DOI: 10.1002/cad.20387

RESEARCH ARTICLE

WILEY

Effect size measures for longitudinal growth analyses: Extending a framework of multilevel model R-squareds to accommodate heteroscedasticity, autocorrelation, nonlinearity, and alternative centering strategies

Jason D. Rights[1] | **Sonya K. Sterba**[2]

[1] Department of Psychology, University of British Columbia, Vancouver, British Columbia, Canada

[2] Department of Psychology and Human Development, Vanderbilt University, Nashville, Tennessee

Correspondence
Jason D. Rights, Department of Psychology, 2136 West Mall, University of British Columbia, Vancouver, BC V6T1Z4, Canada.
Email: jrights@psych.ubc.ca

Abstract

Developmental researchers commonly utilize multilevel models (MLMs) to describe and predict individual differences in change over time. In such growth model applications, researchers have been widely encouraged to supplement reporting of statistical significance with measures of effect size, such as R-squareds (R^2) that convey variance explained by terms in the model. An integrative framework for computing R-squareds in MLMs with random intercepts and/or slopes was recently introduced by Rights and Sterba and it subsumed pre-existing MLM R-squareds as special cases. However, this work focused on cross-sectional applications, and hence did not address how the computation and interpretation of MLM R-squareds are affected by modeling considerations typically arising in longitudinal settings: (a) alternative centering choices for time (e.g., centering-at-a-constant vs. person-mean-centering), (b) nonlinear effects of predictors such as *time*, (c) heteroscedastic level-1 errors and/or (d) autocorrelated level-1 errors. This paper addresses these gaps by extending the Rights and Sterba R-squared framework to longitudinal contexts. We: (a) provide a full framework of total and level-specific R-squared measures for MLMs that utilize any type of centering, and contrast these with Rights and Sterba's measures assuming cluster-mean-centering, (b) explain and derive which measures are applicable for MLMs with nonlinear terms, and extend the R-squared computation to accommodate (c) heteroscedastic and/or (d) auto-correlated errors. Additionally, we show how to use differences in R-squared (ΔR^2) measures between growth models (adding, for instance, time-varying covariates as level-1 predictors or time-invariant covariates as level-2 predictors) to obtain effects sizes for individual terms. We provide R software (*r2MLMlong*) and a running pedagogical example analyzing growth in adolescent self-efficacy to illustrate these methodological developments. With these developments, researchers will have greater ability to consider effect size when analyzing and predicting change using MLMs.

KEYWORDS

effect size, longitudinal analyses, mixed effects modeling, multilevel modeling, R-squared

Researchers studying child and adolescent development commonly analyze repeated measures data using multilevel models (MLMs) in order to describe and predict individual differences in change over time (see, e.g., Hoffman, 2015; Raudenbush & Bryk, 2002; Singer & Willett, 2003). For instance, developmental researchers use MLMs (also called hierarchical linear models or linear mixed effects models) to assess longitudinal changes in newborns' sleep duration (Hoeksma & Koomen, 1992), children's language skills (Siller & Sigman, 2008), adolescents' cortical thickness (King et al., 2018), and young adults' risky behavior (Fergus, Zimmerman, & Caldwell, 2007). Multilevel modeling has proven a useful framework for such applications in that it readily accommodates the inherent dependency of individual observations nested within persons by allowing regression coefficients to vary across persons.

In such longitudinal growth applications of MLMs, researchers have been widely encouraged to compute measures of effect size to convey practical or clinical significance of results (e.g., APA, 2009; Hoffman, 2015; Kwok et al., 2008; LaHuis, Blackmore, & Bryant-Lees, 2019; Lorah, 2018; Nezlek, 2012). One such effect size measure ubiquitously reported in *single-level* regression analyses is R-squared, indicating the proportion of variance explained by a given model. In contrast, for MLM applications, the reporting of R-squared measures has been historically complicated by the fact that many R-squared measures for MLMs were separately developed and there was little understanding of how to relate, interpret, or choose among them (LaHuis, Hartman, Hakoyama, & Clark, 2014). To address these issues, Rights and Sterba (2019) recently developed an integrative framework of R-squared measures for MLMs that analytically related pre-existing measures (from, e.g., Aguinis & Culpepper, 2015; Hox, 2010; Johnson, 2014; Kreft & de Leeuw, 1998; Nakagawa & Schielzeth, 2013; Raudenbush & Bryk, 2002; Snijders & Bosker, 2012; Vonesh & Chinchilli, 1997; Xu, 2003) as special cases, supplied new measures to fill gaps, and provided a unified, accessible approach for visualizing and interpreting these measures.

However, Rights and Sterba (2019) focused primarily on *cross-sectional* multilevel applications, and their R-squared framework did not fully accommodate the following five modeling features that developmental researchers commonly utilize in *longitudinal* applications:

1. *Utilizing alternative centering strategies.* Researchers fitting growth models are often interested in interpreting coefficients after centering *time* at a constant value such as the first or last occasion (e.g., Biesanz, Deeb-Sossa, Papadakis, Bollen, & Curran, 2004), rather than cluster-mean-centering *time*. However, Rights and Sterba's (2019) full framework of measures was provided under the assumption of cluster-mean-centering (called person-mean-centering in longitudinal contexts). They provided only a limited subset of their framework's R-squared measures (some[1] of their *total* measures and none of their *level-specific* measures) for other centering strategies (e.g., centering-at-a-constant) that are more common in longitudinal analyses.
2. *Specifying a nonlinear functional form of time.* Developmental theory frequently leads researchers to specify growth models with nonlinear functions of *time* (Bollen & Curran, 2006; Hoffman, 2015; Singer & Willett, 2003). However, Rights and Sterba's (2019) examples involved only linear effects of predictors. The inclusion of nonlinear terms has implications for which R-squared measures are available and how to interpret them.
3. *Allowing heteroscedastic error variances across timepoints.* In growth model applications there are commonly theoretical reasons to expect level-1 error variances to differ over time (e.g., Crowder & Hand, 1990; Goldstein, 2011). However, Rights and Sterba's (2019) framework (as well as nearly all previous measures subsumed as special cases of their framework) pertained strictly to homoscedastic (i.e., constant across time) error variances.
4. *Incorporating non-diagonal error covariance structures.* In longitudinal settings, level-1 errors may be expected to covary across time, especially when assessments are close in time (e.g., with daily diary data; Bolger & Laurenceau, 2013). However, Rights and Sterba's (2019) framework (as well as prior measures subsumed therein) assumed level-1 errors do not covary across time (i.e., a diagonal error covariance structure).
5. *Including time-varying covariates (i.e., level-1 predictors other than time) and time-invariant covariates (i.e., level-2 predictors).* Once developmental researchers have identified their unconditional model of change, they typically want to investigate the impact of time-varying and/or time-invariant covariates. Rights and Sterba (2020) provided procedures for using R-squared differences (ΔR^2) between models as descriptive effect sizes for individual terms; however, they focused on cross-sectional cluster-mean-centered models, and some interpretations would change under model specifications common in longitudinal analyses, complicating the quantification of effect size for time-varying and time-invariant covariates.

The current paper addresses these gaps by extending the Rights and Sterba (2019, 2020) framework to longitudinal contexts, providing R-squared measures and delineating the implications of R-squared computation for each of the five aforementioned modeling features. We will use a running example modeling self-efficacy across adolescence to first consider alternative centering options, to next consider nonlinear change over time, to then allow error variances to vary over time, to subsequently allow errors to correlate across

[1] For instance, Rights and Sterba's (2019) *total* measures for cluster-mean-centered MLMs distinguished the contribution of level-1 versus level-2 predictors via fixed slopes (called sources f_1 and f_2). However, their total measures for *non*-cluster-mean-centered MLMs did not distinguish between these two sources (which is unideal when substantive interest lies in distinguishing within-person versus between-person sources of explained variance).

timepoints, and, finally, to include both time-varying and time-invariant covariates. For each of these models, we will illustrate how the computation, choice, and interpretation of MLM R-squareds incorporates these modeling features. Mathematical derivations underlying these methodological developments are presented in our appendices so that we can focus the main body of the text on the practical application and interpretation of these R-squared measures in a developmental context. Additionally, to aid researchers in implementing these measures, in the Supporting Information, we provide the example dataset (*selfeffdata.txt*) as well as an R script (*r2MLMlong_supplementalfile.R*) to fit each model and compute all R-squared measures.

1 | BACKGROUND: MULTILEVEL GROWTH MODELS FOR LONGITUDINAL ANALYSES

For illustrative purposes, we will use a running pedagogical simulated example in which we describe and predict adolescents' individual differences in self-efficacy over time. This data set has a hierarchical structure in which, at level-1 (i.e., the observation level), there are repeated observations, and at level-2 (i.e., the person level), there are adolescents. Although data collection was attempted at the same 10 measurement occasions for all adolescents in this example, due to attrition there were only on average 7.5 observations per adolescent. Our level-2 sample size (i.e., number of adolescents) is 400. For now, we are concerned with only two variables in this dataset, namely, our outcome *selfeff* (a numerical rating of self-efficacy) and *time* (given as age in years, i.e., 14, 14.5, 15, 15.5, 16, 16.5, 17, 17.5, 18, 18.5); we will later consider additional predictors that are theoretically anticipated to be related to self-efficacy (GPA, time spent volunteering in the community, and gender; e.g., Johnson, Beebe, Mortimer, & Snyder, 1998; Larose et al., 2008).

We start with an unconditional growth model in which our only predictor is *time*, and in which we assume linear change, as shown in Equation (1).

$$\begin{aligned}
\text{Level 1:}\quad & selfeff_{ij} = \beta_{0j} + \beta_{1j} time_{ij} + e_{ij}, \\
\text{Level 2:}\quad & \beta_{0j} = \gamma_{00} + u_{0j}, \\
& \beta_{1j} = \gamma_{10} + u_{1j}, \\
& e_{ij} \sim N(0, \sigma^2) \\
& \begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix} \sim MVN\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \tau_{00} & \\ \tau_{01} & \tau_{11} \end{bmatrix} \right).
\end{aligned} \tag{1}$$

The i subscript denotes a repeated measure ($i = 1 \ldots I$) where in our running example $I = 10$. The subscript j denotes adolescent ($j = 1 \ldots J$) where in our running example $J = 400$. Hence, $selfeff_{ij}$ represents the ith repeated observation of the outcome for adolescent j. Next we define the person-specific growth coefficients. The person-specific intercept, given as β_{0j}, is adolescent j's expected self-efficacy when $time_{ij} = 0$, and the person-specific slope of $time_{ij}$, given as β_{1j}, is adolescent j's expected change in self-efficacy per unit increase in time. The γ's are the *fixed components* of the person-specific growth coefficients, which reflect across-adolescent averages. The fixed component of the intercept, γ_{00}, is the across-adolescent average expected self-efficacy when $time_{ij} = 0$. The fixed component of the slope, γ_{10}, is the across-adolescent average expected change in self-efficacy for each year increase in $time_{ij}$. The u's, in contrast, are the *random components* of the person-specific growth coefficients, which reflect adolescent-specific deviations from the fixed components. Hence, the random component of the intercept, u_{0j}, is the difference between the adolescent-specific intercept (β_{0j}) and the across-adolescent average intercept (γ_{00}), and the random component of the slope, u_{1j}, is the difference between the adolescent-specific slope (β_{1j}) and the across-adolescent average slope (γ_{10}). The random components are assumed multivariate normally distributed, with the random intercept variance given as τ_{00}, random slope variance as τ_{11}, and intercept-slope covariance as τ_{01}. The level-1 error, e_{ij}, is the difference between an adolescent's expected self-efficacy ($\beta_{0j} + \beta_{1j}time_{ij}$) and their actual self-efficacy ($selfeff_{ij}$), and, for now, it is assumed to be normally distributed with mean 0 and variance σ^2.

2 | BACKGROUND: OVERVIEW OF RIGHTS AND STERBA (2019) MLM R-SQUARED FRAMEWORK

Rights and Sterba (2019) developed an integrative framework for defining and relating R-squared measures for MLMs and it subsumed popular pre-existing MLM R-squared measures (from, e.g., Aguinis & Culpepper, 2015; Hox, 2010; Johnson, 2014; Kreft & de Leeuw, 1998; Nakagawa & Schielzeth, 2013; Raudenbush & Bryk, 2002; Snijders & Bosker, 2012; Vonesh & Chinchilli, 1997; Xu, 2003) as special cases, as well as provided novel ways of quantifying explained variance. We briefly review their framework here as a foundation for subsequent methodological developments. Recall that generically an R-squared can be defined as the ratio of the explained portion of the variance to the overall outcome variance:

$$R^2 = \frac{\text{explained variance}}{\text{outcome variance}}. \tag{2}$$

This yields an effect size indicating the proportion of the outcome variance that is explained by the model. Further defining a particular R-squared measure then requires one to consider (a) which sources of variation are considered to give rise to *explained variance* (which defines the numerator) and (b) what *outcome variance* is of interest (which defines the denominator). Rights and Sterba (2019) first delineated how to address these considerations for cluster-mean-centered models, which, in the present longitudinal context with persons as clusters, amounts to *person*-mean-centered models. Person-mean centering implies that, for each level-1 predictor (e.g., *time* in Equation 1), the person's mean value is subtracted from the raw value. Such person-mean centering assures that the predictor explains only within-person variance, and that its slope reflects a purely within-person effect (e.g., Curran & Bauer, 2011; Enders & Tofighi, 2007; Hoffman, 2015).

To define the denominator and numerator of MLM R-squared measures, Rights and Sterba (2019) provided a novel decomposition of outcome variance whereby one can quantify the proportion of outcome variance attributable to each of a set of distinct sources:

- level-1 (e.g., occasion-level) predictors via the fixed components of slopes (shorthand: f_1)
- level-1 predictors via random slope variation (shorthand: v)
- level-1 errors
- level-2 (e.g., adolescent-level) predictors via fixed components of slopes (shorthand: f_2)
- level-2 (e.g., adolescent-specific) outcome means via random intercept variation (shorthand: m)

In defining the denominator of R-squared measures in person-mean-centered MLMs, the model-implied outcome variance of interest could be the total variance (i.e., variance attributable to all five above sources), the within-person variance (i.e., variance attributable to the first three sources), or the between-person variance (i.e., variance attributable to the last two sources), leading respectively to *total* measures, *within-person* measures, or *between-person* measures. A *within-person* R-squared measure is the most directly relevant effect size metric if, for example, a researcher modeling growth in self-efficacy were most interested in understanding *within-adolescent fluctuations*, and the extent to which these can be understood through growth trajectories and/or time-varying covariates. In contrast, a *between-person* R-squared measure is the most directly relevant effect size metric if, for example, a researcher modeling growth in self-efficacy were instead most interested in understanding why some adolescents are, on average, more self-efficacious than others. As a third option, a *total R-squared measure* provides a more omnibus effect size metric, quantifying the proportion of both the within- and between-adolescent variance that is explained.

In defining the numerator of a *total* MLM R-squared, variance attributable to f_1, f_2, v, and/or m, can be considered singly or in combination. Variance attributable to f_1 and/or v can be considered for the numerator of a *within-person* R-squared measure, and variance attributable to f_2 or m can be considered for the numerator of a *between-person* R-squared measure. When a single source of explained variance appears in the numerator of an R-squared measure (e.g., f_1) we term this a *single-source measure*, but when sums of multiple sources appear in the numerator (e.g., $f_1 + v$), we term this a *combination-source measure*. For further discussion on the distinction between these types of measures and examples of their application and interpretation, see Rights and Sterba (2019, 2020).

Rather than choosing one particular measure to report, however, Rights and Sterba (2019, 2020) argue that one can instead obtain a more complete understanding by considering the entire breakdown of individual sources to which (total, within, or between) variance is attributable. Such a full decomposition of variance can be easily visualized in a barchart, an example of which is given in Figure 1 for the Equation (1) model fit to the self-efficacy example data set, in which we currently person-mean center *time*.[2] In Figure 1, the shaded segments of the left-most column displays a breakdown of how much each individual source f_1, f_2, v, m contributes to the total outcome variance (with the contribution of level-1 errors shown in white). For instance, here we see that an estimated 8% of the total variability in self-efficacy repeated measures is explained by the linear effect of *time* via its fixed component, indicated by the R-squared measure $\hat{R}_t^{2(f_1)} = .08$—with the *superscript* ("f_1") indicating the source of explained variance, and the *subscript* ("t") indicating the outcome variance under consideration (t = total, w = within-person, and b = between-person).[3] In terms of quantifying the degree of *individual differences* in growth over time, we see that an estimated 11% of the total variance is attributable to the linear effect of *time* via random slope variation, indicated by $\hat{R}_t^{2(v)} = .11$. These results highlight the utility of quantifying the contribution of not only the marginal (or average) trajectory, but also individual differences around it, as both can account for meaningful amounts of outcome variance. We also can use the left-most bar to further quantify evidence of meaningful individual differences, as an estimated 48% of outcome variance is attributable to person-specific differences in self-efficacy via intercept variation, indicated by $\hat{R}_t^{2(m)} = .48$. The shaded segments of the middle bar then shows the breakdown of how much each within-person source f_1 and v contributes to within-person outcome variance (with the

[2] Note that point estimates and standard errors for this fitted model are in Table 1.

[3] Using formulas in Appendix A and estimates in Table 1, this $\hat{R}_t^{2(f_1)}$ is computed as the ratio of the variance attributable to *time* via its fixed component (here given as $1.88^2\text{var}(time)$) to the model-implied total outcome variance (given as $1.88^2\text{var}(time) + 4.70\text{var}(time) + 130.43 + 90.94$). The ratio of these two variances, .08, is reflected in the solid red portion of the leftmost bar in Figure 1.

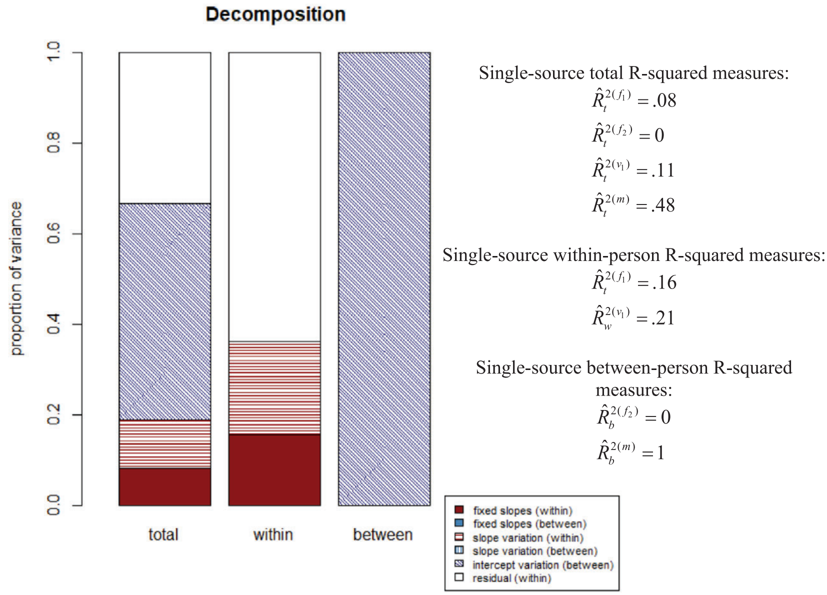

FIGURE 1 Visualizing R-squared results for the unconditional linear growth model of self-efficacy that person-mean centered *time* and assumed a homoscedastic, diagonal error covariance structure (Equation 1)
Notes. The shaded segments of the first bar graphically depict each single-source total R-squared measure. The shaded segments of the second bar graphically depict each single-source within-person R-squared measure. The shaded segments of the third bar depict each single-source between-person R-squared measure. Corresponding numerical values for all of these single-source measures are listed on the far right. The white space in the first and second columns refers to the contribution of level-1 error variance, to either the total or within-person variance. *Combination-source measures* can be visualized as the combination of multiple shaded patterns within a given bar. The shaded patterns in the legend each refer to a different source of variance accounted for. The same shaded pattern appears in multiple bars when the same source is counted as explained variance in total and level-specific measures. Detailed definitions of each measure depicted in Figure 1 were given in Table 3 column 1. Finally, note that for person-mean-centered MLMs (where there is no v_2) the v_1 can simply be referred to as v (as in Rights & Sterba, 2019, 2020).

contribution of level-1 errors again shown in white). Here, the linear effect of *time* accounts for 16% of variance in self-efficacy repeated measures via its fixed component and 21% percent via random slope variation. The shaded segments of the right bar show the breakdown of how much each between-person source (f_2 and m) contributes to between-person variance. Because *time* is presently person-mean-centered, it necessarily has only within-person variability, and as such, we see that no between-person variance is explained by any predictors in this unconditional growth model ($\hat{R}_t^{2(f_2)} = \hat{R}_b^{2(f_2)} = 0$). For the remainder of the current paper, in this fashion, we will focus on interpreting the entire breakdown of individual sources to which variance is attributable, such as that found in Figure 1. Corresponding formulas for each R-squared measure depicted in Figure 1 can be found in Appendix A.

