Similar to what we have seen before, the indirect effect estimates are inverted due to the $\beta = AR - 1$ relationship in the direct effects. The corresponding indirect effects that include β_x and AR_x, while opposite in sign, are relatively similar in magnitude and significance (e.g., $x_1 \rightarrow \Delta x_{21} \rightarrow z_3 = 0.019$, $p = .028$; $x_1 \rightarrow x_2 \rightarrow z_3 =$ -0.023, $p = .028$) because the $AR_x \approx 0.5$. The β_y and AR_y estimates, by contrast, are more unbalanced ($\beta_y = -0.648$ vs. $AR_y = 0.352$), which leads to larger differences in the magnitude of the resulting indirect effects (e.g., $y_2 \rightarrow \Delta y_{32} \rightarrow \Delta z_{43} = -0.043$, $p = .015$; $y_2 \rightarrow y_3 \rightarrow z_4 = 0.024$, $p = .017$). Thus, as the AR effect weakens, we can expect that these coefficients will further diverge.

The indirect effects built from our equivalent parameters, by contrast, show greater disparities across model types. Because of the $b = e - f$ equivalency, the coefficients corresponding to the e weight in the LCS model ($\theta_{1,LCS}$, $\theta_{3,LCS}$, etc.) have a tendency to be larger than the parameters corresponding to the b weight in the AR model ($\theta_{1,AR}$, $\theta_{3,AR}$, etc.) which can be seen by re-arranging the above expression to $e = b + f$. When b and f have similar signs—which is likely in these models—this should lead to larger e estimates, which in turn will inflate the indirect effect involving these pathways. This manifests in the stronger indirect effects in the LCS model compared with the AR model, as given in Table 7. However, note that this inflation in the indirect effect is not due to changes in a path that involves the Δy or contemporaneous y predictors which distinguish the two models. Rather it is due to the lagged $y_2 \rightarrow z_3$ (θ_7) relationship, which involves identical variables across the two versions of the model. This somewhat unintuitive change highlights the caution we need to exercise when transitioning between versions of the model. Because we are controlling for different time-varying predictors, the relationships among the same variables can shift out from under us. These models are likelihood equivalent thus there is no difference in empirical fit that might motivate one version over another—and we emphasize that neither is wrong. We will simply have use other considerations besides fit to select the theoretically optimal version of our time-varying measures. However, as we will discuss in the following section, it seems that the AR model might be a useful default approach to avoid the inversions in sign that are commonly encountered in the change score model.

The issues highlighted here extend to a broad class of models where change versus lagged effects might be of interest. With additional repeated measures, the LCS model can be extended to include a growth component and LCSs from early time intervals can be used to predict change in future time intervals—that is, change scores are both predictors and outcomes—to capture additional dynamics (Estrada et al., 2019; Grimm et al., 2012). The AR model can

Table 7
Indirect Effect Estimates From Latant Change and Autoregressive Mediation Models

Parameter	Latent change score and lagged mediation model	Contemporaneous and lagged mediation model
$x_1 \beta_x \Delta x_{21} \theta_2 y_2$	−0.027** (0.010)	
$x_1 AR_x x_2 \theta_2 y_2$		0.033** (0.013)
$x_1 \beta_x \Delta x_{21} \theta_4 \Delta y_{32}$	−0.004 (0.009)	
$x_1 AR_x x_2 \theta_4 y_3$		0.005 (0.011)
$x_1 \beta_x \Delta x_{21} \theta_6 z_3$	0.019* (0.009)	
$x_1 AR_x x_2 \theta_6 z_3$		−0.023* (0.011)
$x_1 \beta_x \Delta x_{21} \theta_{10} \Delta z_{43}$	−0.007 (0.006)	
$x_1 AR_x x_2 \theta_{10} z_4$		0.009 (0.008)
$y_2 \beta_y \Delta y_{32} \theta_8 z_3$	−0.137*** (0.024)	
$y_2 AR_y y_3 \theta_8 z_3$		0.075*** (0.014)
$y_2 \beta_y \Delta y_{32} \theta_{12} \Delta z_{43}$	−0.043* (0.018)	
$y_2 AR_y y_3 \theta_{12} z_4$		0.024* (0.010)
$x_1 \theta_1 y_2 \theta_7 z_3$	0.018*** (0.005)	−0.001 (0.001)
$x_1 \theta_1 y_2 \theta_{11} \Delta z_{43} \vee z_4$	0.004 (0.003)	−0.001 (0.001)
$\Delta x_{21} \vee x_2 \theta_2 y_2 \theta_7 z_3$	0.011* (0.005)	−0.001 (0.002)
$\Delta x_{21} \vee x_2 \theta_2 y_2 \theta_{11} \Delta z_{43} \vee z_4$	0.002 (0.002)	−0.002 (0.002)
$x_1 \theta_3 \Delta y_{32} \vee y_3 \theta_8 z_3$	0.001 (0.004)	−0.001 (0.004)
$x_1 \theta_3 \Delta y_{32} \vee y_3 \theta_{12} \Delta z_{43} \vee z_4$	0.000 (0.001)	0.000 (0.001)
$\Delta x_{21} \vee x_2 \theta_4 \Delta y_{32} \vee y_3 \theta_8 z_3$	0.002 (0.004)	0.002 (0.004)
$\Delta x_{21} \vee x_2 \theta_4 \Delta y_{32} \vee y_3 \theta_{12} \Delta z_{43} \vee z_4$	0.001 (0.001)	0.001 (0.001)
-2ℓ	13,492.1	13,492.1

