This is a "preproof" accepted article for *Psychometrika*.
This version may be subject to change during the production process.
DOI: 10.1017/psy.2024.2

Deriving models of change with interpretable parameters: linear estimation with nonlinear inference

Ethan M. McCormick*[1]

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

September 6, 2024

*Correspondence regarding this work should be addressed to Ethan M. McCormick, Methodology & Statistics Department, Institute of Psychology, Leiden University, Leiden 2333 AK, Netherlands. Email: e.m.mccormick@fsw.leidenuniv.nl. The code required to reproduce all analyses, as well as more-extensive derivations, can be found at https://doi.org/10.17605/OSF.IO/5DMY3.

Abstract

In the modeling of change over time, there is often a disconnect between developmental theories advanced in substantive research and statistical models specified in longitudinal analysis. That is, theory is understood and advanced in terms of meaningful developmental quantities (e.g., peaks, inflections, timing, and tempo) while common polynomial models estimate the effect of powered terms of time in a linear, additive form. This linear parameterization approach has many advantages, especially its computation efficiency in obtaining stable results, but the quantities estimated in these models are often difficult to directly connect to theoretical ideas of change over time. To bridge the gap between estimation and theory development, I propose a series of approaches for linear estimation with nonlinear inference (LENI), where the results of the stable, easily-estimated linear model are converted through a set of principled transformation functions into nonlinear estimates which align more closely with theoretical quantities of interest. I first lay out how to derive these interpretable nonlinear parameters, then show how to transform the results of the linear model – including fixed and random effects and the conditional effects of covariates – into the effects we would have obtained by fitting a nonlinear version of the model. I conclude by summarizing a linearized structural equation model approach which can be flexibly applied to model any known nonlinear target function into a linearly-estimable model. I conclude with recommendations for applied researchers and directions for fruitful future work in this area.

1 Introduction

The issue of interpretability is one of widespread concern in modeling outcomes of interest across many different disciplines. At the extremes, approaches like large language or deep learning models can be composed of hundreds of thousands of parameters, none of which can be meaningfully interpreted individually. These issues are not isolated to large machine learning approaches, however, and filter down to more common linear models within the social and behavioral sciences. Often, standard models are fit using equations that optimize numerical stability and have relatively simple fit functions – often a crucial feature that allowed these models to be feasibly estimated before modern computing power – rather than the interpretability of the parameters obtained. While simple linear effects solve for both ease of fit and interpretability, they are limited in their ability to test more complex and specific substantive hypotheses (e.g., timing, inflection points). Some fields have sought a balance between these extremes. For instance, the field of cognitive computational modeling has developed a wide array of mathematical expressions that seek to describe mental state representations guiding overt action (Farrell & Lewandowsky, 2018; Wilson & Collins, 2019). While these expressions are relatively complex, a strong emphasis of computational modeling is that the parameters of these models are linked with specific cognitive or behavioral processes. As such, individual (or group) differences in these parameters can ideally be linked directly back to cognitive or neural processes of interest (e.g., Mareschal & Thomas, 2007; Patzelt et al., 2018; Pleskac et al., 2019; Wilson & Collins, 2019). Outside of these computational models, however, there has only been slow progress in the adoption of interpretable parameter models.

The bio-behavioral and clinical sciences are dominated by the use of linear models – specifically models which are linear with respect to the parameters, including popular polynomial models (e.g., quadratic and cubic) which chart out a non-linear relationship. These linear parameter models have many attractive features for estimating relationships between variables – they are identified across an infinite range of parameter values associated with the predictors (i.e., β's), they are purely additive in form, and they are widely implemented in available software. As such, results can be easily and efficiently obtained, often with closed-form solutions (e.g., ordinary least squares regression) or other well-behaved likelihood functions. Unfortunately, the parameters of these models also often do not test specific hypotheses that are of substantive interest (Cudeck & du Toit, 2002; Preacher & Hancock, 2015; Ram & Grimm, 2007). To address these issues, prior work has derived alternative nonlinear expressions (Cudeck & du Toit, 2002; McNeish et al., 2021), or worked to reparameterize known nonlinear expressions into linear forms (Blozis, 2004; Feng et al., 2019; Grimm et al., 2013; Johnson & Hancock, 2019; Preacher & Hancock, 2012, 2015; Zhang et al., 2012). Unfortunately, these models

https://doi.org/10.1017/psy.2024.2 Published online by Cambridge University Press

do not appear as standard options in major software packages, and many applied researchers remain unaware of their potential utility. Additionally, nonlinear expressions present additional estimation challenges – especially in growth modeling contexts with random effects – and for these reasons, largely remain the provenance of researchers with training in and access to more advanced statistical methods and software options.