3 | ALTERNATIVE CENTERING STRATEGIES: IMPLICATIONS FOR THE MLM R-SQUARED FRAMEWORK

In the model in Equation (1), the meaning of the adolescent-specific intercept, β_{0j}, and its mean, variance, and intercept-slope covariance, will change depending on how *time* is centered (Biesanz et al., 2004; Mehta & West, 2000). In longitudinal growth applications of

TABLE 1 Point estimates and standard error results from fitting our illustrative self-efficacy growth models

Estimate	Linear growth/person-mean ctr *time* (Equation 1, Figure 1)	Linear growth/grand mean ctr *time* (Equation 1, Figure 2)	Linear growth/ctr time at origin (Equation 1, Figure 2)	Quadratic growth/ctr time at origin (Equation 3, Appendix Figure E1)	Linear growth/ hetero. error/ctr time at origin (Equation. 4, Appendix Figure E2)	Linear growth/ hetero. autoreg. error cov/ctr time at origin (Equation 5, Appendix Figure E3)	Conditional linear growth/ hetero., autoreg. error cov/ctr time origin (Equation. 6, Figure 3)
Fixed component of intercept	27.50 (0.60)*	27.95 (0.60)*	20.67 (0.55)*	21.07 (0.59)*	20.74 (0.55)*	20.68 (0.54)*	20.13 (0.55)*
Fixed component of *time*	1.88 (0.13)*	1.94 (0.14)*	1.94 (0.14)*	1.60 (0.29)*	1.93 (0.14)*	1.96 (0.13)*	0.82 (0.16)*
Fixed component of $time^2$	—	—	—	0.05 (0.03)	—	—	—
Fixed component of GPA	—	—	—	—	—	—	9.85 (0.32)*
Fixed component of volunteer	—	—	—	—	—	—	1.34 (0.21)*
Fixed component of female	—	—	—	—	—	—	1.16 (0.76)
Fixed component of person-mean GPA	—	—	—	—	—	—	1.61 (0.51)*
Fixed component of person-mean volunteer	—	—	—	—	—	—	7.45 (0.35)*
Fixed component female × *time*	—	—	—	—	—	—	2.14 (0.23)*
Random intercept variance	130.43	128.99	84.81	83.27	89.91	62.17	11.80
Random slope variance of *time*	4.70	5.06	5.06	15.32	5.35	3.48	1.93
Intercept, slope-of-*time* covar.	14.58	15.38	−3.58	−8.39	−4.72	0.41	0.37
Random slope of $time^2$	—	—	—	0.15	—	—	—
Random slp. variance of GPA	—	—	—	—	—	—	1.67
Random slp. variance volunteer	—	—	—	—	—	—	9.32

(Continues)

TABLE 1 (Continued)

Estimate	**Linear growth/person-mean ctr *time* (Equation 1, Figure 1)**	**Linear growth/grand mean ctr *time* (Equation 1, Figure 2)**	**Linear growth/ctr time at origin (Equation 1, Figure 2)**	**Quadratic growth/ctr time at origin (Equation 3, Appendix Figure E1)**	**Linear growth/ hetero. error/ctr time at origin (Equation. 4, Appendix Figure E2)**	**Linear growth/ hetero. autoreg. error cov/ctr time at origin (Equation 5, Appendix Figure E3)**	**Conditional linear growth/ hetero., autoreg. error cov/ctr time origin (Equation. 6, Figure 3)**
Intercept, slope-of-*time*^{2} covar.	—	—	—	0.51	—	—	—
Slope-of-*time*, *time*^{2} covar.	—	—	—	−1.24	—	—	—
Intercept, slp.-of-GPA covar.	—	—	—	—	—	—	−3.69
Intercept, slp.-of-volunteer covar.	—	—	—	—	—	—	2.12
Slope of *time*, GPA covar.	—	—	—	—	—	—	0.84
Slope of *time*, volunteer covar.	—	—	—	—	—	—	0.12
Slope of GPA, volunteer covar.	—	—	—	—	—	—	0.32
Autoregressive correlation ($\hat{\rho}$)	—	—	—	—	—	0.32	0.00
Level-1 error variance	90.94	90.31	90.31	82.86	67, 66, 71, 76, 69, 76, 83, 104, 95, 97	84, 87, 90, 94, 88, 94, 103, 134, 118, 109	56, 61, 61, 69, 63, 71, 79, 105, 90, 82

Notes. Models were fit with the *lme* function in the *nlme* package (Pinheiro et al., 2020) in R. Each unit increase in *time* denotes a single increment in assessment period (each 6 months apart).

MLMs, researchers often consider alternative centering choices for level-1 predictors (e.g., *time*). Longitudinal researchers can choose a centering strategy for *time* that best maps onto their substantive interpretational goals (as also recommended by, e.g., Biesanz et al., 2004; Hoffman, 2015). For instance, if a researcher is interested in assessing self-efficacy at baseline (both in terms of its mean value as well as its across-person differences), then they can center time such that *time* = 0 (and hence the intercept) would then indicate baseline. Alternatively, if a researcher is interested in assessing self-efficacy at the *end* of the assessment, they can center time such that *time* = 0 indicates the final assessment. In contrast, if a researcher is interested in assessing self-efficacy at the *mean* timepoint, grand-mean-centering (subtracting from each raw value of time the overall mean of time) can be employed. All of these examples (in addition to leaving time uncentered) are instances of *centering at a constant value* for everyone in the sample.

In contrast, *person-mean-centering* is another strategy that does not necessarily involve subtracting a constant value from every raw value of *time*. If time is *balanced*, meaning that every adolescent has the exact same set of values for *time* and there are no missing outcomes (termed *observed data are balanced*; Sterba, 2014), then there is no between-person variability in *time*, and as such, person-mean-centering *time* is equivalent to grand-mean-centering *time*. However, if time is *unbalanced* (either that people would have had the exact same set of values for *time* if all data had been non-missing—termed *complete data are balanced* in Sterba [2014]—or that people have completely unique measurement occasions—termed *truly individually varying measurements* in Sterba [2014]), there is between-person variability in *time*. Person-mean-centering then will remove the between-person variability, creating a variable that purely varies within-person. In doing so, each person-specific intercept, β_{0j}, can be interpreted as the adolescent's expected self-efficacy at their own personal mean value of time, with γ_{00} reflecting its average across adolescents; researchers desiring this interpretation can choose to person-mean-center.

Here we extend the original R-squared framework of Rights and Sterba (2019), whose focus was providing R-squared measures for MLMs that used person-mean-centering, to fully accommodate alternative centering strategies beyond person-mean-centering (e.g., centering at the first occasion, last occasion, or grand mean, or even leaving level-1 predictors uncentered). This newly enables researchers—using any centering strategy—to consider (a) the proportion of variance explained within-person versus between-person, and (b) the relative impact of predictors at the within-person versus between-person level.[4] As detailed in Table 2 and derived in Appendix B, we developed an expanded decomposition of outcome variance for *non*-person-mean-centered models which distinguishes among variance attributable to:

- the *within-person-varying portion* of level-1 predictors via the fixed components of slopes (f_1)
- the *within-person-varying portion* of level-1 predictors via random slope variation (v_1)
- level-1 errors
- the *between-person-varying portion* of level-1 and/or level-2 predictors via the fixed components of slopes (f_2)
- the *between-person-varying portion* of level-1 predictors via random slope[5] variation (v_2)

[4] For further discussion on the utility of considering within-person and between-person measures and separately considering the impact of predictors within-person versus between-person, see Rights and Sterba (2020) sections titled *Comments on using total versus level-specific R^2 for MLMs* and *Comments on using combination-source versus single-source R^2 for MLMs*.

[5] Though rarely employed in practice except when modeling level-2 heteroscedasticity (not considered here), if a level-2 predictor also had a random slope (see Rights & Sterba, 2016; [blinded], under review; Goldstein, 2011; Snijders & Bosker, 2012), it would also contribute to variance accounted for by v_2.

TABLE 2 Decomposition of model-implied outcome variance for person-mean-centered MLMs versus non-person-mean-centered MLMs

Source to which variance can be attributable	Definition of source when person-mean-centering all level-1 predictors, including *time*	Definition of source when *not* person-mean-centering all level-1 predictors, including *time* (e.g., centering at origin or at grand mean)
f_1	Level-1 predictors via the fixed components of slopes	The within-person-varying portion of level-1 predictors via the fixed components of slopes
f_2	Level-2 predictors via the fixed components of slopes	The between-person-varying portion of level-1 and/or level-2 predictors via the fixed components of slopes
v_1	Level-1 predictors via random slope variation*	The within-person-varying portion of level-1 predictors via random slope variation
v_2	*N/A*	The between-person-varying portion of level-1 predictors[a] via random slope variation
m	Person-specific outcome means via random intercept variation	Person-specific outcome means via random intercept variation at the mean of all predictors with random slopes
level-1 errors†		

Note:
*= For person-mean-centered MLMs (where there is no v_2) the v_1 can simply be referred to as v (as in Rights & Sterba, 2019, 2020).
†= All sources other than the level-1 errors can potentially be considered as numerator terms in constructing MLM R-squared measures.
a= Though rarely employed in practice and not considered here, if a level-2 predictor also had a random slope (see Rights & Sterba, 2016; [blinded], under review; Goldstein, 2011; Snijders & Bosker, 2012 for conceptual explanation), then this text would read "level-1 or level-2."

- person-specific outcome means via random intercept variation *at the mean of all predictors with random slopes* (m)

From this decomposition, we can now create a full set of single-source R-squared measures for *non*-person-mean-centered models that includes total, within-person, and between-person measures. The definitions of the full set of single-source R-squared measures under the assumption of person-mean-centering versus *non*-person-mean-centering is given in Table 3. Specific formulas for each measure in Table 3 Column 2 are given in Appendix A and specific formulas for each measure in Table 3 Column 3 are given in Appendix C.

One way this new decomposition (Column 3 of Table 2) differs from Rights and Sterba's (2019) decomposition for person-mean-centered models (Column 2 of Table 2) is that level-1 predictors can explain both within-person variance (via sources f_1 and v_1) and between-person variance (via sources f_2 and v_2) in non-person-mean-centered models, whereas in person-mean-centered models level-1 predictors can only explain within-person variance (via sources f_1 and v_1). If there is no between-person variability in all level-1 predictors (including, e.g., *time*), however, then f_1 and f_2 will retain their definitions from the person-mean-centered decomposition, and there will be no variance attributable to v_2.

Table 1 shows that a second difference between this new decomposition and the decomposition for person-mean-centered models concerns the definition of source m. Although in random slope MLMs the intercept variance is known to change depending on the constant at which predictors are centered (e.g., Biesanz et al., 2004), Appendix D shows analytically that the proportion of variance attributable to source m does not change. However, as indicated in Table 2, the definition of source m does change—when centering at a constant (rather than centering at the person-mean), variance attributable to person-specific outcome means via random intercept variation is now defined *at the mean of all predic-*

TABLE 3 Definitions of single-source[†] MLM R^2 measures under person-mean versus non-person-mean centering

Single-source[†] R^2 measure	Definition when person-mean-centering all level-1 predictors	Definition when *not* person-mean-centering all level-1 predictors
Total R-squared measures		
$R_t^{2(f_1)}$	Proportion of total outcome variance attributable to level-1 predictors via the fixed components of slopes	Proportion of total outcome variance attributable to the *within-cluster-varying portion* of level-1 predictors via the fixed components of slopes
$R_t^{2(f_2)}$	Proportion of total outcome variance attributable to level-2 predictors via the fixed components of slopes	Proportion of total outcome variance attributable to the *between-cluster-varying portion* of level-1 and/or level-2 predictors via the fixed components of slopes
$R_t^{2(v_1)}$	Proportion of total outcome variance attributable to level-1 predictors via random slope variation*	Proportion of total outcome variance attributable to the *within-cluster-varying portion* of level-1 predictors via the random slope variation
$R_t^{2(v_2)}$	*N/A*	Proportion of total outcome variance attributable to the *between-cluster-varying portion* of level-1 predictors[a] via random slope variation
$R_t^{2(m)}$	Proportion of total outcome variance attributable to person-specific outcome means via random intercept variation	Proportion of total outcome variance attributable to person-specific outcome means via random intercept variation at mean of all predictors with random slopes
Within-person R-squared measures		
$R_w^{2(f_1)}$	Proportion of within-person outcome variance attributable to level-1 predictors via fixed components of slopes	Proportion of within-person outcome variance explained by the within-person-varying portion of level-1 predictors via fixed components of slopes
$R_w^{2(v_1)}$	Proportion of within-person outcome variance attributable to level-1 predictors via random slope variation*	Proportion of within-person outcome variance attributable to the within-person-varying portion of level-1 predictors via random slope variation
Between-person R-squared measures		
$R_b^{2(f_2)}$	Proportion of between-person outcome variance attributable to level-2 predictors via fixed components of slopes	Proportion of between-person outcome variance explained by the between-person-varying portion of level-1 and/or level-2 predictors via fixed components of slopes
$R_b^{2(v_2)}$	*N/A*	Proportion of between-person outcome variance attributable to the between-person-varying portion of level-1 predictors[a] via random slope variation
$R_b^{2(m)}$	Proportion of between-cluster outcome variance attributable to person-specific outcome means via random intercept variation	Proportion of between-person outcome variance attributable to person-specific outcome means via random intercept variation at the mean of all predictors with random slopes

Note:
*= For cluster-mean centered MLMs (where there is no v_2) the $R_t^{2(v_1)}$ can be simply referred to as $R_t^{2(v)}$ and the $R_w^{2(v_1)}$ can be simply referred to as $R_w^{2(v_1)}$ (as in Rights & Sterba, 2019, 2020).
[a]= See Table 2 note.
[†]= Rights and Sterba (2019, 2020) contrasted single-source measures with *combination-source* measures that are simple sums of these single-source measures, and hence quantify variance explained jointly by multiple sources. For example, $R_t^{2(f)} = R_t^{2(f_1)} + R_t^{2(f_2)}$, or $R_t^{2(fv)} = R_t^{2(f_1)} + R_t^{2(f_2)} + R_t^{2(v_1)} + R_t^{2(v_2)}$, or $R_t^{2(fvm)} = R_t^{2(f_1)} + R_t^{2(f_2)} + R_t^{2(v_1)} + R_t^{2(v_2)} + R_t^{2(m)}$.

tors with random slopes (Snijders & Bosker, 2012). This distinction is not necessary in the person-mean-centered decomposition, as all level-1 predictors have a mean of 0 by definition, and thus variance attributable to source m is simply the random intercept variance.

To illustrate the implications of different types of centering for MLM R-squared measures, we re-fit the unconditional growth model in Equation (1), but instead of person-mean-centering *time*, we now centered *time* at each of two different constant values. The first constant value we chose was the grand mean, and hence we *grand-mean*[6] *centered time* and computed the set of single-source R-squared measures from Table 3; parameter estimates and standard errors (SE) for this model are found in Table 1. The second constant value we chose was the origin, and hence we *centered time at the origin* and re-computed this set of R-squared measures from Table 3; parameter estimates and SEs for this model are again found in Table 1. First we focus on comparing results between these two centering-at-a-constant models. As expected based on statistical theory (e.g., Biesanz et al., 2004), the only difference between the parameter estimates of these centering-at-a-constant models is in the estimated fixed component of the intercept, the intercept variance, and the intercept-slope covariance, as the intercept has a different meaning in the two centering-at-a-constant models. However, consistent with the derivation in Appendix D, the set of R-squared measure results is nonetheless the same for both centering-at-a-constant models, as depicted in Figure 2.

Now we focus on comparing R-squared results between the centering-at-a-constant models (in Figure 2) and the person-mean-centered model (in Figure 1). In this illustrative example, recall that there was between-person variability in *time* scores when centering-at-a-constant (here created by attrition, leading to 13% of the variance in *time* being between-persons) but not when person-mean-centering *time*. Comparing Figures 1 and 2, we can see the only difference is that the person-mean-centered model R-squared results in Figure 1 do not contain any variance attributable to f_2 or v_2. This is because person-mean-centering *time* removes f_2 and v_2 as possible sources of explained variance, as *time* is then made a purely within-person variable. As such, although in Figure 1 the *between-person-varying portion* of *time* was nonexistent and thus explained 0% of total outcome variance and 0% of between-person outcome variance, in the left-most bar in Figure 2 we see that the *between-person-varying portion* of *time* explains 1% of the total outcome variance via its fixed component and 2% via its random component. In the right-most bar in Figure 2 the *between-person-varying portion* of *time* also now explains 3% of the between-person variance via its fixed component and 3% of the between-person variance via its random component.

In practice, the amount of between-person outcome variance explained by *time* will depend on how much between-person variance there is in *time*, that is, the degree to which people's mean values of *time* differ. Because the between-person variance in *time* is not usually reported in applications, it is difficult to know what is typical of practice. See Sterba (2014) and Mehta and West (2000) for further discussion. The within-person versus between-person variance explained by *time* itself could be due to missing-at-random attrition processes, as in our running example, or could be due to some substantively meaningful underlying between-person differences in an unmodeled variable (e.g., conscientiousness) leading to differences in attrition rates. Thus, finding that time explains a meaningful portion of between-person variance can aid researchers in considering ways

[6] When centering time at a constant as we do here (rather than person-mean-centering time as done earlier) researchers implicitly make the assumption that there is no contextual effect of time, meaning that the slope associated with within-person variability in *time* is equivalent to the slope associated with between-person variability in *time*. If this did not hold, then the estimated slope of *time* would then implicitly reflect a weighted average of the within-person and between-person slopes (see Enders & Tofighi, 2007; Rights, Preacher, & Cole, 2019).

