

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

Table of Contents

Appendix A. Review from Rights & Sterba (2019) of definitions of R-squared measures and corresponding formulas, assuming person-mean-centering	1
Appendix B. New extension: Full decomposition of outcome variance for <i>non</i>-person-mean-centered models.....	3
Appendix C. New extension: Full set of definitions of R-squared measures and corresponding formulas under <i>non</i>-person-mean-centering	6
Appendix D. Proof that the proportion of variance attributable to source <i>m</i>, defined in Table 2, does not change when centering predictors by a constant, regardless of the chosen centering constant's value.....	7
Appendix E. Supplementary results from fitting self-efficacy growth model.....	10
Appendix F: Proof that when <i>time</i> by itself explains only within-person variability, <i>time</i>² can still explain some between-person variability.....	10
Appendix G: Proof that heteroscedastic level-1 error variance can be accommodated in the R-squared framework by replacing σ^2 with the expected value of σ_i^2 across all timepoints.....	15
Appendix H. Proof that the mathematical computation of the framework's R-squared measures is unaffected by the inclusion of any kind of autocorrelation.....	18
Appendix I. Specifications used to obtain ΔR^2 effect sizes for individual terms in our illustrative conditional growth model of self-efficacy.....	.19
Appendix J. Software implementation of R-squared framework via R function <i>r2MLMlong</i>...	20

Appendix A. Review from Rights & Sterba (2019) of definitions of R-squared measures and corresponding formulas, assuming person-mean-centering

Here we briefly review the Rights & 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

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

Where \mathbf{x}_{ij}^w 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}_{ij}^w , \mathbf{x}_j^b 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 & Sterba (2019), is then given as

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

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: $\boldsymbol{\gamma}^{w'} \boldsymbol{\Phi}^w \boldsymbol{\gamma}^w$ denotes variance attributable to level-1 predictors via fixed components of slopes, $\boldsymbol{\gamma}^{b'} \boldsymbol{\Phi}^b \boldsymbol{\gamma}^b$ variance attributable to level-2 predictors via fixed components of slopes, $\text{tr}(\mathbf{T}\boldsymbol{\Sigma})$ 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 ($\boldsymbol{\gamma}^{w'} \boldsymbol{\Phi}^w \boldsymbol{\gamma}^w$, $\text{tr}(\mathbf{T}\boldsymbol{\Sigma})$, and σ^2) reflect purely within-person variation, whereas the other two ($\boldsymbol{\gamma}^{b'} \boldsymbol{\Phi}^b \boldsymbol{\gamma}^b$ 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:

$$\begin{aligned}
R_t^{2(f_1)} &= \frac{\boldsymbol{\gamma}^{w'} \boldsymbol{\Phi}^w \boldsymbol{\gamma}^w}{\boldsymbol{\gamma}^{w'} \boldsymbol{\Phi}^w \boldsymbol{\gamma}^w + \boldsymbol{\gamma}^{b'} \boldsymbol{\Phi}^b \boldsymbol{\gamma}^b + \text{tr}(\mathbf{T}\boldsymbol{\Sigma}) + \tau_{00} + \sigma^2} \\
R_t^{2(f_2)} &= \frac{\boldsymbol{\gamma}^{b'} \boldsymbol{\Phi}^b \boldsymbol{\gamma}^b}{\boldsymbol{\gamma}^{w'} \boldsymbol{\Phi}^w \boldsymbol{\gamma}^w + \boldsymbol{\gamma}^{b'} \boldsymbol{\Phi}^b \boldsymbol{\gamma}^b + \text{tr}(\mathbf{T}\boldsymbol{\Sigma}) + \tau_{00} + \sigma^2} \\
R_t^{2(v)} &= \frac{\text{tr}(\mathbf{T}\boldsymbol{\Sigma})}{\boldsymbol{\gamma}^{w'} \boldsymbol{\Phi}^w \boldsymbol{\gamma}^w + \boldsymbol{\gamma}^{b'} \boldsymbol{\Phi}^b \boldsymbol{\gamma}^b + \text{tr}(\mathbf{T}\boldsymbol{\Sigma}) + \tau_{00} + \sigma^2} \\
R_t^{2(m)} &= \frac{\tau_{00}}{\boldsymbol{\gamma}^{w'} \boldsymbol{\Phi}^w \boldsymbol{\gamma}^w + \boldsymbol{\gamma}^{b'} \boldsymbol{\Phi}^b \boldsymbol{\gamma}^b + \text{tr}(\mathbf{T}\boldsymbol{\Sigma}) + \tau_{00} + \sigma^2}
\end{aligned} \tag{A3}$$