Note. -2ℓ is the −2 log-likelihood. AR = autoregressive.
* $p < .05$. ** $p < .01$. *** $p < .001$.

likewise be extended in a myriad of ways, with lagged relationships being a key feature of various forms of cross-lagged panel models (Hamaker et al., 2015; Usami et al., 2019), AR latent trajectory models (Bollen & Curran, 2004), latent curve models with structured residuals (Curran et al., 2014), and dynamic structural equation models (Asparouhov et al., 2018). While these extensions grow increasingly complex, the principles we have outlined here extend naturally into these models, suggesting that there multiple equivalent ways to capture the same dynamics that might give superficially different (and inverted) inferences.

Recommendations for Applied Research

Staring with simple arithmetic expressions all the way up to the complexity of multivariate longitudinal models, we have seen how the apparent differences between using contemporaneous (x_t), lagged (x_{t-1}), and change (Δx) forms of a TVC belie underlying equivalencies. In particular, we showed that all forms of the model contain the same information, and simply package the predictive effect in different ways depending on which form of the covariate used. While useful for understanding the models themselves, these equivalencies might leave applied researchers unsure how to proceed—as all models fit precisely equally, and we can get all alternative model results from any given exemplar.

From one perspective, these equivalencies can be used to justify any of the approaches outlined here, similar to the residualized versus change score as outcomes debate (Castro-Schilo & Grimm, 2018). As such, if researchers have strong theoretical reasons to frame the effect of the TVC in terms of change scores, then they can proceed without compromising the ability to explain variation in the outcome. Additionally, change scores can be more intuitive compared with residualized change models when interpreting these prospective associations (Willett, 1997). However, by the same logic, neither should researchers privilege the inferences of the change predictor as being theoretically distinct from using the contemporaneous and lagged versions of the TVC, they are merely transformations of one another.

From another perspective, our results suggest several reasons why the form of the model with the contemporaneous and lagged forms of the TVC Equation 4—and not the change score—would be useful as a default approach. First, x_t and x_{t-1} are variables that we directly measure, while Δx is a derived composite—whether computed as a data step or modeled directly as a latent difference. However, unlike some composites, like product terms used to estimate interactions, purely additive composites (like difference scores) cannot explain additional variance net their constituent parts. This means that we could not include x_t, x_{t-1}, and Δx in the same model and still obtain unique estimates (McCormick et al., 2022), which we can do in the case of product composites. As we saw in Table 2, we also cannot avoid this by only including the Δx predictor because of the highly unlikely constraint that places on the model—where it is equivalent to x_t and x_{t-1} having regression coefficients of equal magnitude but opposite sign. Additionally, due to the composite nature of Δx, its effect can be completely driven by the contemporaneous x_t rather than any form of prospective relationship, which is the putative aim of including this form of the TVC.