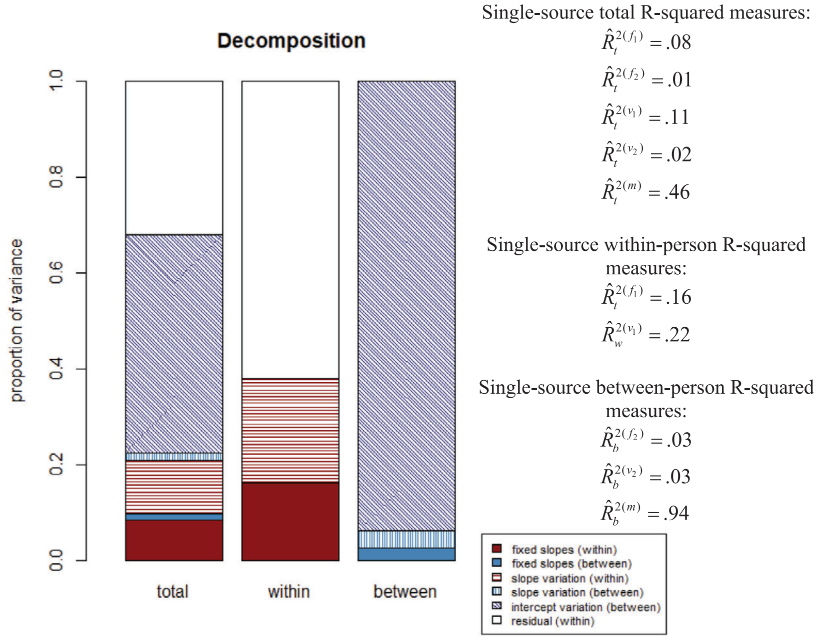

FIGURE 2 Visualizing R-squared results for the unconditional linear growth model of self-efficacy that centered *time* at a constant (e.g., the origin or the final assessment or the grand mean) and assumed a homoscedastic, diagonal error covariance structure (Equation 1)
Notes. See Figure 1 notes. Additionally, detailed definitions of each measure depicted in Figure 2, and subsequent figures, were given in Table 3 column 2. As explained in the manuscript text, the R-squared results in this Figure 2 would look the same regardless of which constant was chosen for centering time (e.g. the origin vs. the final assessment vs. the grand mean).

of explicitly accounting for such individual differences using other predictors in future modeling (e.g., by including a main effect of conscientiousness, as well as a cross-level interaction with *time*).

4 | NONLINEAR FUNCTIONAL FORM OF TIME: IMPLICATIONS FOR THE MLM R-SQUARED FRAMEWORK

Thus far our example analysis has assumed a linear relationship between self-efficacy and time; however, it is possible for the expected rate of change in self-efficacy itself to change over time. For instance, if the beginning of the assessment period involves a particular time of increased independence (e.g., adolescents obtaining a driver's license) or increased opportunity (e.g., joining extracurricular activities at the beginning of high school) it may be that adolescents tend to initially see large increases in self-efficacy but that self-efficacy eventually stops increasing as rapidly.

The most common method of modeling such nonlinear changes over time is to use polynomial functions of *time* (e.g., Bollen & Curran, 2006; Singer & Willett, 2003). For instance, to allow for a mean quadratic relationship between self-efficacy and *time* whose degree of acceleration varies across adolescents, we can add fixed and random components for $time^2$ to the model. We could additionally include fixed and random components of $time^3$ to allow for a cubic relationship, and so on. For simplicity, here will focus on a random quadratic model for self-efficacy:

$$
\begin{aligned}
\text{Level 1:}\quad & selfeff_{ij} = \beta_{0j} + \beta_{1j} time_{ij} + \beta_{2j} time^2_{ij} + e_{ij}, \\
& \beta_{0j} = \gamma_{00} + u_{0j}, \\
\text{Level 2:}\quad & \beta_{1j} = \gamma_{10} + u_{1j}, \\
& \beta_{2j} = \gamma_{20} + u_{2j}, \\
& e_{ij} \sim N(0, \sigma^2), \\
& \begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim MVN\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \tau_{00} & & \\ \tau_{01} & \tau_{11} & \\ \tau_{02} & \tau_{12} & \tau_{22} \end{bmatrix} \right).
\end{aligned} \tag{3}
$$

Though the interpretation of β_{0j} (and all components thereof) is unchanged from Equation (1), the interpretation of the slope of *time* has changed. The slope of the linear component, β_{1j}, is the person-specific expected rate of change in self-efficacy when $time = 0$ (with γ_{10} reflecting the across-person average and u_{1j} the person-specific deviation), whereas the slope of the quadratic component, β_{2j}, is the expected change in the rate of change in self-efficacy for each half-unit increase in time (with γ_{20} reflecting the across-person average and u_{2j} the person-specific deviation). The β_{1j} coefficient can be thought of as the "velocity" of self-efficacy when time = 0, whereas β_{2j} describes the "acceleration" of self-efficacy as a function of time (for further discussion and illustrations of such quadratic patterns of change, see Hoffman, 2015, p. 213).

A complication for R-squared computation when modeling such higher-order terms is that, *even* if the *time* variable varies exclusively within-person (e.g., if *time* is person-mean-centered), it is possible that $time^2$ can still vary between-person. In other words, when *time* by itself only explains within-person variability, $time^2$ can still explain some between-person variability (for instance, with our example self-efficacy dataset, even when person-mean-centering *time*, the intraclass correlation for $time^2$ is nonetheless .15 (i.e., 15% of the variability in $time^2$ lies at the between-person level). This is not well recognized in the context of R-squared computation. As derived in Appendix F, this occurs in the setting where there are different variances of *time* across persons, which arises in situations where there are truly individually varying measurement occasions and also in situations where persons had the same attempted measurement occasions but some persons had missing outcomes. Consequently, in such situations, some persons have more variability in measurement occasions than others.

Researchers modeling nonlinear terms involving any level-1 variable (e.g., $time^2$)—*regardless* of how its linear component (e.g., *time*) is centered—can obtain the outcome variance decomposition for *non*-person-mean-centered models, in Table 2 Column 3, and can use the set of R-squared measures for *non*-person-mean-centered models, in Table 3 column 3. These measures allow researchers to quantify, for instance, how much $time^2$ explains both within-person (via f_1 and v_1) and between-person (via f_2 and v_2). The amount of between-person outcome variance that is explained by $time^2$ when *time* itself varies solely within-person will reflect the degree to which persons have different variability in measurement occasions (as shown mathematically in Appendix F). Such variability in occasions may not itself be of substantive interest, and the $time^2$ term will explain extremely small amounts of between-person outcome variance when all persons have similar variability in occasions. On the other hand, if $time^2$ accounts for a large portion of between-person outcome variance, this can aid researchers in considering ways of accounting for such between-person differences in variability in measurement occasions in future modeling (e.g., considering whether treatment assignment leads to different degrees of variability in completing assessments).

To illustrate the computation and interpretation of MLM R-squared measures for growth models with polynomial functions of *time*, we fit the model in Equation (3) to our illustrative self-efficacy dataset, with *time* centered such that 0 = baseline. Parameter estimates and SEs are found in Table 1. We find that the estimated initial rate of change is positive ($\hat{\gamma}_{10} = 1.60$), and that the estimated rate of change slightly increases over time ($\hat{\gamma}_{20} = 0.05$); however, the latter estimate is non-significant (i.e., there is no evidence of any acceleration or deceleration in the rate of change). The Table 3 Column 3 R-squared estimates for this model are given in Appendix E Figure E1.[7] In a later section we delve more deeply into the interpretation and construction of differences-in-R^2s (i.e., ΔR^2), but for now note the following increments in variance explained by $time^2$ via its fixed and random components–going from the random linear growth model in Figure 2 to the random quadratic growth model in Appendix E Figure E1 we see the following increments: $\Delta\hat{R}_t^{2(f_1)} < .01$, $\Delta\hat{R}_t^{2(f_2)} < .01$, $\Delta\hat{R}_t^{2(v_1)} = .01$, and $\Delta\hat{R}_t^{2(v_2)} < .01$. In short, these R-squared values show little to no change when adding fixed and random components of $time^2$. For all further models, we will assume a linear trend over time for parsimony, excluding $time^2$.

5 | LEVEL-1 HETEROSCEDASTICITY: IMPLICATIONS FOR THE MLM R-SQUARED FRAMEWORK

In our running example, we have thus far assumed that there is only a single, homoscedastic, level-1 error variance, given by σ^2. Often times in longitudinal settings, however, one would theoretically expect that the level-1 error variance would differ across timepoints (i.e., have level-1 heteroscedasticity across time; Grimm & Widaman, 2010). For instance, at the earlier stages of data collection, the adolescents in our sample might be in relatively similar situations relevant to independence and self-efficacy (e.g., all having recently started high school, living with parents, etc.), whereas in later stages they might be more dissimilar (e.g., some having dropped out of high school, some moving out of their parent's house, some attending boarding school, etc.). If the model does not explicitly account for these differences with the included predictors, then we might expect the level-1 error variance to be greater at later timepoints than others. Erroneously constraining the level-1 error variance to be constant across all timepoints would then inaccurately reflect the actual generating process, and could manifest, for instance, as biased random slope variance and standard errors associated with the fixed effect of time (Snijders & Berkhof, 2008).

To relax the homoscedasticity constraint for σ^2, we can expand the model in Equation (3) by allowing level-1 error variances to differ across time. There are two basic approaches to doing so: (a) freely estimating the level-1 error variance for each individual timepoint, or (b) specifying a functional form by which the level-1 error variance changes as function of time. Though the former approach can accommodate complex patterns by which the error variance changes, the latter approach is more parsimonious, and can more readily

[7] Specifically, in Appendix E Figure E1 we see that *time* and $time^2$ together explain an estimated 9% of the total variance and 17% of the within-person variance via f_1 ($\hat{R}_t^{2(f_1)} = .09$, $\hat{R}_w^{2(f_1)} = .17$), 1% of the total variance and 3% of the between-person variance via f_2 ($\hat{R}_t^{2(f_2)} = .01$, $\hat{R}_b^{2(f_2)} = .03$), 13% of the total variance and 26% of the within-person variance via v_1 ($\hat{R}_t^{2(v_1)} = .13$, $\hat{R}_w^{2(f_1)} = .26$), and 2% of the total variance and 4% of the between-person variance via v_2 ($\hat{R}_t^{2(v_2)} = .02$, $\hat{R}_b^{2(f_2)} = .04$).

allow for individually varying timepoints (e.g., Diggle, Heagerty, Liang, & Zeger, 2002). For our running example that has a discrete set of measurement occasions, we chose to flexibly allow the error variances to differ across timepoints, as shown in Equation (4):

$$\begin{bmatrix} e_{1j} \\ e_{2j} \\ \vdots \\ e_{10j} \end{bmatrix} \sim MVN\left(\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & & & \\ 0 & \sigma_2^2 & & \\ \vdots & \ddots & \ddots & \\ 0 & \cdots & 0 & \sigma_I^2 \end{bmatrix}\right). \tag{4}$$

Here, rather than a single σ^2, there is now a separately estimated σ_i^2 specific to each timepoint $i = 1 \ldots I$. (For a demonstration of instead specifying the variance as a continuous parametric function of *time*, rather than estimating a discrete set of variances, see Appendix G.)

In the original Rights and Sterba (2019) framework, level-1 errors were assumed to be homoscedastic. This means that previously researchers could not obtain R-squared measures from this framework when fitting an MLM with level-1 heteroscedasticity. To expand this R-squared framework to accommodate level-1 heteroscedasticity, we show analytically in Appendix G that the framework can accommodate heteroscedastic level-1 error variance by replacing the σ_i^2 term with the expected value of σ_i^2 across all timepoints. With balanced measurement occasions and freely estimating each time-specific variance, this resolves to the average of the σ_i^2 estimates across timepoints. When the level-1 error variance is instead modeled as a continuous parametric function of time—which readily accommodates individually varying measurement occasions—we provide formulas to compute the estimated expected value of σ_i^2 when σ_i^2 follows a linear or a quadratic functional form in Appendix G Equations (G7)–(G10).

Our running simulated self-efficacy example was generated with heteroscedastic time-specific level-1 error variances in the population. When we added time-specific level-1 error variances, as in Equation (4), to our fitted linear self-efficacy growth model, the error variance is indeed estimated to be smaller at the earlier timepoints, and larger at the later timepoints (see Table 1 for parameter estimates and SEs). Appendix E Figure E2 shows that R-squared results, however, are relatively unchanged from the prior homoscedastic linear growth model in Equation (3) (Figure 2). It is likely, however, that under certain circumstances R-squared measures computed from a homoscedastic model can be biased when level-1 heteroscedasticity exists, given that inappropriately constraining the error variance to be homoscedastic can induce bias in random slope variances (Snijders & Berkof, 2008).

It is important to note that this R-squared framework extension to accommodate level-1 heteroscedasticity is not designed to be used to *diagnose* the presence of level-1 heteroscedasticity because, as illustrated here, R-squared measures assuming homoscedasticity versus allowing level-1 heteroscedasticity can be similar, even when level-1 heteroscedasticity exists in the population. Rather, the advantage of using the framework extension presented here to accommodate level-1 heteroscedasticity is to allow MLM parameters and their standard errors to be accurately estimated by correctly modeling theorized level-1 heteroscedasticity in the fitted model, without needing to fit an underspecified model simply to obtain R-squareds.

6 | NON-DIAGONAL RESIDUAL COVARIANCE STRUCTURE: IMPLICATIONS FOR THE MLM R-SQUARED FRAMEWORK

Though we extended the level-1 error covariance structure in the prior section by allowing for heteroscedasticity, we have thus far assumed a *diagonal* error covariance structure in that only the variances (i.e., the diagonal elements of the covariance matrix of the e's in Equation 4) were estimated, whereas the covariances (i.e., the off-diagonal elements of the covariance matrix of the e's in Equation 4) were fixed at 0. This assumption is reasonable in cross-sectional data in which there is no meaningful sequential order of observations within persons, and may be reasonable in longitudinal contexts where measurement occasions are spaced far apart in time. However, particularly in longitudinal designs in which measurements are spaced close together in time (e.g., in ecological momentary assessments or daily diary data; see Bolger & Laurenceau, 2013), it is likely that the predictors and random effects included in the model do not perfectly capture the similarity between observations (e.g., Bollen & Curran, 2004; Campbell & Kenny, 1999; Goldstein, Healy, & Rasbash, 1994; Kwok et al., 2008; Singer & Willet, 2003). In the event of such *autocorrelated* errors, failing to properly account for this autocorrelation can lead to biased standard errors for fixed coefficients as well as biased random effect variances (Ferron, Dailey, & Yi, 2002; Kwok, West, & Green, 2007; Sivo, Fan, & Witta, 2005).

Rights and Sterba's (2019) original R-squared framework dealt exclusively with diagonal covariance structures, meaning there was no autocorrelation. As we prove in Appendix H, the mathematical computation of the original framework's R-squared measures (i.e., the specific set of formulae used) is unaffected by the inclusion of any kind of autocorrelation (e.g., unstructured, Toeplitz, compound symmetric, or first-order autoregressive error covariance structures, to name a few; see Wolfinger (1993) or Diggle (1990) for a more extensive review of possibilities).[8] However, given that ignoring autocorrelation in the level-1 errors when fitting MLMs can lead to biased random effect variances, R-squared measures obtained from a model that fails to properly include autocorrelation can similarly be distorted.

To illustrate, we revisit our running simulated example which was generated with autocorrelation in the population; specifically, it was generated with a first-order autoregressive structure, which is commonly employed in developmental applications of growth models (e.g., Curran & Bollen, 2001; Goldstein et al., 1994). Here, error covariance between two timepoints a and b is given as $\sqrt{\sigma_a^2\sigma_b^2}\rho^{|a-b|}$. This yields an error correlation between adjacent timepoints of ρ, and a weaker correlation for timepoints separated by more than one unit of time. This first-order autoregressive error covariance structure (with heteroscedastic variances) is given as:

$$\begin{bmatrix} e_{1j} \\ e_{2j} \\ \vdots \\ e_{Ij} \end{bmatrix} \sim MVN\left(\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & & & \\ \sqrt{\sigma_1^2\sigma_2^2}\rho & \sigma_2^2 & & \\ \vdots & \ddots & \ddots & \\ \sqrt{\sigma_1^2\sigma_I^2}\rho & \cdots & \sqrt{\sigma_{I-1}^2\sigma_I^2}\rho & \sigma_I^2 \end{bmatrix}\right). \tag{5}$$

We specified this heteroscedastic first-order autoregressive error covariance structure in Equation (5) when fitting our linear self-efficacy growth model and computed the R-squared results. From the parameter estimates in Table 1, we see that the ρ parameter is estimated to be .31, indicating a modest correlation between subsequent timepoints. From the R-squared results shown in Appendix E Figure E3, of note is that the proportion of variance attributable to v and m decreased vis-a-vis the previous model that ignored

[8] This proof also applies to many other such measures that utilize the estimated level-1 error variance in their computation (e.g., Hox, 2010; Johnson, 2014; Kreft & de Leeuw, 1998; Nakagawa & Schielzeth, 2013; Raudenbush & Bryk, 2002; Snijders & Bosker, 2012). Thus, in the current paper, we supply a novel proof that the formulas of these pre-existing measures additionally hold under autocorrelation.

the autocorrelation, that is, Appendix E Figure E2 (in which $\Delta\hat{R}_t^{2(v_1)} = -.04$, $\Delta\hat{R}_t^{2(v_2)} = -.01$, and $\Delta\hat{R}_t^{2(m)} = -.05$—indicating that, cumulatively, v and m accounted for 10% less outcome variance in the properly specified model). Hence, a researcher who ignored autocorrelated errors by erroneously fitting a diagonal error covariance structure would obtain a misleadingly large estimate of the degree of between-adolescent differences in terms of growth trajectories and mean levels of self-efficacy. This finding is consistent with the prior findings that failing to appropriately model autocorrelated errors can yield inflated estimates of random effect variances, as the random effects in the underspecified model help to account for covariances among repeated measures that are actually due to omitted autocorrelations (e.g., Ferron et al., 2002; Kwok et al., 2007; Sivo et al.,2005).

7 | ADDING TIME-VARYING COVARIATES (LEVEL-1 PREDICTORS OTHER THAN *TIME*) AND TIME-INVARIANT COVARIATES (LEVEL-2 PREDICTORS): ILLUSTRATION OF MLM R-SQUARED DIFFERENCES

Usually, once the centering location for *time* has been chosen, the functional form of *time* has been identified, and the level-1 and level-2 covariance structure of the unconditional growth model have been determined, substantive interest turns to the potential impact of additional predictors, rather than just change over time. For instance, when examining self-efficacy over time, a researcher can consider level-1 predictors other than *time*, that is, time-varying covariates (e.g., "do adolescents have higher levels of self-efficacy at time i when they are actively volunteering versus when they are not actively volunteering at that timepoint?"), as well as level-2 predictors, that is, time-invariant covariates (e.g., "do adolescents with higher GPA tend to have higher levels of self-efficacy than adolescents with lower GPA?"). Additionally, researchers can examine how change over time may differ across levels of certain predictors by allowing cross-level interactions with *time* (e.g., "do female adolescents have a higher rate of change in self-efficacy than male adolescents?").

In so doing, beyond considering a set of overall model R-squareds for the researcher's final (full) MLM, the researcher can also compute R-squared differences (ΔR^2) between models to quantify the effect size associated with individual term(s). To explain, we denote the difference in a generic single-source R-squared measure between models as $R^2_{\text{mod } B} - R^2_{\text{mod } A} = \Delta R^2$, wherein $R^2_{\text{mod } B}$ is obtained from the more complex model that contains the added term(s) of interest, and $R^2_{\text{mod } A}$ is obtained from the model that excludes these particular term(s).

Importantly, the impact of each kind of term added to the Model A MLM to form the Model B MLM can be detected by a particular single-source ΔR^2 measure (Rights & Sterba, 2020). For instance, the single-source R-squared difference measure $\Delta R_t^{2(f_1)}$ is suited to detect the overall contribution of a person-mean-centered level-1 predictor via its fixed component (source f_1), whereas $\Delta R_w^{2(f_1)}$ is suited to detect the *level-specific* contribution of that person-mean-centered level-1 predictor via its fixed component (source f_1). Using the wrong single-source ΔR^2 to assess the impact of an added term can lead the researcher to erroneously conclude that there is no contribution of that predictor, when that predictor may indeed have an impact, just via a different source (Rights & Sterba, 2020). Rights and Sterba (2020) also showed that using combination-source ΔR^2 measures to assess the impact of an added term can also be problematic because the added term can have opposite effects on different sources that cancel out when combined, thus preventing the researcher from realizing that there was any impact of the predictor whatsoever. As such, we recommend that single-source, not combination-source, ΔR^2 be computed and

TABLE 4 Which target single-source ΔR^2 can be used to quantify the impact of individual term(s) added to the MLM during model building?