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

$$\begin{aligned}
R_w^{2(f_1)} &= \frac{\boldsymbol{\gamma}^{w'} \boldsymbol{\Phi}^w \boldsymbol{\gamma}^w}{\boldsymbol{\gamma}^{w'} \boldsymbol{\Phi}^w \boldsymbol{\gamma}^w + \text{tr}(\mathbf{T}\boldsymbol{\Sigma}) + \sigma^2} \\
R_w^{2(v)} &= \frac{\text{tr}(\mathbf{T}\boldsymbol{\Sigma})}{\boldsymbol{\gamma}^{w'} \boldsymbol{\Phi}^w \boldsymbol{\gamma}^w + \text{tr}(\mathbf{T}\boldsymbol{\Sigma}) + \sigma^2}
\end{aligned} \tag{A4}$$

A combination-source within-person measure is given as $R_w^{2(f_1v)} = R_w^{2(f_1)} + R_w^{2(v)}$. The single-source between-person measures for person-mean-centered models are likewise:

$$\begin{aligned}
R_b^{2(f_2)} &= \frac{\boldsymbol{\gamma}^{b'} \boldsymbol{\Phi}^b \boldsymbol{\gamma}^b}{\boldsymbol{\gamma}^{b'} \boldsymbol{\Phi}^b \boldsymbol{\gamma}^b + \tau_{00}} \\
R_b^{2(m)} &= \frac{\tau_{00}}{\boldsymbol{\gamma}^{b'} \boldsymbol{\Phi}^b \boldsymbol{\gamma}^b + \tau_{00}}
\end{aligned}$$

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 & 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:

$$\begin{aligned} y_{ij} &= \mathbf{x}'_{ij}\boldsymbol{\gamma} + \mathbf{w}'_{ij}\mathbf{u}_j + e_{ij} \\ \mathbf{u}_j &\sim MVN(\mathbf{0}, \mathbf{T}) \\ e_{ij} &\sim N(0, \sigma^2) \end{aligned} \quad (\text{B1})$$

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:

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

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 $E_{ij}[\mathbf{x}_{ij} - \mathbf{x}_{.j}] = 0$ 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

$\mathbf{x}_{\cdot j}$ and $\mathbf{w}_{\cdot j}$, respectively (since $\mathbf{x}_{\cdot j}$ and $\mathbf{w}_{\cdot j}$ are vectors of constants for each cluster, i.e., $\text{var}_{ij}[\mathbf{x}_{\cdot j}] = \text{var}_{ij}[\mathbf{w}_{\cdot j}] = 0$).

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

$$\begin{aligned} y_{ij} &= \mathbf{x}'_{ij}\boldsymbol{\gamma} + \mathbf{w}'_{ij}\mathbf{u}_j + e_{ij} \\ &= (\mathbf{x}_{ij} - \mathbf{x}_{\cdot j} + \mathbf{x}_{\cdot j})'\boldsymbol{\gamma} + (\mathbf{w}_{ij} - \mathbf{w}_{\cdot j} + \mathbf{w}_{\cdot j})'\mathbf{u}_j + e_{ij} \\ &= (\mathbf{x}_{ij} - \mathbf{x}_{\cdot j})'\boldsymbol{\gamma} + \mathbf{x}'_{\cdot j}\boldsymbol{\gamma} + (\mathbf{w}_{ij} - \mathbf{w}_{\cdot j})'\mathbf{u}_j + \mathbf{w}'_{\cdot j}\mathbf{u}_j + e_{ij} \end{aligned} \quad (\text{B3})$$

We can then compute the model-implied variance as

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

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

$$\text{var}((\mathbf{x}_{ij} - \mathbf{x}_{\cdot j})'\boldsymbol{\gamma}) = \boldsymbol{\gamma}'\boldsymbol{\Phi}_w\boldsymbol{\gamma} \quad (\text{B5})$$