Here, I address several extant issues for formulating polynomial models with interpretable and meaningful parameters, with an eye for expanding the utility and accessibility of these approaches. First, I review a history of interpretable parameter models and walk through a general approach for deriving new parameters of interest, highlighting the quadratic form outlined by Cudeck & du Toit (2002). I then extend these principles and derive two alternative forms of a cubic polynomial with meaningful parameters and show how this new model is related to the standard linear parameter version. I also discuss a multiphase version of this model which can serve as an approximation of S-shaped nonlinear models (e.g., logistic functions). To address the common estimation issues with nonlinear versions of these alternative models, I lay out an approach of linear estimation with nonlinear inference (LENI), where the standard linear parameter model is estimated, and then results are transformed *post hoc* into the parameters of interest from the nonlinear alternative models. I derive transformation equations for the point estimates and standard errors of fixed, random, and conditional effects, allowing inferences to be made on the meaningful parameters as if we had directly estimated the nonlinear equation. Finally, extending prior work (Feng et al., 2019; Preacher & Hancock, 2015), I derive a linearized structural equation model for all of the models discussed, focusing on implementation in freely-available software. Throughout, I discuss these models largely in the context of growth models using mixed-effects multilevel or latent curve structural equation models (McCormick et al., 2023; McNeish & Matta, 2018; Meredith & Tisak, 1990), with artificial and real data examples, but the discussion of fixed effects derivations would apply equally to traditional regression analysis with no additional variance components. I then end with a discussion of implementation options for applied researchers and open avenues for future work in this area.

2 Interpretable Parameter Models

In efforts to address mismatches between theoretical and statistical models, prior work has focused on deriving new expressions which equivalently trace out the same nonlinear curves as standard polynomial models, but using parameters that more-closely match theoretical quantities of interest. A prime example of these efforts is work by Cudeck & du Toit (2002), who derived a new quadratic expression with meaningful parameters. The familiar linear parameter version of the quadratic for

https://doi.org/10.1017/psy.2024.2 Published online by Cambridge University Press

repeated measures outcome y_{ti} is

$$y_{ti} = \beta_0 + \beta_1 x_{ti} + \beta_2 x_{ti}^2 \tag{1}$$

for person i at time t, where β_1 and β_2 are the linear and quadratic effect of the covariate x_{ti} respectively, while the alternative expression takes the following form (Cudeck & du Toit, 2002):

$$y_{ti} = \alpha_y - (\alpha_y - \alpha_0)\left(\frac{x_{ti}}{\alpha_x} - 1\right)^2 \tag{2}$$

Here, α_x and α_y represent the (x, y) location for the vertex (i.e., peak or trough) of the quadratic parabola, while α_0 is equivalent to β_0 in Equation 1 – that is, the predicted level of y when $x = 0$. This alternative quadratic model is nonlinear with respect to the parameters (e.g., α_x is in the denominator) but otherwise describes the exact same parabolic shape as in Equation 1 (note that we can ignore the residual term, ε_{ti}, here when discussing alternative models because if we have done our job correctly, it will be identical across model versions). The advantage of Equation 2 is that the α_x and α_y parameters give a direct estimate of the location of the vertex, which in developmental contexts might relate to the timing of changes in sensitivity to the external environment (Braams et al., 2015; McCormick et al., 2021; Nunes et al., 2020; Orben et al., 2022; Shaw et al., 2008; Somerville et al., 2013), or reflect the optimal arousal levels or dosage needed to maximize the desired response (Chaiken, 1994; Cudeck & du Toit, 2002; Preacher & Hancock, 2015). Cudeck & du Toit (2002) also described a second alternative model