	When using R^2 measures that assume person-mean-centering:		When using R^2 measures that do not assume person-mean-centering:	
The term added to the MLM whose unique contribution is to be detected is a…	**To detect the overall impact of the term, use this single-source ΔR^2 measure**	**To detect the level-specific impact of the term, use this single-source ΔR^2 measure**	**To detect the overall impact of the term, use this single-source ΔR^2 measure**	**To detect the level-specific impact of the term, use this single-source ΔR^2 measure**
…fixed component of a level-1 predictor's slope (including a level-1 × level-1 or level-1 × level-2 predictor product term's slope)	$\Delta R_t^{2(f_1)}$	$\Delta R_w^{2(f_1)}$	$\Delta R_t^{2(f_1)}$ (also $\Delta R_t^{2(f_2)}$ if level-1 predictor has between-person variance*)	$\Delta R_w^{2(f_1)}$ (also $\Delta R_b^{2(f_2)}$ if level-1 predictor has between-person variance*)
… level-2 predictor's slope (including a level-2 × level-2 predictor product term's slope)	$\Delta R_t^{2(f_2)}$	$\Delta R_b^{2(f_2)}$	$\Delta R_t^{2(f_2)}$	$\Delta R_b^{2(f_2)}$
…random component of a level-1 predictor's slope	$\Delta R_t^{2(v_1)}$	$\Delta R_w^{2(v_1)}$	$\Delta R_t^{2(v_1)}$ (also $\Delta R_t^{2(v_2)}$ if level-1 predictor has between-person variance*)	$\Delta R_w^{2(v_1)}$ (also $\Delta R_b^{2(v_2)}$ if level-1 predictor has between-person variance*)

Notes: For person-mean-centered MLMs (where there is no v_2) the v_1 can simply be referred to as v (as in Rights & Sterba, 2019, 2020).

*This level-1 predictor could be a higher-order polynomial term (e.g., quadratic or cubic term or a product term involving multiple level-1 predictors). Higher-order polynomial terms associated with level-1 predictors (e.g., x_{ij}^2) as well as product terms involving multiple level-1 predictors (e.g., $x_{ij} \times time_{ij}$) can have between-cluster variance, even when their constituent level-1 predictors (x_{ij} and $time_{ij}$) are themselves cluster-mean-centered (see our derivation in Appendix F), and in this event the procedures in the fourth and fifth columns of Table 4 can be followed.

examined to detect the impact of added terms. For didactic purposes, Table 4 (first column) lists each possible kind of term that can be added to the Model A MLM to form the Model B MLM; remaining columns list which *target* single-source ΔR^2 measure should be used to detect the (overall or level-specific) impact of that added term. More specifically, contents of the second and third columns of Table 4 pertain to person-mean-centered models and are reviewed from Rights and Sterba (2020). Contents of the fourth and fifth columns of Table 4 pertain to *non*-person-mean-centered models and are a novel contribution of the present paper.

Next we illustrate the use of Table 4 by assessing the unique contribution of two time-varying covariates and two time-invariant covariates, as well as a cross-level interaction of a time-invariant-covariate and *time*, for our running self-efficacy example. Following widespread recommendations, we person-mean-centered our time-varying covariates to avoid conflation of level-specific effects (Enders & Tofighi, 2007; Raudenbush & Bryk, 2002; Rights, Preacher, & Cole, 2019); these included person-mean-centered grade-point-average (GPA, i.e., an adolescent's GPA at a given timepoint relative to their average GPA) and person-mean-centered time spent volunteering in the community (*volunteer*, i.e., an adolescent's community service involvement at the time of assessment relative to their average number of hours of community service each week). For the time-invariant covariates, we included person-mean GPA, person-mean *volunteer*, and *female* (a dichotomous variable in which 1 indicates female and 0 indicates male). Including both the within-person and between-person components of GPA and *volunteer* allows us to separately assess potential within-person and between-person effects (e.g., Curran & Bauer, 2011). The cross-level interaction term is *female* × *time*, which allows the growth rate in self-efficacy to differ between females and males. These predictors are each added to our self-efficacy unconditional linear growth model with heteroscedastic first-order autoregressive errors, and with time centered at the origin, to form the full model of interest:

$$
\begin{aligned}
&\text{Level 1: } selfeff_{ij} = \beta_{0j} + \beta_{1j} time_{ij} + \beta_{2j}\left(gpa_{ij} - gpa_{.j}\right) + \beta_{3j}\left(volunteer_{ij} - volunteer_{.j}\right) + e_{ij},\\
&\text{Level2: } \beta_{0j} = \gamma_{00} + \gamma_{01} gender_j + \gamma_{02} gpa_{.j} + \gamma_{03} volunteer_{.j} + u_{0j},\\
&\qquad \beta_{1j} = \gamma_{10} + \gamma_{11} gender_j + u_{1j},\\
&\qquad \beta_{2j} = \gamma_{20} + u_{2j},\\
&\qquad \beta_{3j} = \gamma_{30} + u_{3j}
\end{aligned} \tag{6}
$$

where:

$$
\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim MVN\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \tau_{00} & & \\ \tau_{01} & \tau_{11} & \\ \tau_{02} & \tau_{12} & \tau_{22} \end{bmatrix}\right),
$$

and

$$
\begin{bmatrix} e_{1j} \\ e_{2j} \\ \vdots \\ e_{Ij} \end{bmatrix} \sim MVN\left(\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & & & \\ \sqrt{\sigma_1^2\sigma_2^2}\rho & \sigma_2^2 & & \\ \vdots & \ddots & \ddots & \\ \sqrt{\sigma_1^2\sigma_I^2}\rho & \cdots & \sqrt{\sigma_{I-1}^2\sigma_I^2}\rho & \sigma_I^2 \end{bmatrix}\right).
$$

Parameter estimates and SE results for the Equation (6) model are found in Table 1. Here we see that each predictor has a positive and significant association with self-efficacy. While these raw coefficients are informative, it is difficult to quantify and compare their associated effect sizes, as coefficients are on differing metrics and we cannot tell how much variance each term uniquely contributes. We hence supplement these results using R-squared difference measures from Table 4 Columns 4–5 (rather than Columns 2–3 because *time* was not person-mean-centered). Here we compute increments in variance for (i.e., ΔR^2) using a *simultaneous* model-building approach.[9] For simplicity of presentation, we

[9] In defining the unique proportion of variance accounted for by a particular added term, Rights and Sterba (2020) distinguish between a *hierarchical* model-building approach (wherein the proportion of variance contributed by each newly added term is defined controlling for previously added terms, but not subsequently added terms) versus a *simultaneous* approach (wherein the proportion of variance contributed by an added term controls for all other terms in the full model). Under a hierarchical approach, one starts with a baseline model (e.g., an unconditional growth model) and adds terms sequentially until forming the full MLM of interest, computing single-source R-squared differences associated with each added term at each step. Under a simultaneous approach, one compares the full MLM of interest to a set of MLMs, each of which excludes a term, and the researcher computes single-source R-squared differences to quantify the proportion of variance uniquely attributable to each excluded single term, in

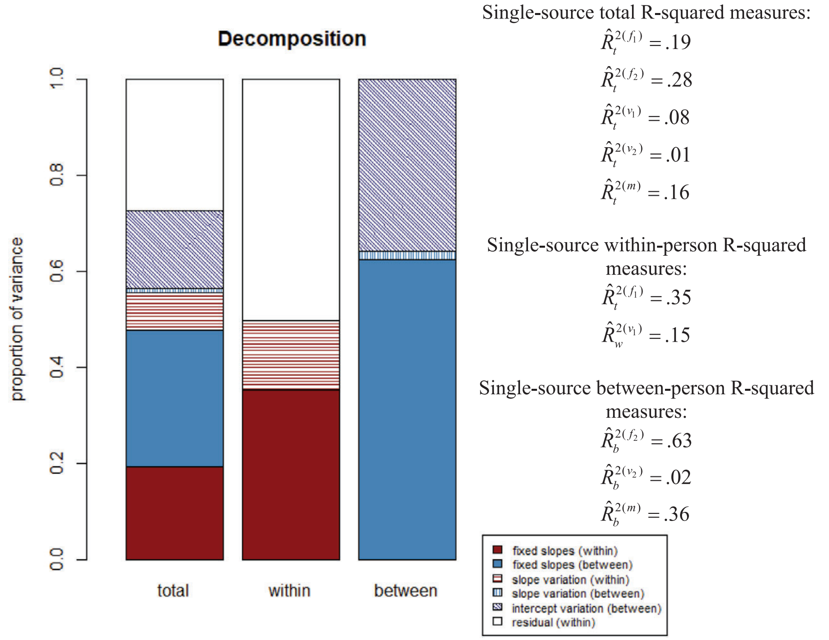

FIGURE 3 Visualizing R-squared results for the conditional linear growth model of self-efficacy that centered *time* at a constant and specified a heteroscedastic, first-order autoregressive covariance structure and included time-varying and time-invariant predictors (Equation 6)
Notes. See Figure 1 and 2 notes.

focus interpretation total ΔR^2 results here, but report corresponding level-specific ΔR^2s in footnotes.

We first estimate the proportion of variance uniquely explained by GPA. Note that this involves three separate model terms each having a different source of explained variance—the fixed component of person-mean-centered GPA, the random component of person-mean-centered GPA, and the fixed component of person-mean GPA. We will thus compare the full model in Equation (6) (whose R-squared results were provided in Figure 3) to a model that removes these three components (given in Appendix I Equation I1) and we quantify the proportion of total variance uniquely explained by these terms with, $\Delta R_t^{2(f_1)}$, $\Delta R_t^{2(f_2)}$, and $\Delta R_t^{2(v_1)}$, respectively. We estimate that GPA explains more variance via its within-person fixed component than via its between-person fixed component, as $\Delta \hat{R}_t^{2(f_1)} = .06$ and $\Delta \hat{R}_t^{2(f_2)} < .01$. This implies a stronger impact of *current* grades at a given assessment, relative to an adolescent's own typical grades, with adolescents having meaningfully more self-efficacy when doing well relative to their own average, and less self-efficacy when doing worse relative to their average. We additionally find that the within-person varying portion of GPA does not contribute a meaningful amount of outcome variance via random slopes, as $\Delta \hat{R}_t^{2(v_1)} < .01$.[10]

turn (see Rights & Sterba, 2020 for more details). Additionally, rather than adding/removing only a single term at a time, unique term-specific contributions can also be identified when multiple terms are added at a given step (under a hierarchical approach), or multiple terms are removed from the full model (under the simultaneous approach), so long as these multiple terms each reflect different sources of explained variance (which we illustrate for GPA and *volunteer* in our example).

[10] For these three terms, total measures were reported in the manuscript text and level-specific measures are provided here: $\Delta \hat{R}_w^{2(f_1)} = .12$, $\Delta \hat{R}_w^{2(v_1)} < .01$, $\Delta \hat{R}_b^{2(f_2)} < .01$.

We will next repeat this procedure for volunteer hours by comparing the full model in Equation (6) to a model that removes the fixed component of person-mean-centered volunteer hours, the random component of person-mean-centered volunteer hours, and the fixed component of person-mean volunteer hours (given in Appendix I Equation I2). We find largely the opposite pattern than what we found for GPA—more variance is explained by volunteerism via its between-person fixed component ($\Delta\hat{R}_t^{2(f_2)} = .19$) than its within-person fixed component ($\Delta R_t^{2(f_1)} = 0.01$). This suggests that adolescents who over time *tend* to volunteer more have meaningfully higher levels of self-efficacy than those who rarely volunteer, but that the *relative* amount of volunteering at a given assessment explains little of the variance. Another pattern we find here that differs from GPA, is that volunteerism accounts for a non-zero proportion of variance in self-efficacy via its random slope, quantified as $\Delta\hat{R}_t^{2(v_1)} = .03$;[11] this highlights the possibility for future modeling to consider including cross-level interactions to account for such random slope variation (see Aguinis & Culpepper, 2015; Rights & Sterba, 2019). As an example, it could be that adolescents higher in altruism have a more positive association of volunteering and self-efficacy than adolescents who are less altruistic.

Lastly, we can compute an effect size associated with the interaction of *time* and *female* by estimating the proportion of total variation explained by their product term via its fixed component. We compare the full model in Equation (6) to a reduced model that excludes this product term, given in Appendix I Equation (I3). Because this product term involves a non-cluster-mean-centered level-1 predictor (*time*) and a level-2 predictor (*female*), this product term can vary both within-cluster and between-cluster. As such, in accordance with Table 4, we need to examine both $\Delta R_t^{2(f_1)}$ and $\Delta R_t^{2(f_2)}$; we see that this term cumulatively explains an estimated 8% of the total outcome variance ($\Delta\hat{R}_t^{2(f_1)} = .03$ and $\Delta\hat{R}_t^{2(f_2)} = .05$), providing an indication of the importance of considering gender-specific differences in trajectories.[12]

8 | DISCUSSION

8.1 | Summary

Developmental researchers often fit longitudinal growth models to examine change over time in children and adolescents' behaviors, symptoms, or abilities. For such multilevel analyses of repeated measures data, however, developmental researchers have received limited methodological guidance on employing and interpreting effect size metrics such as R-squared measures. Though a recently developed framework of MLM R-squared computation (Rights & Sterba, 2019, 2020) integrated, analytically related, and filled gaps in previously published MLM R-squareds to provide a coherent approach for reporting, visualizing, and interpreting these measures, it concentrated on cross-sectional applications that used cluster-mean-centering. The current paper extended the Rights and Sterba (2019, 2020) original framework to longitudinal contexts by (a) clarifying how the interpretation and computation of R-squared (R^2) and R-squared difference (ΔR^2) measures are modified to accommodate alternative centering strategies common in longitudinal applications (e.g., centering-at-a-constant such as the origin), and (b) clarifying what measures

[11] For these three terms, total measures were reported in the manuscript text and level-specific measures are provided here: $\Delta\hat{R}_w^{2(f_1)} < .01, \Delta\hat{R}_w^{2(v_1)} = .05$, $\Delta\hat{R}_b^{2(f_2)} = .43$.

[12] For this term, total measures were reported in the manuscript text and level-specific measures are provided here: $\Delta\hat{R}_w^{2(f_1)} = .06$ and $\Delta\hat{R}_b^{2(f_2)} = .09$.

to use when incorporating: nonlinear effects (e.g., of *time*), level-1 heteroscedasticity, and autocorrelation of level-1 errors. We illustrated the application of this extended R-squared framework in a running illustrative growth model that described and predicted change over time in adolescent self-efficacy.

8.2 | Software implementation

To aid researchers in implementing this expanded R-squared framework in practice, we have a provided an R function, *r2MLMlong*. In Appendix I, we provide the code for this function, as well as example input and a description of all of the function arguments. This function expands upon the original *r2MLM* function (Rights & Sterba, 2019)[13] by (a) allowing researchers to specify either homoscedastic or heteroscedastic level-1 residual variance and (b) providing a full decomposition of variance and full set of R-squared measures for both person-mean-centered and *non*-person-mean-centered models (as shown in Tables 3 and 4).

8.3 | Future directions

Some limitations of the current work serve as future research directions. Here we restricted focus to models that are "linear in the parameters," meaning that the expected value of the outcome is a simple sum of coefficients weighted by predictors—which still allows for the possibility for the predictors to be polynomial (e.g., quadratic or cubic) functions of time, as discussed earlier. Longitudinal researchers sometimes wish to employ more complex functional forms where the coefficients do not necessarily enter the model linearly. Examples include exponential, logarithmic, or power functions (see Cudeck & Harring, 2007, for a review). Similarly, researchers may also wish to model the error variance using one of these such intrinsically nonlinear functions (e.g., Browne & Du Toit, 1991). Though empirical applications of such intrinsically nonlinear functions are much rarer than the polynomial growth models discussed in this paper (Grimm, Ram, & Hamagami, 2011), future work can expand the current framework for models that are nonlinear in the parameters.

Second, here we focused on growth models fit in an MLM framework. However, it is widely appreciated that growth models specified as MLMs can alternatively be fit within a structural equation modeling (SEM) framework (e.g., Bauer, 2003; Curran, 2003). The models discussed in this paper could be fit via SEM, and our provided software could be used to compute R-squared measures (although the dataset itself would need to be converted from wide format—to fit the SEM—to long format—to use our function).

Last, in the current work, we did not specify effect size benchmarks for what constitutes a "small," "medium," or "large" contribution of variance for a particular kind of term in the MLM via a particular source. Though such rules of thumb are often employed in single-level regression analyses (e.g., Cohen, 1988), benchmarks have neither been proposed nor systematically evaluated for R-squared measures in longitudinal or cross-sectional multilevel contexts. For the latter contexts, reporting a common framework of MLM R-squared measures needs to become routine in applied practice before subfields can amass substantive grounding to inform rence points for these measures (see Schafer & Schwarz, 2019).

[13] Shaw, Rights, Sterba, and Flake (2020) recently provided an R package in which users can obtain R-squared measures from Rights & Sterba's (2019) original framework by directly entering fit objects obtained from the MLM model-fitting packages *lme4* (Bates, Maechler, Bolker, & Walker, 2014) or *nlme* (Pinheiro et al., 2020); a future direction is to incorporate the *r2MLMlong* function into this package. The current provided *r2MLMlong* function offers the flexibility of obtaining estimates via any MLM software.

8.4 | Conclusions

Though longitudinal growth models have proven to be an invaluable and popular analytic tool for developmental researchers, there has been little guidance provided on how to construct and interpret MLM R-squareds for these models. This stands at odds with the current widespread recommendations to consider effect size, rather than relying on the exclusive reporting of statistical significance, when reporting MLM results (e.g., APA, 2009; Hoffman, 2015; Kwok et al., 2008; Lorah, 2018; Nezlek, 2012). This paper fills this gap by providing a framework by which researchers can compute, visualize, compare, and interpret R-squared measures for longitudinal MLMs. In so doing, we delineated how the various model specification decisions faced by longitudinal researchers have important implications for R-squared measures.

ORCID
Jason D. Rights https://orcid.org/0000-0003-3784-9772

REFERENCES

Aguinis, H., & Culpepper, S. A. (2015). An expanded decision-making procedure for examining cross-level interaction effects with multilevel modeling. *Organizational Research Methods*, *18*, 155–176.
American Psychological Association. (2009). Publication manual of the *American Psychological Association* (6th ed.). Washington, DC: American Psychological Association.
Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). Lme4: Linear mixed-effects models using Eigen and S4. *R Package Version*, *1*, 1–23.
Bauer, D. J. (2003). Estimating multilevel linear models as structural models. *Journal of Educational and Behavioral Statistics*, *28*, 135–167.
Biesanz, J. C., Deeb-Sossa, N., Papadakis, A. A., Bollen, K. A., & Curran, J. C. (2004). The role of coding time in estimating and interpreting growth curve models. *Psychological Methods*, *9*, 30–52.
Bolger, N., & Laurenceau, J. (2013). *Intensive longitudinal methods: An introduction to diary and experience sampling research*. New York: Guilford Press.
Bollen, K. A., & Curran, P. J. (2006). *Latent curve models: A structural equation perspective*. Hoboken, NJ: Wiley.
Bollen, K. A., & Curran, P. J. (2004). Autoregressive latent trajectory (ALT) models: A synthesis of two traditions. *Sociological Methods and Research*, *32*, 336–383.
Browne, M. W., & DuToit, S. H. C. (1991). Models for learning data. In L. Collins & J. Horn (Eds.), *Best methods for the analysis of change. Recent advances, unanswered questions, future directions* (pp. 47–68). Washington DC: American Psychological Association.
Campbell, D. T., & Kenny, D. A. (1999). *A primer on regression artifacts*. New York: Guilford.
Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
Crowder, M. J., & Hand, D. J. (1990). *Analysis of repeated measures*. Boca Raton, FL: Chapman & Hall/CRC.
Cudeck, R., & Harring, J. R. (2007). Analysis of nonlinear patterns of change with random coefficient models. *Annual Review of Psychology*, *58*, 615–637.
Curran, P. J. (2003). Have multilevel models been structural equation models all along? *Multivariate Behavioral Research*, *38*, 529–569.
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., & Bollen, K. A. (2001). The best of both worlds: Combining autoregressive and latent curve models. In L. M. Collins & A. G. Sayer (Eds.), *New methods for the analysis of change* (pp. 105–136). Washington, DC: American Psychological Association.
Diggle, P. J., Heagerty, P., Liang, K.-Y., & Zeger, S. L. (2002). *Analysis of longitudinal data* (2nd ed.). Oxford: Oxford University Press.
Diggle, P. J. (1990). *Time series: A biostatistical introduction*. Oxford: Clarendon Press.
Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods*, *12*, 121–138.
Fergus, S., Zimmerman, M., & Caldwell, C. (2007). Growth trajectories of sexual risk behavior in adolescence and young adulthood. *American Journal of Public Health*, *97*, 1096–1101.
Ferron, J., Dailey, R., & Yi, Q. (2002). Effects of misspecifying the first-level error structure in two-level models of change. *Multivariate Behavioral Research*, *37*, 379–403.
Goldstein, H. (2011). *Multilevel statistical models* (4th ed.). New York: Wiley.