Another issue becomes apparent in the multivariate mediation models we discussed, which is the relationship between the AR and proportional parameters, as given in Table 6. What is of potential concern is that as the autoregressive effect tends toward zero, the proportional effect tends toward −1. In other words, when the observed repeated measures are completely unrelated over time (i.e., zero autocorrelation), there becomes a deterministic inverse relationship between the lagged version of the TVC and the change score which introduces possibilities for misinterpretation. That is because while the proportional parameter in the LCS model is most often interpreted as how prior status predicts subsequent change, it more accurately reflects the strength of the AR effect minus the perfect prediction of x_{t-1} on itself within the change score (e.g., $x_{t-1} \rightarrow \Delta x = (x_t - x_{t-1})$. While we stress that this is not wrong, per se, and equivalently describes the data, this composite nature of Δx leads to a higher likelihood for misinterpretation of the time-varying relationships for these reasons.

Summary and Conclusions

Here, we extended a long history of concern for the use of residualized versus difference scores in the study of change from its traditional focus on outcomes to instead examine predictors. We showed a general derivation for how the effects of contemporaneous, lagged, and difference score versions of a given TVC relate to one another, with relationships for transforming between these parameters. We then demonstrated how these relationships impact estimates and interpretations in applications of TVCs within the multiple regression model and the multilevel model using empirical data to highlight the inferential challenges related to the choice of TVC model. We showed that these parameter transformations hold across a range of ancillary modeling decisions, including whether to control for baseline status in the outcome and the inclusion of random effects. Finally, we synthesized past and current research to highlight how the use of change scores as both outcome and predictor within a mediation model can alter the estimation of indirect effects using the LCS model, and urged caution regarding the use of change scores in these analyses. These results offer a nice symmetry of considering long-standing issues of change in both outcomes and covariates, and shrink the conceptual distance between considerations for the TVC versus multivariate approaches for modeling change.

References

Arizmendi, C., Gates, K., Fredrickson, B., & Wright, A. (2021). Specifying exogeneity and bilinear effects in data-driven model searches. *Behavior Research Methods*, *53*(3), 1276–1288. https://doi.org/10.3758/s13428-020-01469-2

Asparouhov, T., Hamaker, E. L., & Muthén, B. (2018). Dynamic structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, *25*(3), 359–388. https://doi.org/10.1080/10705511.2017.1406803

Bereiter, C. (1963). Some persisting dilemmas in the measurement of change. In C. W. Harris (Ed.), *Problems in measuring change* (pp. 3–20). University of Wisconsin Press.

Bollen, K. A., & Curran, P. J. (2004). Autoregressive latent trajectory (ALT) models a synthesis of two traditions. *Sociological Methods & Research*, *32*(3), 336–383. https://doi.org/10.1177/0049124103260222

Casey, B. J., Cannonier, T., Conley, M. I., Cohen, A. O., Barch, D. M., Heitzeg, M. M., Soules, M. E., Teslovich, T., Dellarco, D. V., Garavan, H., Orr, C. A., Wager, T. D., Banich, M. T., Speer, N. K., Sutherland, M. T., Riedel, M. C., Dick, A. S., Bjork, J. M., Thomas, K. M., … Dale, A. M. (2018). The adolescent brain cognitive development (ABCD) study: Imaging acquisition across 21 sites. *Developmental Cognitive Neuroscience*, *32*, 43–54. https://doi.org/10.1016/j.dcn.2018.03.001

Castro-Schilo, L., & Grimm, K. J. (2018). Using residualized change vs. difference scores for longitudinal research. *Journal of Social and Personal Relationships*, *35*(1), 32–58. https://doi.org/10.1177/0265407517718387

Cronbach, L. J., & Furby, L. (1970). How we should measure change: Or should we? *Psychological Bulletin*, *74*(1), 68–80. https://doi.org/10.1037/h0029382

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*(1), 583–619. https://doi.org/10.1146/psych.2011.62.issue-1

Curran, P. J., & Hancock, G. R. (2021). The challenge of modeling co-developmental processes over time. *Child Development Perspectives*, *15*(2), 67–75. https://doi.org/10.1111/cdep.v15.2

Curran, P. J., Howard, A. L., Bainter, S. A., Lane, S. T., & McGinley, J. S. (2014). The separation of between-person and within-person components of individual change over time: A latent curve model with structured residuals. *Journal of Consulting and Clinical Psychology*, *82*(5), 879–894. https://doi.org/10.1037/a0035297

Curran, P. J., Obeidat, K., & Losardo, D. (2010). Twelve frequently asked questions about growth curve modeling. *Journal of Cognition and Development*, *11*(2), 121–136. https://doi.org/10.1080/15248371003699969