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

$$\text{var}(\mathbf{x}'_{\cdot j}\boldsymbol{\gamma}) = \boldsymbol{\gamma}'\boldsymbol{\Phi}_b\boldsymbol{\gamma} \quad (\text{B6})$$

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

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

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

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

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

$$\text{var}(e_{ij}) = \sigma^2 \quad (\text{B9})$$

Thus, the total model-implied outcome variance is

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

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

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

Example combination-source total R-squared measures that could be constructed are:

$$\begin{aligned}
 R_t^{2(f)} &= R_t^{2(f_1)} + R_t^{2(f_2)}, R_t^{2(fv)} = R_t^{2(f_1)} + R_t^{2(f_2)} + R_t^{2(v_1)} + R_t^{2(v_2)}, \text{ and} \\
 R_t^{2(fvm)} &= R_t^{2(f_1)} + R_t^{2(f_2)} + R_t^{2(v_1)} + R_t^{2(v_2)} + R_t^{2(m)}.
 \end{aligned}$$

The single-source within-person measures are then

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

Summing these yields a combination-source within-person measure: $R_w^{2(f_1v_1)} = R_w^{2(f_1)} + R_w^{2(v_1)}$.

The single-source between-person measures are:

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

Summing these yields a combination-source between-person measure: $R_b^{2(f_2v_2)} = R_b^{2(f_2)} + R_b^{2(v_2)}$.

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:

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

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

$$y_{ij} = \gamma_{00}^* + u_{0j}^* + \gamma_{10}^*(x_{ij} - a) + u_{1j}^*(x_{ij} - a) + e_{ij}^* \tag{D2}$$

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, Aiken, & de Leeuw, 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:

$$\begin{aligned}
 y_{ij} &= \gamma_{00}^* + u_{0j}^* + \gamma_{10}^*(x_{ij} - a) + u_{1j}^*(x_{ij} - a) + e_{ij}^* \\
 &= \gamma_{00}^* + u_{0j}^* + \gamma_{10}^*x_{ij} - \gamma_{10}^*a + u_{1j}^*x_{ij} - u_{1j}^*a + e_{ij}^* \\
 &= (\gamma_{00}^* - \gamma_{10}^*a) + (u_{0j}^* - u_{1j}^*a) + \gamma_{10}^*x_{ij} + u_{1j}^*x_{ij} + e_{ij}^*
 \end{aligned} \tag{D3}$$

This reexpression 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 like so

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

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

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

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:

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

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

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

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

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

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

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

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

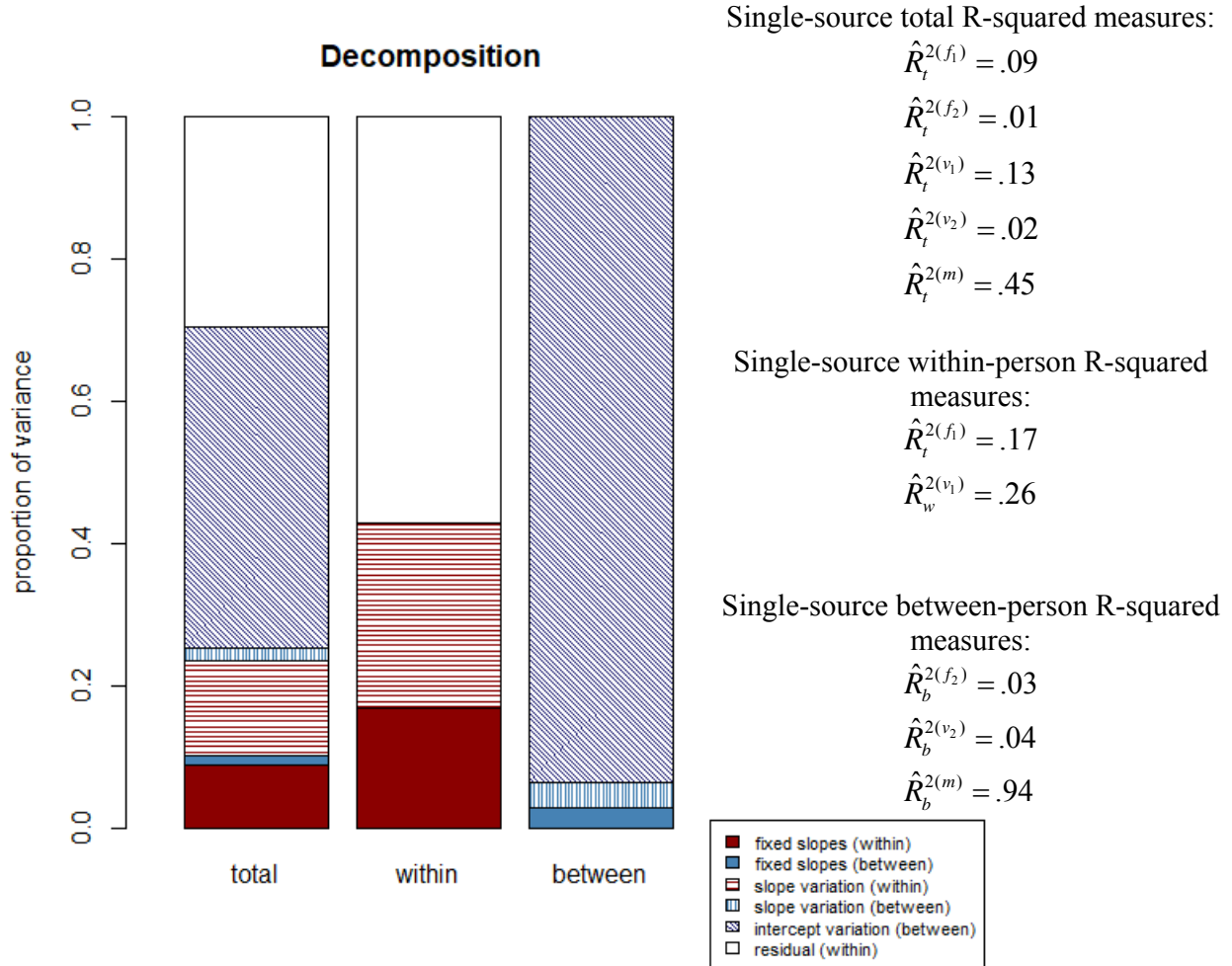
$$\begin{aligned}
\mathbf{m}'^* \mathbf{T}^* \mathbf{m}^* &= \begin{bmatrix} 1 & E[x_{ij} - a] \end{bmatrix} \begin{bmatrix} \tau_{00} + a^2 \tau_{11} + 2a\tau_{01} & \tau_{01} + a\tau_{11} \\ \tau_{01} + a\tau_{11} & \tau_{11} \end{bmatrix} \begin{bmatrix} 1 \\ E[x_{ij} - a] \end{bmatrix} \\
&= \begin{bmatrix} 1 & E[x_{ij} - a] \end{bmatrix} \begin{bmatrix} \tau_{00} + a^2 \tau_{11} + 2a\tau_{01} & \tau_{01} + a\tau_{11} \\ \tau_{01} + a\tau_{11} & \tau_{11} \end{bmatrix} \begin{bmatrix} 1 \\ E[x_{ij} - a] \end{bmatrix} \\
&= \begin{bmatrix} \tau_{00} + a^2 \tau_{11} + 2a\tau_{01} + (E[x_{ij}] - a)(\tau_{01} + a\tau_{11}) & \tau_{01} + a\tau_{11} + (E[x_{ij}] - a)\tau_{11} \end{bmatrix} \begin{bmatrix} 1 \\ E[x_{ij} - a] \end{bmatrix} \\
&= \tau_{00} + a^2 \tau_{11} + 2a\tau_{01} + (E[x_{ij}] - a)(\tau_{01} + a\tau_{11}) + (E[x_{ij}] - a)(\tau_{01} + a\tau_{11}) + (E[x_{ij}] - a)^2 \tau_{11} \\
&= \tau_{00} + a^2 \tau_{11} + 2a\tau_{01} + 2(E[x_{ij}] - a)(\tau_{01} + a\tau_{11}) + E[x_{ij}]^2 \tau_{11} + a^2 \tau_{11} - 2aE[x_{ij}] \tau_{11} \\
&= \tau_{00} + a^2 \tau_{11} + 2a\tau_{01} + 2E[x_{ij}] \tau_{01} - 2a\tau_{01} - 2a^2 \tau_{11} + 2aE[x_{ij}] \tau_{11} + E[x_{ij}]^2 \tau_{11} + a^2 \tau_{11} - 2aE[x_{ij}] \tau_{11} \\
&= \tau_{00} + 2E[x_{ij}] \tau_{01} + E[x_{ij}]^2 \tau_{11}
\end{aligned} \tag{D11}$$