Goldstein, H., Healy, M. J., & Rasbash, J. (1994). Multilevel time series models with applications to repeated measures data. *Statistics in Medicine*, *13*, 1643–1655.
Grimm, K. J., & Widaman, K. F. (2010). Residual structures in latent growth curve modeling. *Structural Equation Modeling*, *17*, 424–442.
Grimm, K. J., Ram, N., & Hamagami, F. (2011). Nonlinear growth curves in developmental research. *Child Development*, *82*, 1357–1371.
Hoeksma, J. B., & Koomen, H. M. Y. (1992). Multilevel models in developmental psychological research: Rationales and applications. *Early Development and Parenting*, *1*, 157–167.
Hoffman, L. (2015). *Longitudinal analysis: Modeling within person fluctuation and change*. New York: Routledge.
Hox, J. J. (2010). *Multilevel analysis. Techniques and applications* (2nd ed.). New York: Routledge.
Johnson, P. C. (2014). Extension of Nakagawa & Schielzeth's R2GLMM to random slopes models. *Methods in Ecology and Evolution*, *5*, 944–946.
Johnson, M. K., Beebe, T., Mortimer, J. T., & Snyder, M. (1998). Volunteerism in adolescence: A process perspective. *Journal of Research in Adolescence*, *8*, 309–332.
King, K. M., Littlefield, A. K., McCabe, C. J., Mills, K. L., Flournoy, J., & Chassin, L. (2018). Longitudinal modeling in developmental neuroimaging research: Common challenges, and solutions from developmental psychology. *Developmental Cognitive Neuroscience*, *33*, 54–72.
Kreft, I. G., & de Leeuw, J. (1998). *Introducing multilevel modeling*. Thousand Oaks, CA: Sage.
Kreft, I. G. G., de Leeuw, J., & Aiken, L. S. (1995). The effects of different forms of centering in hierarchical linear models. *Multivariate Behavioral Research*, *30*, 1–21.
Kwok, O.-M., Underhill, A. T., Berry, J. W., Luo, W., Elliot, T. R., & Yoon, M. (2008). Analyzing longitudinal data with multilevel models. An example with individuals living with lower extremity intra-articular fractures. *Rehabilitation Psychology*, *53*, 370–386.
Kwok, O., West, S. G., & Green, S. B. (2007). The impact of misspecifying the within-subject covariance structure in multiwave longitudinal multilevel models: A Monte Carlo study. *Multivariate Behavioral Research*, *42*, 557–592.
Larose, S., Ratelle, C. F., Guay, F., Senecal, C., Harvey, M., & Drouin, E. (2008). A sociomotivational analysis of gender effects on persistence in science and technology: A 5-year longitudinal study. In H. M. G. Watt & J. S. Eccles (Eds.), *Explaining gendered occupational outcomes: Examining individual and social explanations through school and beyond* (pp. 171–192). Washington, DC: American Psychological Association.
Lorah, J. (2018). Effect size measures for multilevel models: Definition, interpretation, and TIMSS example. *Large-Scale Assessments in Education*, *6*, 6–8.
LaHuis, D. M., Hartman, M. J., Hakoyama, S., & Clark, P. C. (2014). Explained variance measures for multilevel models. *Organizational Research Methods*, *17*, 433–451.
LaHuis, D. M., Blackmore, C. E., & Bryant-Lees, K. B. (2019). Explained variance measures for multilevel models. In S. E. Humphrey & J. M. LeBreton (Eds.), *The handbook of multilevel theory, measurement, and analysis* (pp. 353–364). Washington, DC: American Psychological Association.
Mehta, P. D., & West, S. G. (2000). Putting the individual back in to individual growth curves. *Psychological Methods*, *5*, 23.
Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, *4*, 133–142.
Nezlek, J. B. (2012). Multilevel modeling for psychologists. In H. Cooper (Ed.), *Handbook of research methods in psychology* (pp. 219–241). Washington, DC: American Psychological Association.
Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & Core Team R. (2020). nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1-151, https://CRAN.R-project.org/package=nlme.
Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Newbury Park, CA: Sage.
Rights, J. D., Preacher, K. J., & Cole, D. A. (2019). The danger of conflating level-specific effects of control variables when primary interest lies in level-2 effects. *British Journal of Mathematical and Statistical Psychology*, *73*, 194–211.
Rights, J. D., & Sterba, S. K. (2016). The relationship between multilevel models and non-parametric multilevel mixture models: Discrete approximation of intraclass correlation, random coefficient distributions, and residual heteroscedasticity. *British Journal of Mathematical and Statistical Psychology*, *69*, 316–343.
Rights, J. D., & Sterba, S. K. (2019). Quantifying explained variance in multilevel models: An integrative framework for defining R-squared measures. *Psychological Methods*, *24*, 309–338.
Rights, J. D., & Sterba, S. K. (2020). New recommendations on the use of R-squared differences in multilevel models. *Multivariate Behavioral Research*, *55*, 568–599.
Schäfer, T., & Schwarz, M. (2019). The meaningfulness of effect sizes in psychological research: Differences between sub-disciplines and the impact of potential biases. *Frontiers in Psychology*, *10*, 813.
Shaw, M., Rights, J. D., Sterba, S. K., & Flake, J. K. (2020). r2mlm: R-Squared Measures for Multilevel Models. https://doi.org.10.31234/osf.io/xc4sv, in-progress, https://psyarxiv.com/xc4sv/.

Siller, M., & Sigman, M. (2008). Modeling longitudinal change in the language abilities of children with autism: Parent behaviors and child characteristics as predictors of change. *Developmental Psychology*, *44*, 1691–1704.
Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. London: Oxford University Press.
Sivo, S. A., Fan, X., & Witta, L. (2005). The biasing effects of unmodeled ARMA time series processes on latent growth curve model estimates. *Structural Equation Modeling*, *12*, 215–231.
Snijders, T. A. B., & Berkhof, J. (2008). Diagnostic checks for multilevel models. In J. de Leeuw & E. Meijer (Eds.), *Handbook of multilevel analysis* (pp. 141–176). New York: Springer.
Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling* (2nd ed.). London: Sage.
Sterba, S. K. (2014). Fitting nonlinear latent growth curve models with individually varying time points. *Structural Equation Modeling*, *21*, 630–647.
Wolfinger, R. (1993). Covariance structure selection in general mixed models. *Communications in Statistics, Simulation and Computation*, *22*, 1079–1106.
Xu, R. H. (2003). Measuring explained variation in linear mixed effects models. *Statistics in Medicine*, *22*, 3527–3541.
Vonesh, E. F., & Chinchilli, V. M. (1997). *Linear and nonlinear models for the analysis of repeated measurements*. New York: Marcel Dekker.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Rights, J. D., & Sterba, S. K. (2021). Effect size measures for longitudinal growth analyses: Extending a framework of multilevel model R-squareds to accommodate heteroscedasticity, autocorrelation, nonlinearity, and alternative centering strategies. *New Directions for Child and Adolescent Development*, 2021, 65–110. https://doi.org/10.1002/cad.20387

APPENDIX A: REVIEW FROM RIGHTS AND STERBA (2019) OF DEFINITIONS OF R-SQUARED MEASURES AND CORRESPONDING FORMULAS, ASSUMING PERSON-MEAN-CENTERING

Here we briefly review the Rights and Sterba (2019) decomposition of model-implied outcome variance for cluster-mean-centered models, framing this decomposition in terms of longitudinal models wherein observations are nested within persons (and hence clusters = persons). We will show how this decomposition is used to form each R-squared measure. Importantly, this decomposition here assumes that the level-1 errors are homoscedastic and have no autocorrelation—assumptions we later relax in the current paper.

First note that a two-level person-mean-centered model can be expressed generally as

$$\begin{gathered} y_{ij} = \mathbf{x}'^{w}_{ij}\boldsymbol{\gamma}^{w} + \mathbf{x}'^{b}_{j}\boldsymbol{\gamma}^{b} + \mathbf{w}'_{ij}\mathbf{u}_j + e_{ij}, \\ \mathbf{u}_j \sim MVN(\mathbf{0}, \mathbf{T}), \\ e_{ij} \sim N(0, \sigma^2), \end{gathered} \tag{A1}$$

where $\mathbf{x}^{w}_{ij}$ denotes a vector of all level-1 predictors (each person-mean-centered), $\boldsymbol{\gamma}^{w}$ a vector of fixed components of slopes corresponding to elements in $\mathbf{x}^{w}_{ij}$, $\mathbf{x}^{b}_{j}$ a vector of 1 and all level-2 predictors, $\boldsymbol{\gamma}^b$ a vector of fixed components of slopes corresponding to elements in $\mathbf{x}_j^b$, $\mathbf{w}_{ij}$ a vector with the first element equal to 1 and all subsequent elements being predictors with random slopes, $\mathbf{u}_j$ a vector of random effect errors (with covariance matrix $\mathbf{T}$) corresponding to the elements in $\mathbf{w}_{ij}$, and e_{ij} the level-1 error. The level-1 errors are presently assumed uncorrelated with homoscedastic variance σ^2. The model-implied variance for the person-mean-centered model expression, derived in Rights and Sterba (2019), is then given as

$$\begin{aligned} \text{var}\left(y_{ij}\right) &= \text{var}\left(\mathbf{x'}_{ij}^{w}\boldsymbol{\gamma}^{w} + \mathbf{x'}_{j}^{b}\boldsymbol{\gamma}^{b} + \mathbf{w'}_{ij}\mathbf{u}_j + e_{ij}\right) \\ &= \boldsymbol{\gamma}^{w\prime}\boldsymbol{\Phi}^{w}\boldsymbol{\gamma}^{w} + \boldsymbol{\gamma}^{b\prime}\boldsymbol{\Phi}^{b}\boldsymbol{\gamma}^{b} + tr\left(\mathbf{T\Sigma}\right) + \tau_{00} + \sigma^2, \end{aligned} \tag{A2}$$

where $\boldsymbol{\Phi}^w$, $\boldsymbol{\Phi}^b$, and $\boldsymbol{\Sigma}$ denote covariance matrices of elements of $\mathbf{x}_{ij}^w$, $\mathbf{x}_j^b$, and $\mathbf{w'}_{ij}$, respectively, and τ_{00} denotes the random intercept variance. The five separate terms in Equation (A2) each denote variance attributable to a distinct source: $\boldsymbol{\gamma}^{w\prime}\boldsymbol{\Phi}^{w}\boldsymbol{\gamma}^{w}$ denotes variance attributable to level-1 predictors via fixed components of slopes, $\boldsymbol{\gamma}^{b\prime}\boldsymbol{\Phi}^{b}\boldsymbol{\gamma}^{b}$ variance attributable to level-2 predictors via fixed components of slopes, $tr(\mathbf{T\Sigma})$ variance attributable to level-1 predictors via random slope variance, τ_{00} variance attributable to cluster-specific outcome means via random intercept variation, and σ^2 variance attributable to level-1 errors. Three of these variances ($\boldsymbol{\gamma}^{w\prime}\boldsymbol{\Phi}^{w}\boldsymbol{\gamma}^{w}$, $tr(\mathbf{T\Sigma})$, and σ^2) reflect purely within-person variation, whereas the other two ($\boldsymbol{\gamma}^{b\prime}\boldsymbol{\Phi}^{b}\boldsymbol{\gamma}^{b}$ and τ_{00}) reflect purely between-person variation.

From this decomposition, we can compute the total single-source R-squared measures, quantifying total variance explained by one source at a time, for person-mean-centered models (defined in Table 3 Column 2) as:

$$\begin{aligned} R_t^{2(f_1)} &= \frac{\boldsymbol{\gamma}^{w\prime}\boldsymbol{\Phi}^{w}\boldsymbol{\gamma}^{w}}{\boldsymbol{\gamma}^{w\prime}\boldsymbol{\Phi}^{w}\boldsymbol{\gamma}^{w} + \boldsymbol{\gamma}^{b\prime}\boldsymbol{\Phi}^{b}\boldsymbol{\gamma}^{b} + tr\left(\mathbf{T\Sigma}\right) + \tau_{00} + \sigma^2}, \\ R_t^{2(f_2)} &= \frac{\boldsymbol{\gamma}^{b\prime}\boldsymbol{\Phi}^{b}\boldsymbol{\gamma}^{b}}{\boldsymbol{\gamma}^{w\prime}\boldsymbol{\Phi}^{w}\boldsymbol{\gamma}^{w} + \boldsymbol{\gamma}^{b\prime}\boldsymbol{\Phi}^{b}\boldsymbol{\gamma}^{b} + tr\left(\mathbf{T\Sigma}\right) + \tau_{00} + \sigma^2}, \\ R_t^{2(v)} &= \frac{tr\left(\mathbf{T\Sigma}\right)}{\boldsymbol{\gamma}^{w\prime}\boldsymbol{\Phi}^{w}\boldsymbol{\gamma}^{w} + \boldsymbol{\gamma}^{b\prime}\boldsymbol{\Phi}^{b}\boldsymbol{\gamma}^{b} + tr\left(\mathbf{T\Sigma}\right) + \tau_{00} + \sigma^2}, \\ R_t^{2(m)} &= \frac{\tau_{00}}{\boldsymbol{\gamma}^{w\prime}\boldsymbol{\Phi}^{w}\boldsymbol{\gamma}^{w} + \boldsymbol{\gamma}^{b\prime}\boldsymbol{\Phi}^{b}\boldsymbol{\gamma}^{b} + tr\left(\mathbf{T\Sigma}\right) + \tau_{00} + \sigma^2}. \end{aligned} \tag{A3}$$

The total combination-source R-squared measures, quantifying total variance explained by multiple sources together, are then just combinations of these equations, for instance: $R_t^{2(f)} = R_t^{2(f_1)} + R_t^{2(f_2)}$,$R_t^{2(fv)} = R_t^{2(f_1)} + R_t^{2(f_2)} + R_t^{2(v)}$, and $R_t^{2(fvm)} = R_t^{2(f_1)} + R_t^{2(f_2)} + R_t^{2(v)} + R_t^{2(m)}$. The single-source within-person measures for person-mean-centered models are then given as

$$\begin{aligned} R_w^{2(f_1)} &= \frac{\boldsymbol{\gamma}^{w\prime}\boldsymbol{\Phi}^{w}\boldsymbol{\gamma}^{w}}{\boldsymbol{\gamma}^{w\prime}\boldsymbol{\Phi}^{w}\boldsymbol{\gamma}^{w} + tr\left(\mathbf{T\Sigma}\right) + \sigma^2}, \\ R_w^{2(v)} &= \frac{tr\left(\mathbf{T\Sigma}\right)}{\boldsymbol{\gamma}^{w\prime}\boldsymbol{\Phi}^{w}\boldsymbol{\gamma}^{w} + tr\left(\mathbf{T\Sigma}\right) + \sigma^2}. \end{aligned} \tag{A4}$$

A combination-source within-person measure is given as $R_w^{2(f_1 v)} = R_w^{2(f_1)} + R_w^{2(v)}$. The single-source between-person measures for person-mean-centered models are likewise:

$$R_b^{2(f_2)} = \frac{\boldsymbol{\gamma}^{b\prime}\boldsymbol{\Phi}^b\boldsymbol{\gamma}^b}{\boldsymbol{\gamma}^{b\prime}\boldsymbol{\Phi}^b\boldsymbol{\gamma}^b + \tau_{00}},$$

$$R_b^{2(m)} = \frac{\tau_{00}}{\boldsymbol{\gamma}^{b\prime}\boldsymbol{\Phi}^b\boldsymbol{\gamma}^b + \tau_{00}}.$$

APPENDIX B: NEW EXTENSION: FULL DECOMPOSITION OF OUTCOME VARIANCE FOR *NON*-PERSON-MEAN-CENTERED MODELS

In Appendix B, we derive a complete model-implied variance decomposition for *non*-person-mean-centered models (i.e., models involving uncentered level-1 predictors or involving centering level-1 predictors at a constant value such as the first assessment value or the grand mean). Here we show how this variance is a sum of variances attributable to each of the following distinct sources:

- the within-person-varying portion of level-1 predictors via fixed components of slopes (f_1);
- the between-person-varying portion of level-1 and/or level-2 predictors via fixed components of slopes (f_2);
- the within-person-varying portion of level-1 predictors via random slope variation (v_1);
- the between-person-varying portion of level-1 predictors via random slope variation (v_2);
- person-specific outcome means via random intercept variation at the mean of all predictors with random slopes (m);
- level-1 errors.

From this decomposition, we show how each of the R-squared measures described in Table 3 Column 3 are computed. Rights and Sterba (2019) had provided a more limited decomposition for non-person-mean-centered models which did not break down f into f_1 (reflecting purely within-cluster variance) and f_2 (reflecting purely between-cluster variance), and did not break down v into v_1 and v_2.

An expression for a two-level multilevel model without assuming person-mean-centering is:

$$y_{ij} = \mathbf{x}_{ij}'\boldsymbol{\gamma} + \mathbf{w}_{ij}'\mathbf{u}_j + e_{ij},$$

$$\mathbf{u}_j \sim MVN(\mathbf{0}, \mathbf{T}),$$

$$e_{ij} \sim N\left(0, \sigma^2\right). \tag{B1}$$

Here, y_{ij} denotes the outcome for observation i nested within cluster j, $\mathbf{x}_{ij}$ a vector with the first element equal to 1 and all subsequent elements being predictors for observation i within person j, $\boldsymbol{\gamma}$ a vector of fixed components of coefficients corresponding to the elements in $\mathbf{x}_{ij}$, $\mathbf{w}_{ij}$ a vector with the first element equal to 1 and all subsequent elements being predictors with random slopes, $\mathbf{u}_j$ a vector of random effect errors (with covariance matrix $\mathbf{T}$) corresponding to the elements in $\mathbf{w}_{ij}$, and e_{ij} the level-1 error. The level-1 error covariance matrix is presently assumed diagonal with homoscedastic variance σ^2.