Epskamp, S., Waldorp, L. J., Mõttus, R., & Borsboom, D. (2018). The Gaussian graphical model in cross-sectional and time-series data. *Multivariate Behavioral Research*, *53*(4), 453–480. https://doi.org/10.1080/00273171.2018.1454823

Estrada, E., Ferrer, E., Román, F. J., Karama, S., & Colom, R. (2019). Time-lagged associations between cognitive and cortical development from childhood to early adulthood. *Developmental Psychology*, *55*(6), 1338–1352. https://doi.org/10.1037/dev0000716

Fried, E. I., Papanikolaou, F., & Epskamp, S. (2022). Mental health and social contact during the COVID-19 pandemic: An ecological momentary assessment study. *Clinical Psychological Science*, *10*(2), 340–354. https://doi.org/10.1177/21677026211017839

Goldsmith, K. A., MacKinnon, D. P., Chalder, T., White, P. D., Sharpe, M., & Pickles, A. (2018). Tutorial: The practical application of longitudinal structural equation mediation models in clinical trials. *Psychological Methods*, *23*(2), 191–207. https://doi.org/10.1037/met0000154

Gottfredson, N. C. (2019). A straightforward approach for coping with unreliability of person means when parsing within-person and between-person effects in longitudinal studies. *Addictive Behaviors*, *94*, 156–161. https://doi.org/10.1016/j.addbeh.2018.09.031

Grimm, K. J., An, Y., McArdle, J. J., Zonderman, A. B., & Resnick, S. M. (2012). Recent changes leading to subsequent changes: Extensions of multivariate latent difference score models. *Structural Equation Modeling: A Multidisciplinary Journal*, *19*(2), 268–292. https://doi.org/10.1080/10705511.2012.659627

Grimm, K. J., Zhang, Z., Hamagami, F., & Mazzocco, M. (2013). Modeling nonlinear change via latent change and latent acceleration frameworks: Examining velocity and acceleration of growth trajectories. *Multivariate Behavioral Research*, *48*(1), 117–143. https://doi.org/10.1080/00273171.2012.755111

Hamaker, E. L., Kuiper, R. M., & Grasman, R. P. P. P. (2015). A critique of the cross-lagged panel model. *Psychological Methods*, *20*(1), 102–116. https://doi.org/10.1037/a0038889

Hoffman, L., & Stawski, R. S. (2009). Persons as contexts: Evaluating between-person and within-person effects in longitudinal analysis. *Research in Human Development*, *6*(2-3), 97–120. https://doi.org/10.1080/15427600902911189

Howard, A. L., Patrick, M. E., & Maggs, J. L. (2015). College student affect and heavy drinking: Variable associations across days, semesters, and people. *Psychology of Addictive Behaviors: Journal of the Society of Psychologists in Addictive Behaviors*, *29*(2), 430–443. https://doi.org/10.1037/adb0000023

Kievit, R. A., Brandmaier, A. M., Ziegler, G., van Harmelen, A.-L., de Mooij, S. M. M., Moutoussis, M., Goodyer, I. M., Bullmore, E., Jones, P. B., Fonagy, P., Lindenberger, U., & Dolan, R. J. (2018). Developmental cognitive neuroscience using latent change score models: A tutorial and applications. *Developmental Cognitive Neuroscience*, *33*, 99–117. https://doi.org/10.1016/j.dcn.2017.11.007

Kollen, B., van de Port, I., Lindeman, E., Twisk, J., & Kwakkel, G. (2005). Predicting improvement in gait after stroke. *Stroke*, *36*(12), 2676–2680. https://doi.org/10.1161/01.STR.0000190839.29234.50

Lord, F. M. (1956). The measurement of growth. *ETS Research Bulletin Series*, *1956*(1), i–22. https://doi.org/10.1002/j.2333-8504.1956.tb00058.x

Luciana, M., Bjork, J., Nagel, B., Barch, D., Gonzalez, R., Nixon, S., & Banich, M. (2018). Adolescent neurocognitive development and impacts of substance use: Overview of the adolescent brain cognitive development (ABCD) baseline neurocognition battery. *Developmental Cognitive Neuroscience*, *32*, 67–79. https://doi.org/10.1016/j.dcn.2018.02.006

Maheux, A. J., Nesi, J., Galla, B. M., Roberts, S. R., & Choukas-Bradley, S. (2021). #grateful: Longitudinal associations between adolescents' social media use and gratitude during the COVID-19 pandemic. *Journal of Research on Adolescence*, *31*(3), 734–747. https://doi.org/10.1111/jora.v31.3

McArdle, J. J. (2009). Latent variable modeling of differences and changes with longitudinal data. *Annual Review of Psychology*, *60*(1), 577–605. https://doi.org/10.1146/psych.2009.60.issue-1

McArdle, J. J., & Nesselroade, J. R. (1994). Using multivariate data to structure developmental change. In S. H. Cohen & H. W. Reese (Eds.), *Life-span developmental psychology: Methodological contributions* (pp. 223–267). Lawrence Erlbaum Associates.