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

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

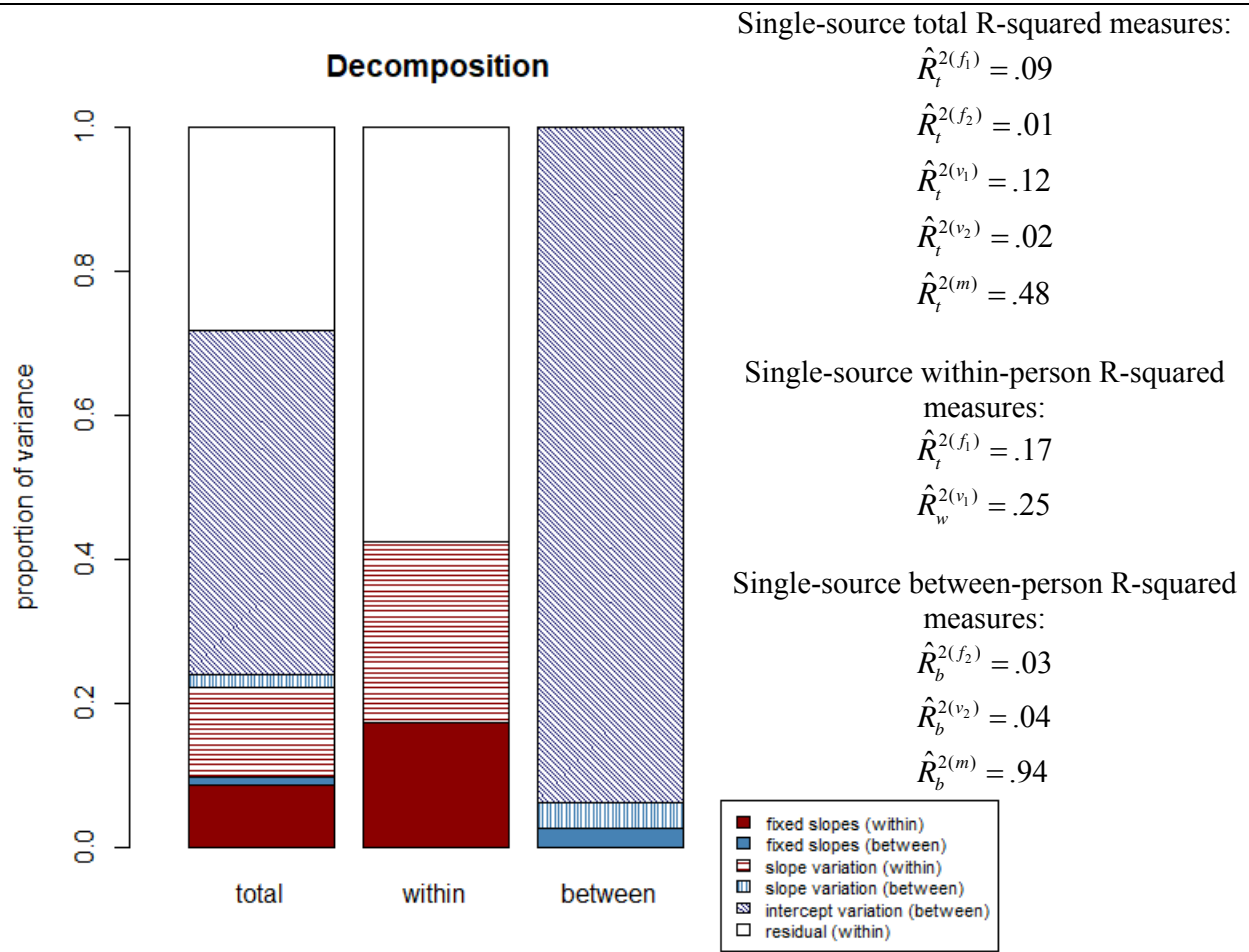
Appendix E. Supplementary results from fitting self-efficacy growth models

Appendix E 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 (Eqn. 3)