First note that we can decompose every predictor into a purely *within-person-varying portion* and a purely *between-person-varying portion* by using the following substitutions:

$$\begin{aligned}\mathbf{x}_{ij} &= (\mathbf{x}_{ij} - \mathbf{x}_{.j}) + \mathbf{x}_{.j},\\ \mathbf{w}_{ij} &= (\mathbf{w}_{ij} - \mathbf{w}_{.j}) + \mathbf{w}_{.j}.\end{aligned} \tag{B2}$$

Here, $\mathbf{x}_{.j}$ and $\mathbf{w}_{.j}$ denote vectors of person means of each element of $\mathbf{x}_{ij}$ and $\mathbf{w}_{ij}$, respectively. Hence, $(\mathbf{x}_{ij} - \mathbf{x}_{.j})$ reflects a vector of variables that are deviations from the person-specific means, and is thus the portion of $\mathbf{x}_{ij}$ that varies exclusively within-person (since $E_{i|j}[\mathbf{x}_{ij} - \mathbf{x}_{.j}] = 0$ for all clusters, and hence $\text{var}_j[\mathbf{x}_{ij} - \mathbf{x}_{.j}] = 0$). Similarly, $(\mathbf{w}_{ij} - \mathbf{w}_{.j})$ is the portion of $\mathbf{w}_{ij}$ that varies exclusively within-person. The parts of $\mathbf{x}_{ij}$ and $\mathbf{w}_{ij}$ that vary exclusively between-person then are $\mathbf{x}_{.j}$ and $\mathbf{w}_{.j}$, respectively (since $\mathbf{x}_{.j}$ and $\mathbf{w}_{.j}$ are vectors of constants for each cluster, i.e., $\text{var}_{i|j}[\mathbf{x}_{.j}] = \text{var}_{i|j}[\mathbf{w}_{.j}] = 0$).

We can then re-write the model expression in Equation (A1) as

$$\begin{aligned}y_{ij} &= \mathbf{x}'_{ij}\boldsymbol{\gamma} + \mathbf{w}'_{ij}\mathbf{u}_j + e_{ij}\\ &= \left(\mathbf{x}_{ij} - \mathbf{x}_{.j} + \mathbf{x}_{.j}\right)'\boldsymbol{\gamma} + \left(\mathbf{w}_{ij} - \mathbf{w}_{.j} + \mathbf{w}_{.j}\right)'\mathbf{u}_j + e_{ij}\\ &= \left(\mathbf{x}_{ij} - \mathbf{x}_{.j}\right)'\boldsymbol{\gamma} + \mathbf{x}'_{.j}\boldsymbol{\gamma} + \left(\mathbf{w}_{ij} - \mathbf{w}_{.j}\right)'\mathbf{u}_j + \mathbf{w}'_{.j}\mathbf{u}_j + e_{ij}.\end{aligned} \tag{B3}$$

We can then compute the model-implied variance as

$$\begin{aligned}\text{var}\left(y_{ij}\right) &= \text{var}\left(\left(\mathbf{x}_{ij} - \mathbf{x}_{.j}\right)'\boldsymbol{\gamma} + \mathbf{x}'_{.j}\boldsymbol{\gamma} + \left(\mathbf{w}_{ij} - \mathbf{w}_{.j}\right)'\mathbf{u}_j + \mathbf{w}'_{.j}\mathbf{u}_j + e_{ij}\right)\\ &= \text{var}\left(\left(\mathbf{x}_{ij} - \mathbf{x}_{.j}\right)'\boldsymbol{\gamma}\right) + \text{var}\left(\mathbf{x}'_{.j}\boldsymbol{\gamma}\right) + \text{var}\left(\left(\mathbf{w}_{ij} - \mathbf{w}_{.j}\right)'\mathbf{u}_j\right) + \text{var}\left(\mathbf{w}'_{.j}\mathbf{u}_j\right) + \text{var}\left(e_{ij}\right).\end{aligned} \tag{B4}$$

The five variances in the second line of Equation (B4) are separable because of the lack of covariance between the following pairs: the fixed components and random components, the purely within-cluster-varying portion of predictors and the purely between-cluster-varying portion, and the level-1 errors and all other terms. The first part of Equation (B4) is computed as

$$\text{var}\left(\left(\mathbf{x}_{ij} - \mathbf{x}_{.j}\right)'\boldsymbol{\gamma}\right) = \boldsymbol{\gamma}'\boldsymbol{\Phi}_w\boldsymbol{\gamma}, \tag{B5}$$

where $\boldsymbol{\Phi}_w$ is the covariance matrix of the within-cluster-varying portions of $\mathbf{x}_{ij}$. The second part of Equation (B4) is computed as

$$\text{var}\left(\mathbf{x}'_{.j}\boldsymbol{\gamma}\right) = \boldsymbol{\gamma}'\boldsymbol{\Phi}_b\boldsymbol{\gamma}, \tag{B6}$$

where $\boldsymbol{\Phi}_b$ is the covariance matrix of the between-cluster-varying portions of $\mathbf{x}_{ij}$. The third part of Equation (B4) is computed using the law of total variance as

$$\text{var}\left(\left(\mathbf{w}_{ij} - \mathbf{w}_{.j}\right)'\mathbf{u}_j\right) = E\left[\text{var}\left(\left(\mathbf{w}_{ij} - \mathbf{w}_{.j}\right)'\mathbf{u}_j|\mathbf{u}_j\right)\right] + \text{var}\left(E\left[\left(\mathbf{w}_{ij} - \mathbf{w}_{.j}\right)'\mathbf{u}_j|\mathbf{u}_j\right]\right)$$

$$
\begin{aligned}
&= E\left[\mathbf{u}_j'\boldsymbol{\Sigma}_w\mathbf{u}_j\right] + \text{var}\left(E\left[\left(\mathbf{w}_{ij} - \mathbf{w}_{\cdot j}\right)'\right]\mathbf{u}_j\right) \\
&= E\left[tr\left(\mathbf{u}_j'\boldsymbol{\Sigma}_w\mathbf{u}_j\right)\right] + \text{var}\,(0) \\
&= E\left[tr\left(\mathbf{u}_j\mathbf{u}_j'\boldsymbol{\Sigma}_w\right)\right] \\
&= tr\left(E\left[\mathbf{u}_j\mathbf{u}_j'\right]\boldsymbol{\Sigma}_w\right) \\
&= tr\left(\mathbf{T}\boldsymbol{\Sigma}_w\right), \qquad \text{(B7)}
\end{aligned}
$$

where $\boldsymbol{\Sigma}_w$ is the covariance matrix of the within-cluster-varying portions of $\mathbf{w}_{ij}$. The fourth part of Equation (B4) is computed, again using the law of total variance, as

$$
\begin{aligned}
\text{var}\left(\left(\mathbf{w}_j\right)'\mathbf{u}_j\right) &= E\left[\text{var}\left(\mathbf{w}_{\cdot j}'\mathbf{u}_j|\mathbf{u}_j\right)\right] + \text{var}\left(E\left[\mathbf{w}_{\cdot j}'\mathbf{u}_j|\mathbf{u}_j\right]\right) \\
&= E\left[\mathbf{u}_j'\boldsymbol{\Sigma}_b\mathbf{u}_j\right] + \text{var}\left(E\left[\mathbf{w}_{\cdot j}'\right]\mathbf{u}_j\right) \\
&= E\left[tr\left(\mathbf{u}_j'\boldsymbol{\Sigma}_b\mathbf{u}_j\right)\right] + \text{var}\left(\mathbf{m}_b\mathbf{u}_j\right) \\
&= E\left[tr\left(\mathbf{u}_j\mathbf{u}'_j\boldsymbol{\Sigma}_b\right)\right] + \mathbf{m}_b'\text{var}\left(\mathbf{u}_j\right)\mathbf{m}_b \\
&= tr\left(E\left[\mathbf{u}_j\mathbf{u}_j'\right]\boldsymbol{\Sigma}_b\right) + \mathbf{m}_b'\mathbf{T}\mathbf{m}_b \\
&= tr\left(\mathbf{T}\boldsymbol{\Sigma}_b\right) + \mathbf{m}_b'\mathbf{T}\mathbf{m}_b, \qquad \text{(B8)}
\end{aligned}
$$

where $\boldsymbol{\Sigma}_b$ is the covariance matrix of the between-cluster-varying portions of $\mathbf{w}_{ij}$ and $\mathbf{m}_b$ is a vector containing the means of all elements of $\mathbf{w}_{ij}$. The fifth part of Equation (B4) is then simply

$$
\text{var}\left(e_{ij}\right) = \sigma^2. \qquad \text{(B9)}
$$

Thus, the total model-implied outcome variance is

$$
\text{var}\left(y_{ij}\right) = \boldsymbol{\gamma}'\boldsymbol{\Phi}_w\boldsymbol{\gamma} + \boldsymbol{\gamma}'\boldsymbol{\Phi}_b\boldsymbol{\gamma} + tr\left(\mathbf{T}\boldsymbol{\Sigma}_w\right) + tr\left(\mathbf{T}\boldsymbol{\Sigma}_b\right) + \mathbf{m}'\mathbf{T}\mathbf{m} + \sigma^2. \qquad \text{(B10)}
$$

These six distinct variances in Equation (B10) denote the variance attributed, in order, to each source listed in bullet points at the beginning of the Appendix B section. (Later, in Appendix G, we show how this expression is modified for heteroscedastic models by replacing σ^2 with $E[\sigma^2]$.)

APPENDIX C: NEW EXTENSION: FULL SET OF DEFINITIONS OF R-SQUARED MEASURES AND CORRESPONDING FORMULAS UNDER *NON*-PERSON-MEAN-CENTERING

From the decomposition of outcome variance provided in Appendix B, we can compute the total single-source R-squared measures for *non*-person-mean-centered models (i.e., models involving uncentered level-1 predictors or involving centering level-1 predictors at a constant value such as the first assessment value or the grand mean) as

$$\begin{aligned}
R_t^{2(f_1)} &= \frac{\boldsymbol{\gamma}'\boldsymbol{\Phi}_w\boldsymbol{\gamma}}{\boldsymbol{\gamma}'\boldsymbol{\Phi}_w\boldsymbol{\gamma} + \boldsymbol{\gamma}'\boldsymbol{\Phi}_b\boldsymbol{\gamma} + tr\left(\mathbf{T}\boldsymbol{\Sigma}_w\right) + tr\left(\mathbf{T}\boldsymbol{\Sigma}_b\right) + \mathbf{m}'\mathbf{T}\mathbf{m} + \sigma^2},\\
R_t^{2(f_2)} &= \frac{\boldsymbol{\gamma}'\boldsymbol{\Phi}_b\boldsymbol{\gamma}}{\boldsymbol{\gamma}'\boldsymbol{\Phi}_w\boldsymbol{\gamma} + \boldsymbol{\gamma}'\boldsymbol{\Phi}_b\boldsymbol{\gamma} + tr\left(\mathbf{T}\boldsymbol{\Sigma}_w\right) + tr\left(\mathbf{T}\boldsymbol{\Sigma}_b\right) + \mathbf{m}'\mathbf{T}\mathbf{m} + \sigma^2},\\
R_t^{2(v_1)} &= \frac{tr\left(\mathbf{T}\boldsymbol{\Sigma}_w\right)}{\boldsymbol{\gamma}'\boldsymbol{\Phi}_w\boldsymbol{\gamma} + \boldsymbol{\gamma}'\boldsymbol{\Phi}_b\boldsymbol{\gamma} + tr\left(\mathbf{T}\boldsymbol{\Sigma}_w\right) + tr\left(\mathbf{T}\boldsymbol{\Sigma}_b\right) + \mathbf{m}'\mathbf{T}\mathbf{m} + \sigma^2},\\
R_t^{2(v_2)} &= \frac{tr\left(\mathbf{T}\boldsymbol{\Sigma}_b\right)}{\boldsymbol{\gamma}'\boldsymbol{\Phi}_w\boldsymbol{\gamma} + \boldsymbol{\gamma}'\boldsymbol{\Phi}_b\boldsymbol{\gamma} + tr\left(\mathbf{T}\boldsymbol{\Sigma}_w\right) + tr\left(\mathbf{T}\boldsymbol{\Sigma}_b\right) + \mathbf{m}'\mathbf{T}\mathbf{m} + \sigma^2},\\
R_t^{2(m)} &= \frac{\mathbf{m}'\mathbf{T}\mathbf{m}}{\boldsymbol{\gamma}'\boldsymbol{\Phi}_w\boldsymbol{\gamma} + \boldsymbol{\gamma}'\boldsymbol{\Phi}_b\boldsymbol{\gamma} + tr\left(\mathbf{T}\boldsymbol{\Sigma}_w\right) + tr\left(\mathbf{T}\boldsymbol{\Sigma}_b\right) + \mathbf{m}'\mathbf{T}\mathbf{m} + \sigma^2}.
\end{aligned} \tag{C1}$$

Example combination-source total R-squared measures that could be constructed are: $R_t^{2(f)} = R_t^{2(f_1)} + R_t^{2(f_2)}$, $R_t^{2(fv)} = R_t^{2(f_1)} + R_t^{2(f_2)} + R_t^{2(v_1)} + R_t^{2(v_2)}$, and $R_t^{2(fvm)} = R_t^{2(f_1)} + R_t^{2(f_2)} + R_t^{2(v_1)} + R_t^{2(v_2)} + R_t^{2(m)}$.

The single-source within-person measures are then

$$\begin{aligned}
R_w^{2(f_1)} &= \frac{\boldsymbol{\gamma}'\boldsymbol{\Phi}_w\boldsymbol{\gamma}}{\boldsymbol{\gamma}'\boldsymbol{\Phi}_w\boldsymbol{\gamma} + tr(\mathbf{T}\boldsymbol{\Sigma}_w) + \sigma^2},\\
R_w^{2(v_1)} &= \frac{tr(\mathbf{T}\boldsymbol{\Sigma}_w)}{\boldsymbol{\gamma}'\boldsymbol{\Phi}_w\boldsymbol{\gamma} + tr(\mathbf{T}\boldsymbol{\Sigma}_w) + \sigma^2}.
\end{aligned} \tag{C2}$$

Summing these yields a combination-source within-person measure: $R_w^{2(f_1 v_1)} = R_w^{2(f_1)} + R_w^{2(v_1)}$.

The single-source between-person measures are:

$$\begin{aligned}
R_b^{2(f_2)} &= \frac{\boldsymbol{\gamma}'\boldsymbol{\Phi}_b\boldsymbol{\gamma}}{\boldsymbol{\gamma}'\boldsymbol{\Phi}_b\boldsymbol{\gamma} + tr(\mathbf{T}\boldsymbol{\Sigma}_b) + \mathbf{m}'\mathbf{T}\mathbf{m}},\\
R_b^{2(v_2)} &= \frac{tr(\mathbf{T}\boldsymbol{\Sigma}_b)}{\boldsymbol{\gamma}'\boldsymbol{\Phi}_b\boldsymbol{\gamma} + tr(\mathbf{T}\boldsymbol{\Sigma}_b) + \mathbf{m}'\mathbf{T}\mathbf{m}},\\
R_b^{2(v_2)} &= \frac{\mathbf{m}'\mathbf{T}\mathbf{m}}{\boldsymbol{\gamma}'\boldsymbol{\Phi}_b\boldsymbol{\gamma} + tr(\mathbf{T}\boldsymbol{\Sigma}_b) + \mathbf{m}'\mathbf{T}\mathbf{m}}.
\end{aligned} \tag{C3}$$

Summing these yields a combination-source between-person measure: $R_b^{2(f_2 v_2)} = R_b^{2(f_2)} + R_b^{2(v_2)}$.

Previous work (Rights & Sterba, 2019) had provided a more limited set of measures for non-person-mean-centered models compared to those given above (i.e., previous work only provided total measures, not level-specific measures, for non-person-mean-centered models).

APPENDIX D: PROOF THAT THE PROPORTION OF VARIANCE ATTRIBUTABLE TO SOURCE *m*, DEFINED IN TABLE 2, DOES NOT CHANGE WHEN CENTERING PREDICTORS BY A CONSTANT, REGARDLESS OF THE CHOSEN CENTERING CONSTANT'S VALUE

In Appendix D, we show that, in the population, the following are invariant to centering predictors by a constant value: the proportion of total variance attributable to predictors via fixed components of slopes, the proportion of total variance attributable to predictors via random slope variation, and the proportion of variance attributable to source m (defined separately for cluster-mean-centered and non-cluster-mean-centered in Table 2).

We start with the unconditional linear growth model defined in Equation (1), written here in reduced form:

$$\begin{gathered} y_{ij} = \gamma_{00} + u_{0j} + \gamma_{10}x_{ij} + u_{1j}x_{ij} + e_{ij}, \\ \begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix} \sim MVN\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \tau_{00} & \\ \tau_{01} & \tau_{11} \end{bmatrix} \right), \\ e_{ij} \sim N\left(0, \sigma^2\right). \end{gathered} \tag{D1}$$

For generality, we denote y as the outcome (e.g., self-efficacy) and x as the level-1 predictor (e.g., *time*). We will compare this to a model that centers x by an arbitrary constant, a. We will show that the aforementioned proportions will always be the same as those obtained from the uncentered model, regardless of the value of a. The centered-by-a model is thus given as

$$y_{ij} = \gamma^*_{00} + u^*_{0j} + \gamma^*_{10}\left(x_{ij} - a\right) + u^*_{1j}\left(x_{ij} - a\right) + e^*_{ij}. \tag{D2}$$

We use asterisks to denote terms and parameters from the centered-by-a model. It is well-established that these two models (Equation D1 and D2) are equivalent models in that their likelihoods are maximized at the same value and they generate the same set of expectations and dispersions (Kreft, de Leeuw, & Aiken 1995). As such, we can write each component of the centered-by-a model in terms of the components of the uncentered model by rearranging terms like so:

$$\begin{aligned} y_{ij} &= \gamma^*_{00} + u^*_{0j} + \gamma^*_{10}\left(x_{ij} - a\right) + u^*_{1j}\left(x_{ij} - a\right) + e^*_{ij} \\ &= \gamma^*_{00} + u^*_{0j} + \gamma^*_{10}x_{ij} - \gamma^*_{10}a + u^*_{1j}x_{ij} - u^*_{1j}a + e^*_{ij} \\ &= \left(\gamma^*_{00} - \gamma^*_{10}a\right) + \left(u^*_{0j} - u^*_{1j}a\right) + \gamma^*_{10}x_{ij} + u^*_{1j}x_{ij} + e^*_{ij}. \end{aligned} \tag{D3}$$

This re-expression highlights the following equivalencies between the two models:

$$\begin{aligned} \gamma_{00} &= \gamma^*_{00} - \gamma^*_{10}a, \\ u_{0j} &= u^*_{0j} - u^*_{1j}a, \\ \gamma_{10} &= \gamma^*_{10}, \\ u_{1j} &= u^*_{1j}, \\ e_{ij} &= e^*_{ij}. \end{aligned} \tag{D4}$$

Hence, we can rewrite the centered-by-a model using terms from the uncentered model as

$$y_{ij} = (\gamma_{00} + \gamma_{10}a) + (u_{0j} + u_{1j}a) + \gamma_{10}x_{ij} + u_{1j}x_{ij} + e_{ij}. \tag{D5}$$

The variance component from the centered-by-a model can then be written as

$$\begin{aligned}
\text{var}\left(u^*_{0j}\right) &= \text{var}\left(u_{0j} + u_{1j}a\right) \\
&= \text{var}\left(u_{0j}\right) + \text{var}\left(u_{1j}a\right) + 2\text{cov}\left(u_{0j}, u_{1j}a\right) \\
&= \tau_{00} + a^2\tau_{11} + 2a\tau_{01}, \\
\text{var}\left(u^*_{1j}\right) &= \text{var}\left(u_{1j}\right) \\
&= \tau_{11}, \\
\text{cov}\left(u^*_{0j}, u_{1j}\right) &= \text{cov}\left(u_{0j} + u_{1j}a, u_{1j}\right) \\
&= \text{cov}\left(u_{0j}, u_{1j}\right) + \text{cov}\left(u_{1j}a, u_{1j}\right) \\
&= \tau_{01} + a\tau_{11}.
\end{aligned} \tag{D6}$$