McCormick, E. M. (2021). Multi-level multi-growth models: New opportunities for addressing developmental theory using advanced longitudinal designs with planned missingness. *Developmental Cognitive Neuroscience*, *51*, Article 101001. https://doi.org/10.1016/j.dcn.2021.101001

McCormick, E. M. (2023a). *Deriving models of change with interpretable parameters: Linear estimation with nonlinear inference*. PsyArXiv. https://doi.org/10.31234/osf.io/r4vxb

McCormick, E. M. (2023b). *How should we model the effect of "change"—or should we?* https://osf.io/yc96v/

McCormick, E. M., & Bauer, D. J. (2023). *How should we model the effect of "change"—or should we?* PsyArXiv. https://doi.org/10.31234/osf.io/pbhmq

McCormick, E. M., Borgeest, G. S., & Kievit, R. (2022). *Interrupted mediation: A cautionary note on using derived metrics as intervening variables in path models*. PsyArXiv. https://doi.org/10.31234/osf.io/48xj5

McCormick, E. M., Byrne, M. L., Flournoy, J. C., Mills, K. L., & Pfeifer, J. H. (2023). The Hitchhiker's guide to longitudinal models: A primer on model selection for repeated-measures methods. *Developmental Cognitive Neuroscience*, *63*, Article 101281. https://doi.org/10.1016/j.dcn.2023.101281

McNeish, D., & Matta, T. (2018). Differentiating between mixed-effects and latent-curve approaches to growth modeling. *Behavior Research Methods*, *50*(4), 1398–1414. https://doi.org/10.3758/s13428-017-0976-5

McNeish, D., & Matta, T. H. (2020). Flexible treatment of time-varying covariates with time unstructured data. *Structural Equation Modeling: A Multidisciplinary Journal*, *27*(2), 298–317. https://doi.org/10.1080/10705511.2019.1627213

Michel, L. C., McCormick, E. M., & Kievit, R. A. (2023). *Grey and white matter metrics demonstrate distinct and complementary prediction of differences in cognitive performance in children: Findings from ABCD (n = 11,876)*. bioRxiv. https://doi.org/10.31234/osf.io/r4vxb

Nakagawa, S., Johnson, P. C. D., & Schielzeth, H. (2017). The coefficient of determination R2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *Journal of The Royal Society Interface*, *14*(134), Article 20170213. https://doi.org/10.1098/rsif.2017.0213

O'Laughlin, K. D., Martin, M. J., & Ferrer, E. (2018). Cross-sectional analysis of longitudinal mediation processes. *Multivariate Behavioral Research*, *53*(3), 375–402. https://doi.org/10.1080/00273171.2018.1454822

Rogosa, D. R., & Willett, J. B. (1983). Demonstrating the reliability of the difference score in the measurement of change. *Journal of Educational Measurement*, *20*(4), 335–343. https://doi.org/10.1111/jedm.1983.20.issue-4

Rogosa, D. R., & Willett, J. B. (1985). Understanding correlates of change by modeling individual differences in growth. *Psychometrika*, *50*(2), 203–228. https://doi.org/10.1007/BF02294247

Selig, J. P., & Preacher, K. J. (2009). Mediation models for longitudinal data in developmental research. *Research in Human Development*, *6*(2), 144–164. https://doi.org/10.1080/15427600902911247

Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. *Sociological Methodology*, *13*, 290–312. https://doi.org/10.2307/270723

Usami, S., Murayama, K., & Hamaker, E. L. (2019). A unified framework of longitudinal models to examine reciprocal relations. *Psychological Methods*, *24*(5), 637–657. https://doi.org/10.1037/met0000210