$$y_{ti} = \alpha_0 - \gamma\left[\left(\frac{x_{ti}}{\alpha_x} - 1\right)^2 - 1\right] \qquad \text{or} \qquad y_{ti} = \alpha_y - \gamma\left(\frac{x_{ti}}{\alpha_x} - 1\right)^2 \tag{3}$$

where γ represents the difference in the level of the outcome y between the intercept and the vertex ($\gamma = \alpha_y - \alpha_0$) rather than estimating α_y (Equation 3, left) or α_0 (right) directly. To highlight the fact that we can retain any combination of meaningful parameters that we wish, we can also lay out another alternative quadratic form below where we estimate α_x and α_y but retain β_2 as α_c (i.e., half of the acceleration, which controls the degree of curvature for the parabola) rather than α_0:

$$y_{ti} = \alpha_y + \alpha_c\left(x_{ti} - \alpha_x\right)^2 \tag{4}$$

Detailed derivations for all of these models are available in the Supplemental Material.

https://doi.org/10.1017/psy.2024.2 Published online by Cambridge University Press

2.1. Deriving New Quantities

While these parameters represent mathematically meaningful points on a quadratic parabola (e.g., location of the vertex), we might also wish to understand something about the curve at a particular value of x. Examples of this kind of question include things like the number of words acquired by 20 months-of-age (Huttenlocher et al., 1991), or level of drug and alcohol use upon entry to university (Derefinko et al., 2016). This is an interesting inversion of time-to-criterion models (Johnson & Hancock, 2019), where the focus is on estimating how long it takes for some outcome to reach a pre-defined level, or nonlinear models for measuring potential (e.g., McNeish & Dumas, 2017) where the rate of approach towards an average or individual-level learning "capacity" is of interest. Here, we might instead be interested in group- or individual-level status achieved on some outcome by a certain developmental milestone. We can use this example to highlight how to go about translating new quantities of interest into statistical parameters that we can estimate from alternative model expressions. For a quadratic of the form derived by Cudeck & du Toit (2002), we can express the level of the outcome at any pre-defined value of $x = s$ through the parameter α_s with the equation

$$y_{ti} = \alpha_s - (\alpha_s - \alpha_0)\left(\frac{x_{ti} - s}{\alpha_x} - 1\right)^2 \tag{5}$$

To emphasize, s is predefined when estimating the model based on a theoretically interesting value of x, and not estimated as a unique parameter. This could be accomplished alternatively as part of a data management step, where s is subtracted from x before entering the model. In that approach, the original expression in Equation 2 should be used, as the identity of α_0 has been changed through centering (Aiken & West, 1991) rather than through the model equation.

It might occur to the reader that this new parameter α_s in particular is easily achieved by simply entering the theoretically interesting value of $x = s$ into the linear parameter quadratic equation (Equation 1) and generating a predicted value of y. However, this approach only returns a point estimate for α_s and not a standard error or random effect variance, therefore should preclude us from making inferences on that predicted value (although applied research frequently does so). For instance, many developmental applications calculate and interpret the vertex of a quadratic trajectory (Eggleston et al., 2004; Giedd et al., 2015; Lenroot et al., 2007; LeWinn et al., 2017) or inflection of a cubic (LeWinn et al., 2017; Mills et al., 2016) as a quantity of interest without also obtaining a standard error. Without some measure of uncertainty, interpreting these point estimates can lead to erroneous conclusions about developmental timing, or differences among individuals or groups thereof (Giedd et al., 2015; Karriker-Jaffe et al., 2008; Pfefferbaum et al., 2018) . We will return to the issue of obtaining these measures of uncertainty (see Linear Estimation, Nonlinear Inference (LENI)).

https://doi.org/10.1017/psy.2024.2 Published online by Cambridge University Press

When considering new interpretable quantities, it is important that we be able to generate specific mathematical definitions. Assuming that we can do this, we can generate an infinite combination of meaningful model expressions that directly estimate parameters of interest. In the following section, I will build on the principles outlined here to derive nonlinear alternatives to the standard cubic model, with interpretable parameters linked to meaningful developmental phenomena.