Using the formulas outlined in Appendix B and in Rights and Sterba (2019), we can then compute the total variance attributable to predictors via fixed components in the centered-by-a model as:

$$\begin{aligned}
\boldsymbol{\gamma}^{*\prime}\boldsymbol{\Phi}^*\boldsymbol{\gamma}^* &= \begin{bmatrix}\gamma_{00} + \gamma_{10}a & \gamma_{10}\end{bmatrix} \begin{bmatrix}0 & 0 \\ 0 & \text{var}(x_{ij})\end{bmatrix} \begin{bmatrix}\gamma_{00} + \gamma_{10}a \\ \gamma_{10}\end{bmatrix} \\
&= \begin{bmatrix}0 & \gamma_{10}\text{var}(x_{ij})\end{bmatrix} \begin{bmatrix}\gamma_{00} + \gamma_{10}a \\ \gamma_{10}\end{bmatrix} \\
&= \gamma_{10}^2\text{var}(x_{ij}).
\end{aligned} \tag{D7}$$

And we see that this is exactly equal to that obtained from the uncentered model:

$$\begin{aligned}
\boldsymbol{\gamma}'\Phi\gamma &= \begin{bmatrix}\gamma_{00} & \gamma_{10}\end{bmatrix} \begin{bmatrix}0 & 0 \\ 0 & \text{var}(x_{ij})\end{bmatrix} \begin{bmatrix}\gamma_{00} \\ \gamma_{10}\end{bmatrix} \\
&= \begin{bmatrix}0 & \gamma_{10}\text{var}(x_{ij})\end{bmatrix} \begin{bmatrix}\gamma_{00} \\ \gamma_{10}\end{bmatrix} \\
&= \gamma_{10}^2\text{var}(x_{ij}).
\end{aligned} \tag{D8}$$

We can additionally compute the total variance attributable to predictors via random slope variation in the centered-by-a model as:

$$\begin{aligned}
tr(\mathbf{T}^*\boldsymbol{\Sigma}^*) &= tr\left(\begin{bmatrix}\tau_{00} + a^2\tau_{11} + 2a\tau_{01} & \tau_{01} + a\tau_{11} \\ \tau_{01} + a\tau_{11} & \tau_{11}\end{bmatrix} \begin{bmatrix}0 & 0 \\ 0 & \text{var}(x_{ij})\end{bmatrix}\right) \\
&= tr\left(\begin{bmatrix}0 & \tau_{01}\text{var}(x_{ij}) + a\tau_{11}\text{var}(x_{ij}) \\ 0 & \tau_{11}\text{var}(x_{ij})\end{bmatrix}\right) \\
&= \tau_{11}\text{var}(x_{ij}).
\end{aligned} \tag{D9}$$

And we again see that this is exactly equal to that obtained from the uncentered model:

$$
\begin{aligned}
tr(\mathbf{T}\Sigma) &= tr\left(\begin{bmatrix}\tau_{00} & \tau_{01}\\ \tau_{01} & \tau_{11}\end{bmatrix}\begin{bmatrix}0 & 0\\ 0 & \mathrm{var}(x_{ij})\end{bmatrix}\right)\\
&= tr\left(\begin{bmatrix}0 & \tau_{01}\mathrm{var}(x_{ij})\\ 0 & \tau_{11}\mathrm{var}(x_{ij})\end{bmatrix}\right)\\
&= \tau_{11}\mathrm{var}(x_{ij}).
\end{aligned}
\tag{D10}
$$

Lastly, we compute the total variance attributable to source m in the centered-by-a model as:

$$
\begin{aligned}
\mathbf{m}^{*\prime}\mathbf{T}^{*}\mathbf{m}^{*} &= \begin{bmatrix}1 & E\left[x_{ij}-a\right]\end{bmatrix}\begin{bmatrix}\tau_{00}+a^2\tau_{11}+2a\tau_{01} & \tau_{01}+a\tau_{11}\\ \tau_{01}+a\tau_{11} & \tau_{11}\end{bmatrix}\begin{bmatrix}1\\ E\left[x_{ij}-a\right]\end{bmatrix}\\
&= \begin{bmatrix}1 & E\left[x_{ij}\right]-a\end{bmatrix}\begin{bmatrix}\tau_{00}+a^2\tau_{11}+2a\tau_{01} & \tau_{01}+a\tau_{11}\\ \tau_{01}+a\tau_{11} & \tau_{11}\end{bmatrix}\begin{bmatrix}1\\ E\left[x_{ij}\right]-a\end{bmatrix}\\
&= \begin{bmatrix}\tau_{00}+a^2\tau_{11}+2a\tau_{01}+\left(E\left[x_{ij}\right]-a\right)\left(\tau_{01}+a\tau_{11}\right) & \tau_{01}+a\tau_{11}+\left(E\left[x_{ij}\right]-a\right)\tau_{11}\end{bmatrix}\\
&\quad \begin{bmatrix}1\\ E\left[x_{ij}\right]-a\end{bmatrix}\\
&= \tau_{00}+a^2\tau_{11}+2a\tau_{01}+\left(E\left[x_{ij}\right]-a\right)\left(\tau_{01}+a\tau_{11}\right)+\left(E\left[x_{ij}\right]-a\right)(\tau_{01}\\
&\quad +a\tau_{11})+\left(E\left[x_{ij}\right]-a\right)^2\tau_{11}\\
&= \tau_{00}+a^2\tau_{11}+2a\tau_{01}+2\left(E\left[x_{ij}\right]-a\right)\left(\tau_{01}+a\tau_{11}\right)+E\left[x_{ij}\right]^2\tau_{11}\\
&\quad +a^2\tau_{11}-2aE\left[x_{ij}\right]\tau_{11}\\
&= \tau_{00}+a^2\tau_{11}+2a\tau_{01}+2E\left[x_{ij}\right]\tau_{01}-2a\tau_{01}-2a^2\tau_{11}+2aE\left[x_{ij}\right]\tau_{11}+E\left[x_{ij}\right]^2\tau_{11}\\
&\quad +a^2\tau_{11}-2aE\left[x_{ij}\right]\tau_{11}\\
&= \tau_{00}+2E\left[x_{ij}\right]\tau_{01}+E\left[x_{ij}\right]^2\tau_{11}.
\end{aligned}
\tag{D11}
$$

And we see that this is exactly equal to that obtained from the uncentered model:

$$
\begin{aligned}
\mathbf{m}'\mathbf{T}\mathbf{m} &= \begin{bmatrix}1 & E[x_{ij}]\end{bmatrix}\begin{bmatrix}\tau_{00} & \tau_{01}\\ \tau_{01} & \tau_{11}\end{bmatrix}\begin{bmatrix}1\\ E[x_{ij}]\end{bmatrix}\\
&= \begin{bmatrix}\tau_{00}+E[x_{ij}]\tau_{01} & \tau_{01}+E[x_{ij}]\tau_{11}\end{bmatrix}\begin{bmatrix}1\\ E[x_{ij}]\end{bmatrix}\\
&= \tau_{00}+2E[x_{ij}]\tau_{01}+E[x_{ij}]^2\tau_{11}.
\end{aligned}
\tag{D12}
$$

APPENDIX E: SUPPLEMENTARY RESULTS FROM FITTING SELF-EFFICACY GROWTH MODELS

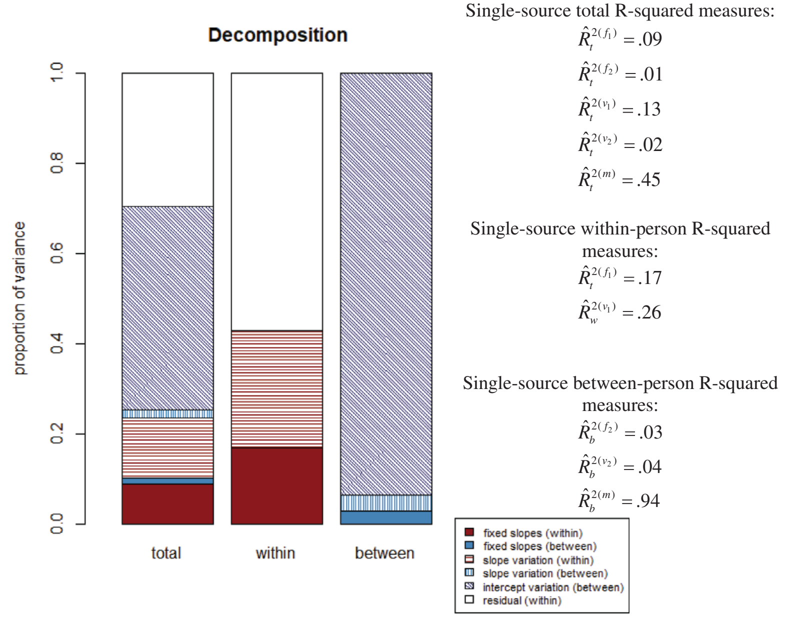

FIGURE E1 Visualizing R-squared results for the unconditional *quadratic* growth model of self-efficacy that centered *time* at-a-constant and assumed a homoscedastic, diagonal error covariance structure (Equation 3) *Notes*. See manuscript Figures 1 and 2 notes.

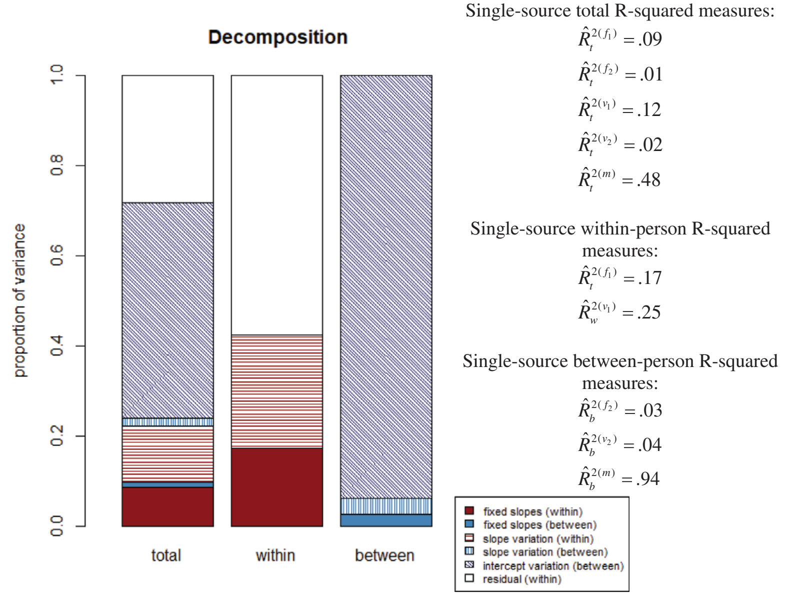

FIGURE E2 Visualizing R-squared results for the unconditional linear growth model of self-efficacy that centered *time* at-a-constant and specified a heteroscedastic, diagonal error covariance structure (see Equation 4) *Notes*. See manuscript Figures 1 and 2 notes.

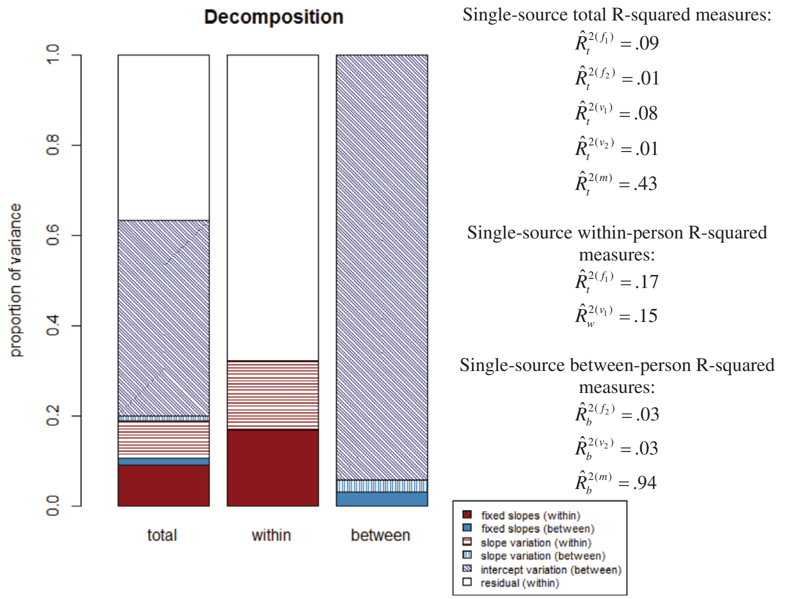

FIGURE E3 Visualizing R-squared results for the unconditional linear growth model of self-efficacy that centered *time* at-a-constant and specified a heteroscedastic, first-order autoregressive covariance structure (Equation 5)
Notes. See manuscript Figures 1 and 2 notes.

APPENDIX F: PROOF THAT WHEN *TIME* BY ITSELF EXPLAINS ONLY WITHIN-PERSON VARIABILITY, *TIME*^{2} CAN STILL EXPLAIN SOME BETWEEN-PERSON VARIABILITY

Here we show that including higher-order (i.e., > 1) polynomial terms associated with level-1 predictors can explain between-cluster variance even when the level-1 predictor has only within-cluster variance. To illustrate, we will consider the reduced form expression for the quadratic model given in Equation (3); for generality, we will denote the level-1 predictors (e.g., *time*) as *x*:

$$y_{ij} = \gamma_{00} + u_{0j} + \gamma_{10}x_{ij} + \gamma_{20}x_{ij}^2 + u_{1j}x_{ij} + u_{2j}x_{ij}^2 + e_{ij}. \tag{F1}$$

Using the formula in Appendix B Equation (B6), we can compute the total variance attributable to the between-cluster-varying portion of *x* via its fixed component as

$$\begin{aligned}
\boldsymbol{\gamma}'\boldsymbol{\Phi}_b\boldsymbol{\gamma} &= \begin{bmatrix}\gamma_{00} & \gamma_{10} & \gamma_{20}\end{bmatrix}\begin{bmatrix}0 & 0 & 0\\ 0 & \text{var}(E_{i|j}[x_{ij}]) & \text{cov}(E_{i|j}[x_{ij}], E_{i|j}[x_{ij}^2])\\ 0 & \text{cov}(E_{i|j}[x_{ij}], E_{i|j}[x_{ij}^2]) & \text{var}(E_{i|j}[x_{ij}^2])\end{bmatrix}\begin{bmatrix}\gamma_{00}\\ \gamma_{10}\\ \gamma_{20}\end{bmatrix}\\
&= \begin{bmatrix}\gamma_{00} & \gamma_{10} & \gamma_{20}\end{bmatrix}\begin{bmatrix}0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & \text{var}(E_{i|j}[x_{ij}^2])\end{bmatrix}\begin{bmatrix}\gamma_{00}\\ \gamma_{10}\\ \gamma_{20}\end{bmatrix}\\
&= \gamma_{20}^2\text{var}(E_{i|j}[x_{ij}^2])\\
&= \gamma_{20}^2\text{var}(\text{var}_{i|j}(x_{ij}) + E_{i|j}[x_{ij}]^2)\\
&= \gamma_{20}^2\text{var}(\text{var}_{i|j}(x_{ij})).
\end{aligned} \tag{F2}$$

Note that $E_{i|j}[x_{ij}]$ is necessarily constant across clusters (as the level-1 predictor has only within-cluster variance), and hence $\text{var}(E_{i|j}[x_{ij}]) = \text{cov}(E_{i|j}[x_{ij}], E_{i|j}[x_{ij}^2]) = 0$. However, $E_{i|j}[x_{ij}^2]$ is not necessarily the same for each cluster, and will vary across clusters when clusters have different degrees of within-cluster variance of x. As shown in Equation (F2), holding all else constant, the amount of between-cluster variance explained by x_{ij}^2 via its fixed component will increase as the amount of across-cluster variability in the within-cluster variability of x increases.

We can similarly compute the total variance attributable to the between-cluster-varying portion of x via random slope variation as

$$\begin{aligned}
tr\left(\mathbf{T\Sigma}_b\right) &= tr\left(\begin{bmatrix}\tau_{00} & \tau_{01} & \tau_{02}\\ \tau_{01} & \tau_{11} & \tau_{12}\\ \tau_{02} & \tau_{12} & \tau_{22}\end{bmatrix}\begin{bmatrix}0 & 0 & 0\\ 0 & \text{var}\left(E_{i|j}\left[x_{ij}\right]\right) & \text{cov}\left(E_{i|j}\left[x_{ij}\right], E_{i|j}\left[x_{ij}^2\right]\right)\\ 0 & \text{cov}\left(E_{i|j}\left[x_{ij}\right], E_{i|j}\left[x_{ij}^2\right]\right) & \text{var}\left(E_{i|j}\left[x_{ij}^2\right]\right)\end{bmatrix}\right)\\
&= tr\left(\begin{bmatrix}\tau_{00} & \tau_{01} & \tau_{02}\\ \tau_{01} & \tau_{11} & \tau_{12}\\ \tau_{02} & \tau_{12} & \tau_{22}\end{bmatrix}\begin{bmatrix}0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & \text{var}\left(E_{i|j}\left[x_{ij}^2\right]\right)\end{bmatrix}\right)\\
&= \tau_{22}\text{var}\left(E_{i|j}\left[x_{ij}^2\right]\right)\\
&= \tau_{22}\text{var}\left(\text{var}_{i|j}\left(x_{ij}\right) + E_{i|j}\left[x_{ij}\right]^2\right)\\
&= \tau_{22}\text{var}\left(\text{var}_{i|j}\left(x_{ij}\right)\right).
\end{aligned} \tag{F3}$$

Similarly, Equation (F3) shows that, holding all else constant, the amount of between-cluster variance explained by x_{ij}^2 via random slope variation will increase as the amount of across-cluster variability in the within-cluster variability of x increases.

APPENDIX G: PROOF THAT HETEROSCEDASTIC LEVEL-1 ERROR VARIANCE CAN BE ACCOMMODATED IN THE R-SQUARED FRAMEWORK BY REPLACING THE σ^2 TERM WITH THE EXPECTED VALUE OF σ_i^2 ACROSS ALL TIMEPOINTS

In Appendix B Equation (B10), the level-1 error was assumed homoscedastic, and thus the variance was given as a single value, σ^2. Here, we will consider the more general case in which the variance of e_{ij} can differ as a discrete or continuous function of covariates, such as allowing it to differ discretely across timepoints, or specifying it as a smooth parametric function of time. We will first let $\mathbf{e}_j$ be a cluster-specific $n_j \times 1$ vector of all e_{ij}'s for cluster j (where n_j is the number of observations in cluster j). We can thus allow for heteroscedastic errors with the following expression:

$$\mathbf{e}_j \sim MVN\left(\begin{bmatrix}0\\0\\ \vdots\\0\end{bmatrix}, \begin{bmatrix}\sigma_1^2 & & & \\ 0 & \sigma_2^2 & & \\ \vdots & \ddots & \ddots & \\ 0 & \cdots & 0 & \sigma_{n_i}^2\end{bmatrix}\right), \tag{G1}$$

where each diagonal element in the covariance matrix denotes the error variance at each particular value of i, and each off-diagonal element of 0 indicates that there is no autocorrelation (an assumption we relax in Appendix H).

We will then let $\mathbf{I}_{ij}$ be a cluster-specific $n_j \times 1$ indicator vector such that the ith element is equal to 1 and all other elements are equal to 0. We can then say that

$$e_{ij} = \mathbf{I}'_{ij}\mathbf{e}_j. \tag{G2}$$

For example, if a given cluster j has four observations, and i is 3, then

$$e_{3j} = [0 \quad 0 \quad 1 \quad 0] \begin{bmatrix} e_{1j} \\ e_{2j} \\ e_{3j} \\ e_{4j} \end{bmatrix} = e_{3j}. \tag{G3}$$

Using the law of total variance, we can thus compute the variance of e_{ij} as

$$\begin{aligned} \text{var}\left(e_{ij}\right) &= \text{var}\left(\mathbf{I}'_{ij}\mathbf{e}_j\right) \\ &= E\left[\text{var}\left(\mathbf{I}'_{ij}\mathbf{e}_j\right)|\mathbf{I}'_{ij}\right] + \text{var}\left(E\left[\mathbf{I}'_{ij}\mathbf{e}_j|\mathbf{I}'_{ij}\right]\right) \\ &= E\left[\sigma_i^2\right] + \text{var}\,(0) \\ &= E\left[\sigma_i^2\right], \end{aligned} \tag{G4}$$

where σ_i^2 is the error variance for observation i. Hence, the overall/across-time error variance is, sensibly, the expected value of the error variance.