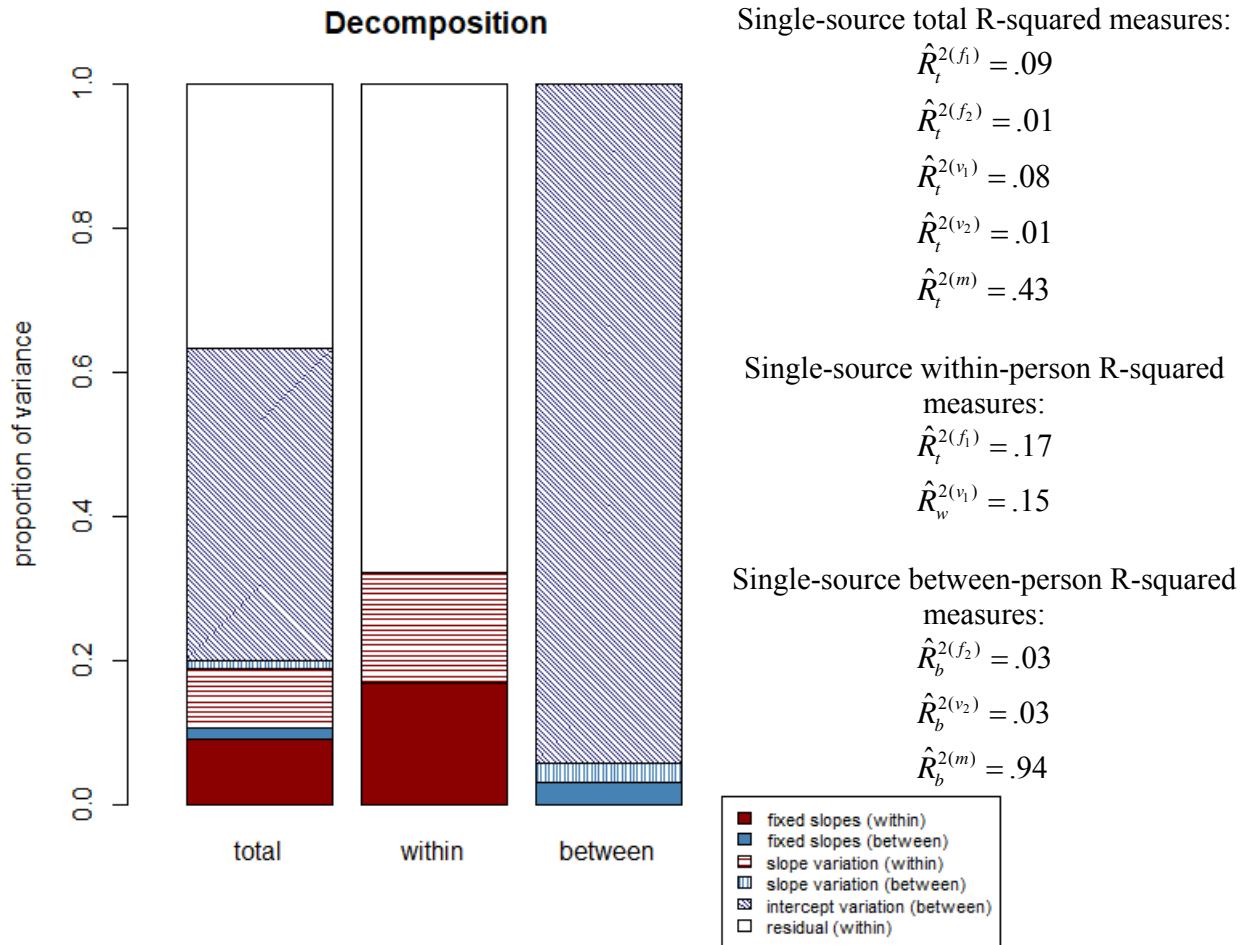
Notes. See manuscript Figure 1 and 2 notes.

Appendix E 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 Eqn. 4)



Notes. See manuscript Figure 1 and 2 notes.