Valente, M. J., Georgeson, A. R., & Gonzalez, O. (2021). Clarifying the implicit assumptions of two-wave mediation models via the latent change score specification: An evaluation of model fit indices. *Frontiers in Psychology*, *12*, Article 709198. https://doi.org/10.3389/fpsyg.2021.709198

Valente, M. J., & MacKinnon, D. P. (2017). Comparing models of change to estimate the mediated effect in the pretest–posttest control group design. *Structural Equation Modeling: A Multidisciplinary Journal*, *24*(3), 428–450. https://doi.org/10.1080/10705511.2016.1274657

Wang, L. P., & Maxwell, S. E. (2015). On disaggregating between-person and within-person effects with longitudinal data using multilevel models. *Psychological Methods*, *20*(1), 63–83. https://doi.org/10.1037/met0000030

Willett, J. B. (1997). Measuring change: What individual growth modeling buys you. In E. Amsel & K. Renninger (Eds.), *Measuring change: What individual growth modeling buys you* (pp. 213–241). Lawrence Erlbaum Associates. https://gseacademic.harvard.edu/~willetjo/pdf%20files/Willett%201997.pdf

Wright, A. G. C., & Simms, L. J. (2016). Stability and fluctuation of personality disorder features in daily life. *Journal of Abnormal Psychology*, *125*(5), 641–656. https://doi.org/10.1037/abn0000169

Appendix

Full Covariance Matrix Transformations

To obtain the relationships between all variance and covariance parameters between the two models simultaneously, we can take a matrix-based approach (see McCormick, 2023a for details on this approach), where we pre and postmultiply the Jacobian matrix of partial derivatives of the fixed effects transformations (i.e., Equation 20) with respect to the parameters of the reference model. For instance, to obtain the covariance matrices (**T**) for Equation 17 (model with $x_{i,t}$ and Δx_i) and Equation 18 (model with $x_{i,t-1}$ and Δx_i) from Equation 16 (model with $x_{i,t}$ and $x_{i,t-1}$), we would compute the following expressions:

$$\begin{aligned} \mathbf{T}_{(c,d)} &= \mathbf{J}'_{(c,d)}\ \mathbf{T}_{(a,b)}\ \mathbf{J}_{(c,d)}, \\ \mathbf{T}_{(e,f)} &= \mathbf{J}'_{(e,f)}\ \mathbf{T}_{(a,b)}\ \mathbf{J}_{(e,f)}, \end{aligned} \tag{A1}$$

(*Appendix continues*)

where $\mathbf{J}_{(c,d)}$ and $\mathbf{J}_{(e,f)}$ contain partial derivatives of the following form:

$$\mathbf{J}_{(c,d)} = \begin{bmatrix} \frac{\partial(c = a+b)}{\partial a} & \frac{\partial(d = -b)}{\partial a} \\ \frac{\partial(c = a+b)}{\partial b} & \frac{\partial(d = -b)}{\partial b} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix},$$
$$\mathbf{J}_{(e,f)} = \begin{bmatrix} \frac{\partial(e = a+b)}{\partial a} & \frac{\partial(f = a)}{\partial a} \\ \frac{\partial(e = a+b)}{\partial b} & \frac{\partial(f = a)}{\partial b} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \tag{A2}$$

and apply the necessary quadratic transformations to both the variances on the diagonal and covariances (or correlations if we standardize $\mathbf{T}$) on the off-diagonal. Note that the -1 in $\mathbf{J}_{(c,d)}$ Equation A2 will result in the random effect correlation of $x_{i,t}$ and Δx_i being opposite in sign to the correlation in the other alternative TVC models.

Incidentally, as outlined by McCormick (2023a), this matrix-based approach could alternatively be used to compute the standard errors by pre and postmultiplying the asymptotic covariance matrix of the fixed effects—ACOV(γ)—by the Jacobian instead, with the form:

$$\begin{aligned} \text{ACOV}(\gamma)_{(c,d)} &= \mathbf{J}'_{(c,d)}\ \text{ACOV}(\gamma)_{(a,b)}\ \mathbf{J}_{(c,d)}, \\ \text{ACOV}(\gamma)_{(e,f)} &= \mathbf{J}'_{(e,f)}\ \text{ACOV}(\gamma)_{(a,b)}\ \mathbf{J}_{(e,f)}, \end{aligned} \tag{A3}$$

and taking the square root of the diagonal of the resulting matrix.

Received August 4, 2023
Revision received January 4, 2024
Accepted March 25, 2024 ■