Computing this expected value when there are a discrete set of error variances, by definition, can be done as

$$\begin{aligned} \text{var}\left(e_{ij}\right) &= E\left[\sigma_i^2\right] \\ &= \sum_{i=1}^{\max(n_j)} \pi_i \sigma_i^2, \end{aligned} \tag{G5}$$

where $\max(n_j)$ is the largest possible cluster size and π_i is the probability of a randomly selected observation being the ith observation within a cluster. As an example, if we had four discrete timepoints with a separate error variance for each, and had an equal number of observations per timepoint, the expected error variance would be given as

$$\begin{aligned} \text{var}(e_{ij}) &= E[\sigma_i^2] \\ &= \sum_{t=1}^{4} \pi_i \sigma_i^2 \\ &= \frac{1}{4}\sigma_1^2 + \frac{1}{4}\sigma_2^2 + \frac{1}{4}\sigma_3^2 + \frac{1}{4}\sigma_4^2, \end{aligned} \tag{G6}$$

which is just the unweighted mean of the four error variances.

When there is not a discrete set of error variances, and the error variance instead varies as a function of continuous covariates, we compute the expected error variance as

$$\begin{aligned}\operatorname{var}\left(e_{ij}\right) &= E\left[\sigma_i^2\right]\\ &= E\left[\boldsymbol{\beta}'\mathbf{X}_{ij}\right]\\ &= \boldsymbol{\beta}'E\left[\mathbf{X}_{ij}\right],\end{aligned} \tag{G7}$$

where $\boldsymbol{\beta}$ is the vector of coefficients that are used to model the error variance, and $\mathbf{X}_{ij}$ is the vector of observation-specific predictors of the error variance. As an example, if the error variance differed as a linear function of *time*, then

$$\begin{aligned}\operatorname{var}(e_{ij}) &= E[\sigma_i^2]\\ &= E\left[[\beta_0 \quad \beta_1]\begin{bmatrix}1\\ time_{ij}\end{bmatrix}\right]\\ &= E\left[\beta_0 + \beta_1 time_{ij}\right]\\ &= \beta_0 + \beta_1 E[time_{ij}].\end{aligned} \tag{G8}$$

Similarly, if the error variance were modeled as a quadratic function of *time*,

$$\begin{aligned}\operatorname{var}\left(e_{ij}\right) &= E\left[\sigma_i^2\right]\\ &= E\left[\beta_0 + \beta_1 time_{ij} + \beta_2 time_{ij}^2\right]\\ &= \beta_0 + \beta_1 E\left[time_{ij}\right] + \beta_2 E\left[time_{ij}^2\right].\end{aligned} \tag{G9}$$

As a last example, if the error variance were modeled as a quadratic function of *time*, and a linear function of some other covariate x_{ij}

$$\begin{aligned}\operatorname{var}\left(e_{ij}\right) &= E\left[\sigma_i^2\right]\\ &= E\left[\beta_0 + \beta_1 time_{ij} + \beta_2 time_{ij}^2 + \beta_3 x_{ij}\right]\\ &= \beta_0 + \beta_1 E\left[time_{ij}\right] + \beta_2 E\left[time_{ij}^2\right] + \beta_3 E\left[x_{ij}\right].\end{aligned} \tag{G10}$$

When estimating these quantities in a sample, the above expectations can be replaced with sample means, and the parameters can be replaced with estimates. For instance, with our example dataset, if we were to specify the error variance to have the form of Equation (G10) (with x_{ij} denoting *female* to allow for different error variances for boys and girls), and we would obtain estimates of $\hat{\beta}_0 = 70$, $\hat{\beta}_1 = 5$, $\hat{\beta}_2 = 0.01$, and $\hat{\beta}_3 = 2$, and we would then compute the expected error variance as $70 + 5(\text{sample mean of time}_{ij}) + 0.01(\text{sample mean of time}_{ij}^2) + 2(\text{sample mean of gender}_{ij})$, which in our case is 90. Hence, our estimate of $E[\sigma_i^2]$ would be 90 (e.g., we would enter 90 in the *sigma2* argument of the *r2MLMlong* function in Appendix J).

Integrating Appendix B and the current Appendix G, the total model-implied outcome variance allowing heteroscedasticity of level-1 errors is given as

$$\mathrm{var}\left(y_{ij}\right) = \boldsymbol{\gamma}'\boldsymbol{\Phi}_w\boldsymbol{\gamma} + \boldsymbol{\gamma}'\boldsymbol{\Phi}_b\boldsymbol{\gamma} + tr\left(\mathbf{T}\boldsymbol{\Sigma}_w\right) + tr\left(\mathbf{T}\boldsymbol{\Sigma}_b\right) + \mathbf{m}'\mathbf{T}\mathbf{m} + E\left[\sigma_i^2\right]. \tag{G11}$$

The sources corresponding to these terms are defined in Table 2. The only difference in this expression and that provided in Appendix B Equation (B10) is that σ^2 is replaced with $E[\sigma_i^2]$.

APPENDIX H: PROOF THAT THE MATHEMATICAL COMPUTATION OF THE FRAMEWORK'S R-SQUARED MEASURES IS UNAFFECTED BY THE INCLUSION OF ANY KIND OF AUTOCORRELATION

In the newly derived expression for var(e_{ij}) given in Appendix G, we assumed there was no autocorrelation. Here we prove that the addition of autocorrelation does not change this formula. We will expand the expression in Appendix G Equation (G1) to allow for autocorrelation as such:

$$\mathbf{e}_j \sim MVN\left(\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & & & \\ \sigma_{12}^2 & \sigma_2^2 & & \\ \vdots & \ddots & \ddots & \\ \sigma_{1n_j}^2 & \cdots & \sigma_{(n_j-1)n_j}^2 & \sigma_{1n_j}^2 \end{bmatrix}\right). \tag{H1}$$

Here, each off-diagonal element denotes the error covariance between two different values of i.

Again letting $\mathbf{I}_{ij}$ be a cluster-specific $n_j \times 1$ indicator vector such that the ith element is equal to 1 and all other elements are equal to 0. We can then say that

$$e_{ij} = \mathbf{I}'_{ij}\mathbf{e}_j. \tag{H2}$$

And can compute the variance of e_{ij} as

$$\begin{aligned} \mathrm{var}\left(e_{ij}\right) &= \mathrm{var}\left(\mathbf{I}'_{ij}\mathbf{e}_j\right) \\ &= E\left[\mathrm{var}\left(\mathbf{I}'_{ij}\mathbf{e}_j\right)|\mathbf{I}'_{ij}\right] + \mathrm{var}\left(E\left[\mathbf{I}'_{ij}\mathbf{e}_j|\mathbf{I}'_{ij}\right]\right) \\ &= E\left[\sigma_i^2\right] + \mathrm{var}\left(0\right) \\ &= E\left[\sigma_i^2\right]. \end{aligned} \tag{H3}$$

This expression here in Equation (H3) is identical to that in Equation (G4), and hence, the autocorrelation does not change the formulas used to compute R-squared measures.

APPENDIX I: SPECIFICATIONS USED TO OBTAIN ΔR^2 EFFECT SIZES FOR INDIVIDUAL TERMS IN OUR ILLUSTRATIVE CONDITIONAL GROWTH MODEL OF SELF-EFFICACY

Letting Model B denote the full model of interest given in Equation (6), and using a *simultaneous* model-building approach (see Rights & Sterba, 2020) we can compute the variance uniquely explained by GPA (via each of f_1, f_2, and v_1) by computing R-squared differences between the full Model B and the following reduced Model A that excludes both person-mean-centered GPA and person-mean GPA:

$$\begin{aligned}
selfeff_{ij} &= \beta_{0j} + \beta_{1j} time_{ij} + \beta_{2j}\left(volunteer_{ij} - volunteer_j\right) + e_{ij},\\
\beta_{0j} &= \gamma_{00} + \gamma_{01} female_j + \gamma_{03} volunteer_j + u_{0j},\\
\beta_{1j} &= \gamma_{10} + \gamma_{11} female_j + u_{1j},\\
\beta_{2j} &= \gamma_{20} + u_{2j}.
\end{aligned} \tag{I1}$$

Specifically, the variance uniquely explained by person-mean-centered GPA via its fixed component is estimated as $\Delta\hat{R}_t^{2(f_1)}$(for total variance) and $\Delta\hat{R}_w^{2(f_1)}$(for within-person variance), the variance uniquely explained by person-mean GPA via its fixed component is estimated as $\Delta\hat{R}_t^{2(f_2)}$(for total variance) and $\Delta\hat{R}_b^{2(f_2)}$ (for between-person variance), and the variance uniquely explained by person-mean-centered GPA via random slope variation is estimated as $\Delta\hat{R}_t^{2(v_1)}$ (for total variance) and $\Delta\hat{R}_w^{2(v_1)}$ (for within-person variance).

We can similarly compute the variance uniquely explained by volunteer hours (via each of f_1, f_2, and v_1) by computing these same R-squared differences between the full Model B and the following reduced Model C that excludes person-mean-centered volunteer hours and person-mean volunteer hours:

$$\begin{aligned}
selfeff_{ij} &= \beta_{0j} + \beta_{1j} time_{ij} + \beta_{2j}\left(gpa_{ij} - gpa_j\right) + e_{ij},\\
\beta_{0j} &= \gamma_{00} + \gamma_{01} female_j + \gamma_{02} gpa_j + u_{0j},\\
\beta_{1j} &= \gamma_{10} + \gamma_{11} female_j + u_{1j},\\
\beta_{2j} &= \gamma_{20} + u_{2j}.
\end{aligned} \tag{I2}$$

Lastly, we can compute the variance uniquely explained by the product term of $time \times female$ by comparing the full Model B and the following reduced Model D that excludes this product term:

$$\begin{aligned}
selfeff_{ij} &= \beta_{0j} + \beta_{1j} time_{ij} + \beta_{2j}\left(gpa_{ij} - gpa_j\right) + \beta_{3j}\left(volunteer_{ij} - volunteer_j\right) + e_{ij},\\
\beta_{0j} &= \gamma_{00} + \gamma_{01} female_j + \gamma_{02} gpa_j + \gamma_{03} work_j + u_{0j},\\
\beta_{1j} &= \gamma_{10} + u_{1j},\\
\beta_{2j} &= \gamma_{20} + u_{2j},\\
\beta_{3j} &= \gamma_{30} + u_{3j}.
\end{aligned} \tag{I3}$$

Specifically, the variance uniquely explained by *time* × *female* via its fixed component is estimated as the sum of $\Delta\hat{R}_t^{2(f_1)}$nd $\Delta\hat{R}_t^{2(f_2)}$(for total variance), by$\Delta\hat{R}_w^{2(f_1)}$ (for within-person variance), and by $\Delta\hat{R}_b^{2(f_2)}$for between-person variance).

APPENDIX J: SOFTWARE IMPLEMENTATION OF *R*-SQUARED FRAMEWORK VIA R FUNCTION r2MLMlong

***r2MLMlong* R function description**:

This R function reads in raw data as well as parameter estimates from the researcher's previously fit longitudinal growth model (hence, any software program can have been used to fit the researcher's longitudinal growth model prior to the use of this R function, so long as parameter estimates from the fitted model are recorded). This function then outputs R-squared measures (shown in manuscript Table 3), as well as variance decompositions and associated barcharts (e.g., Figures 1–3). This function allows researchers to input heteroscedastic residual variance by including multiple estimates, for example, corresponding to individual timepoints. Users need not specify if predictors are person-mean-centered or not—the function will automatically output total, within-person, and between-person variance attributable to each potential source of explained variance (f_1, f_2, v_1, v_2, and m). Note, however, that the interpretations of these sources differ for person-mean-centered versus non-person-mean-centered models (as delineated in manuscript Table 2) and that variance attributable to v_2 will necessarily be 0 for person-mean-centered models.

***r2MLMlong* R function input description**:

data—dataset in long format, in which rows denote individual observations and columns denote variables

covs—list of predictors in the dataset that have fixed components of slopes included in the model (if none, set to NULL)

random_covs—list of predictors in the dataset that have random components of slopes included in the model (if none, set to NULL)

clusterID—variable name in dataset corresponding to cluster (e.g., person) identification

gammas—vector containing estimated fixed components of all slopes, listed in the order specified in *covs* (if none, set to NULL)

Tau—random effect covariance matrix; the first row and the first column denote the intercept variance and covariances and each subsequent row/column denotes a given random slope's variance and covariances (to be entered in the order listed by *random_covs*)

sigma2—level-1 residual variance; can be entered as a single number, or as a set of numbers, for example corresponding to different residual variances at individual timepoints; if entered as a set of numbers, function will assume equal weights and take the raw average of these to estimate the expectation of the error variance

r2MLMlong R function example input:

```
#NOTE: estimates in the input represent hypothetical results for a random slope model with "time" and "x"
  as level-1 predictors and "w1" and "w2" as level-2 predictors; model also allows level-1 residual variance
  to vary across the five timepoints
#in practice a user would have previously obtained these input estimates by fitting their model in MLM
  software
#additionally, the input consists of hypothetical predictor data, whereas in practice a user would read-in
  their actual data

exampledata <- matrix(NA,100*5,5)
time <-rep(seq(5),100)
x <- rnorm(100*5,0,2)
w1<-rnorm(100,2,1)
w2<-rnorm(100,3,2)
exampledata[,1] <- rep(seq(100),each=5)
exampledata[,2:3] <- cbind(time,x)
exampledata[,4] <- rep(w1,each=5)
exampledata[,5] <- rep(w2,each=5)
exampledata<-as.data.frame(exampledata)

colnames(exampledata) <- c("person","time","x","w1","w2")
r2MLMlong(data=exampledata,covs=c("time","x","w1","w2"),random_covs=c("time","x"),
  gammas=c(.25,1.5,-.75,.01),
clusterID="person",Tau=matrix(c(4,1,.75,1,1,.25,.75,.25,.5),3,3),sigma2=c(10,11,12,14,15))
```

r2MLMlong R function code:

```
#need to install the following packages
library(rockchalk)

r2MLMlong <- function(data,covs,random_covs,clusterID,gammas,Tau,sigma2){

  if(is.null(covs)==FALSE){
  centered_data <- gmc(data,covs,clusterID)
  phi_w <- var(centered_data[,c(paste0(covs,"_dev"))])     phi_b <-
  var(centered_data[,c(paste0(covs,"_mn"))])     gammas <- matrix(c(gammas),ncol=1)
  f1<-t(gammas)%*%phi_w%*%gammas
  f2<-t(gammas)%*%phi_b%*%gammas
  }
  else{
  f1<-0
  f2<-0
  }

  if(is.null(random_covs)==FALSE){
  centered_data_rand <- gmc(data,random_covs,clusterID)
  Sig_w <- var(centered_data_rand[,c(paste0(random_covs,"_dev"))])     Sig_b <-
  var(centered_data_rand[,c(paste0(random_covs,"_mn"))])     m_mat <-
  matrix(c(colMeans(cbind(1,data[,c(random_covs)]))),ncol=1)
  v1<-sum(diag(Tau[2:nrow(Tau),2:nrow(Tau)]%*%Sig_w))
  v2<-sum(diag(Tau[2:nrow(Tau),2:nrow(Tau)]%*%Sig_b))
  }
```

```
else{
 v1<-0
 v2<-0
 m_mat <- 1
 }

 m<- t(m_mat)%*%Tau%*%m_mat

 sigma<-mean(sigma2)

 #decompositions

 decomp_fixed_within <- f1/sum(f1,f2,v1,v2,m,sigma)
 decomp_fixed_between <-f2/sum(f1,f2,v1,v2,m,sigma)
 decomp_varslopes_within <- v1/sum(f1,f2,v1,v2,m,sigma)
 decomp_varslopes_between <- v2/sum(f1,f2,v1,v2,m,sigma)
 decomp_varmeans <- m/sum(f1,f2,v1,v2,m,sigma)
 decomp_sigma <- sigma/sum(f1,f2,v1,v2,m,sigma)

 decomp_fixed_within_w <- f1/sum(f1,v1,sigma)
 decomp_fixed_between_b <-f2/sum(f2,v2,m)
 decomp_varslopes_within_w <- v1/sum(f1,v1,sigma)
 decomp_varslopes_between_b <- v2/sum(f2,v2,m)
 decomp_varmeans_b <- m/sum(f2,v2,m)
 decomp_sigma_w <- sigma/sum(f1,v1,sigma)

 #barchart

 contributions_stacked <-
 matrix(c(decomp_fixed_within,decomp_fixed_between,decomp_varslopes_within,decomp_varslopes_
 between,decomp_varmeans,decomp_sigma,
          decomp_fixed_within_w,0,decomp_varslopes_within_w,0,0,decomp_sigma_w,
          0,decomp_fixed_between_b,0,decomp_varslopes_between_b,decomp_varmeans_b,0),6,3)
 colnames(contributions_stacked) <- c("total","within","between")
 rownames(contributions_stacked) <- c("fixed slopes (within)",
          "fixed slopes (between)",
          "slope variation (within)",
          "slope variation (between)",
          "intercept variation (between)",
          "residual (within)")


 barplot(contributions_stacked, main="Decomposition", horiz=FALSE,
       ylim=c(0,1),col=c("darkred","steelblue","darkred","steelblue","midnightblue","white"),
          ylab="proportion of variance",
       density=c(NA,NA,30,40,40,NA),angle=c(0,45,0,90,135,0),xlim=c(0,1.5),width=c(.3,.3))
 legend(1.1,.65,legend=rownames(contributions_stacked_avg),fill=c("darkred",
 "steelblue","darkred","steelblue","midnightblue","white"),
       cex=.7, pt.cex = 1,xpd=T,density=c(NA,NA,30,40,40,NA),angle=c(0,45,0,90,135,0))

 #create tables for output

 decomp_table <-
 matrix(c(decomp_fixed_within,decomp_fixed_between,decomp_varslopes_within,decomp_varslopes_
 between,decomp_varmeans,decomp_sigma,
       decomp_fixed_within_w,"NA",decomp_varslopes_within_w,"NA","NA",decomp_sigma_w,
       "NA",decomp_fixed_between_b,"NA",decomp_varslopes_between_b,decomp_varmeans_b,"NA"),6,3)
```

```
 colnames(decomp_table) <- c("total","within","between")
rownames(decomp_table) <- c("fixed slopes (within)",
            "fixed slopes (between)",
            "slope variation (within)",
            "slope variation (between)",
            "intercept variation (between)",
            "residual (within)")

R2_table <- matrix(c(decomp_fixed_within,decomp_fixed_between,decomp_varslopes_within,
decomp_varslopes_between,decomp_varmeans,
decomp_fixed_within+decomp_fixed_between,decomp_fixed_within+decomp_fixed_between+decomp_
varslopes_within+decomp_varslopes_between,
        decomp_fixed_within+decomp_fixed_between+decomp_varslopes_within+decomp_
            varslopes_between+decomp_varmeans,
        decomp_fixed_within_w,"NA",decomp_varslopes_within_w,"NA","NA","NA",decomp_
            fixed_within_w+decomp_varslopes_within_w,"NA",
"NA",decomp_fixed_between_b,"NA",decomp_varslopes_between_b,decomp_varmeans_b,"NA",
decomp_fixed_between_b+decomp_varslopes_between_b,"NA"),8,3)

colnames(R2_table) <- c("total","within","between")
rownames(R2_table) <- c("f1","f2","v1","v2","m","f","fv","fvm")

Output <- list(noquote(decomp_table),noquote(R2_table))
names(Output) <- c("Decompositions","R2s")

return(Output)
}
```