Appendix E 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 (Eqn. 5)



Notes. See manuscript Figure 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 :

$$y_{ij} = \gamma_{00} + u_{0j} + \gamma_{10}x_{ij} + \gamma_{20}x_{ij}^2 + u_{1j}x_{ij} + u_{2j}x_{ij}^2 + e_{ij} \quad (\text{F1})$$

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

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

Note that $E_{ij}[x_{ij}]$ is necessarily constant across clusters (as the level-1 predictor has only within-cluster variance), and hence $\text{var}(E_{ij}[x_{ij}]) = \text{cov}(E_{ij}[x_{ij}], E_{ij}[x_{ij}^2]) = 0$. However, $E_{ij}[x_{ij}^2]$ 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

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

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'll 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:

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

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

$$e_{ij} = \mathbf{I}'_{ij} \mathbf{e}_j \quad (G2)$$

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

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

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

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

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 so

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

Where $\max(n_j)$ is the largest possible cluster size and π_i is the probability of a randomly selected observation being the i th 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

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

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

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

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

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

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

$$\begin{aligned}\text{var}(e_{ij}) &= E[\sigma_i^2] \\ &= E[\beta_0 + \beta_1 \text{time}_{ij} + \beta_2 \text{time}_{ij}^2] \\ &= \beta_0 + \beta_1 E[\text{time}_{ij}] + \beta_2 E[\text{time}_{ij}^2]\end{aligned}\tag{G9}$$

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}

$$\begin{aligned}\text{var}(e_{ij}) &= E[\sigma_i^2] \\ &= E[\beta_0 + \beta_1 \text{time}_{ij} + \beta_2 \text{time}_{ij}^2 + \beta_3 x_{ij}] \\ &= \beta_0 + \beta_1 E[\text{time}_{ij}] + \beta_2 E[\text{time}_{ij}^2] + \beta_3 E[x_{ij}]\end{aligned}\tag{G10}$$

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

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

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

$$\mathbf{e}_j \sim MVN \left(\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & & & & \\ \sigma_{12}^2 & \sigma_2^2 & & & \\ \vdots & \vdots & \ddots & & \\ \sigma_{(n_i-1)n_i}^2 & \dots & \sigma_{(n_i-1)n_i}^2 & & \sigma_{n_i}^2 \end{bmatrix} \right) \quad (\text{H1})$$

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 i th element is equal to 1 and all other elements are equal to 0. We can then say that

$$e_{ij} = \mathbf{I}'_{ij} \mathbf{e}_j \quad (\text{H2})$$

And can compute the variance of e_{ij} as

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

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:

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

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:

$$\begin{aligned} selfeff_{ij} &= \beta_{0j} + \beta_{1j}time_{ij} + \beta_{2j}(gpa_{ij} - gpa_{.j}) + e_{ij} \\ \beta_{0j} &= \gamma_{00} + \gamma_{01}female_j + \gamma_{02}gpa_{.j} + u_{0j} \\ \beta_{1j} &= \gamma_{10} + \gamma_{11}female_j + u_{1j} \\ \beta_{2j} &= \gamma_{20} + u_{2j} \end{aligned} \quad (12)$$

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:

$$\begin{aligned} selfeff_{ij} &= \beta_{0j} + \beta_{1j}time_{ij} + \beta_{2j}(gpa_{ij} - gpa_{.j}) + \beta_{3j}(volunteer_{ij} - volunteer_{.j}) + e_{ij} \\ \beta_{0j} &= \gamma_{00} + \gamma_{01}female_j + \gamma_{02}gpa_{.j} + \gamma_{03}work_{.j} + u_{0j} \\ \beta_{1j} &= \gamma_{10} + u_{1j} \\ \beta_{2j} &= \gamma_{20} + u_{2j} \\ \beta_{3j} &= \gamma_{30} + u_{3j} \end{aligned} \quad (13)$$

Specifically, the variance uniquely explained by *time* \times *female* via its fixed component is estimated as the sum of $\Delta \hat{R}_t^{2(f_1)}$ and $\Delta \hat{R}_t^{2(f_2)}$ (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-square 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., manuscript Figures 1-3). This function allows researchers to input heteroscedastic residual variance by including multiple estimates, e.g., 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 vs. 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 that 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, e.g., 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 <- rmnorm(100*5,0,2)
w1<-rmnorm(100,2,1)
w2<-rmnorm(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,decomp_varslopes_within_w,0,0,decomp_sigma_w,
0,decomp_fixed_between_b,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)
}

```