3 Deriving Alternative Cubic Models

As a natural extension of the quadratic, the cubic polynomial is another option for modeling nonlinear change over time. The common expression of the cubic model is as follows for y for person i at time t:

$$y_{ti} = \beta_0 + \beta_1 x_{ti} + \beta_2 x_{ti}^2 + \beta_3 x_{ti}^3 \tag{6}$$

which is linear with respect to the parameters. However, like the quadratic expression (Cudeck & du Toit, 2002), the parameters of this model are not readily identifiable with a specific developmental feature that might be of theoretical interest. The lower order terms β_1 and β_2 are conditional effects specific to where $x = 0$, and β_3 is not easily expressible in meaningful terms (i.e., the change in the acceleration of the curve) for most researchers. Here I will draw out two alternative expressions of the linear parameter cubic model with theoretically-interesting parameters. In its initial form, this model expression is best suited to cases where both extrema (i.e., local minimum and maximum) occur within the range of x, although that condition is not necessary for this alternative expression to obtain meaningful results. For modeling increases which subsequently plateau (e.g., Somerville et al., 2013), the multiphase form outlined in Multiphase Cubic Model for S-Shaped Trajectories would likely be preferable.

3.1. Expressing Meaningful Model Parameters

For defining the alternative form of the cubic formula, we can consider 4 parameters of interest, two location parameters – x_N and y_N – which locate the inflection point (N) of the curve, and two stretch parameters – δ (delta) and h (height) – which determine the horizontal and vertical distances respectively between the inflection point and the extrema of the cubic function. These parameters have well-defined geometric properties and can be used to solve polynomials in terms of their roots (Nickalls, 1993). To obtain x_N, we take the second derivative of the linear parameter model

$$\frac{\mathrm{d}^2 f}{\mathrm{d}x^2} = 2\beta_2 + 6\beta_3 x \tag{7}$$

https://doi.org/10.1017/psy.2024.2 Published online by Cambridge University Press

which can be set to zero and rearranged such that

$$x_N = \frac{-\beta_2}{3\beta_3} \tag{8}$$

which nicely resembles the form of α_x from the quadratic expression (Cudeck & du Toit, 2002). For full details on these parameter derivations, see the Supplemental Material.

The value obtained at the inflection point (y_N) is expressed by evaluating Equation 6 at Equation 8 and simplifying

$$\begin{aligned} y_N &= \beta_0 + \beta_1 x_N + \beta_2 x_N^2 + \beta_3 x_N^3 \\ &= \beta_0 + \beta_1 \left(\frac{-\beta_2}{3\beta_3}\right) + \beta_2 \left(\frac{-\beta_2}{3\beta_3}\right)^2 + \beta_3 \left(\frac{-\beta_2}{3\beta_3}\right)^3 \\ &= \beta_0 - \frac{\beta_1\beta_2}{3\beta_3} + \frac{2\beta_2^3}{27\beta_3^2} \end{aligned} \tag{9}$$

Now that we have identified the location parameters of the inflection point for the cubic function, we need to define the stretch parameters. The most convenient way to do so is to lay out expressions which identify the extrema – or local maximum and minimum – and defining our parameters as the distances between the inflection point and extrema. A cubic has two extrema locations ($x_{extrema}$), but since the function is symmetric, we only need to derive a single expression for δ or h that will define the positive and negative distance. A convenient way to obtain δ is to set the first derivative of the cubic equal to 0

$$\frac{\mathrm{d}f}{\mathrm{d}x} = \beta_1 + 2\beta_2 x + 3\beta_3 x^2 = 0 \tag{10}$$

and solve using the quadratic formula where $a = 3\beta_3$, $b = 2\beta_2$, and $c = \beta_1$

$$x_{extrema} = \frac{-(2\beta_2) \pm \sqrt{(2\beta_2)^2 - 4(3\beta_3)(\beta_1)}}{2(3\beta_3)} \tag{11}$$

which, through substitution and simplification (see here), results in

$$x_{extrema} = x_N \pm \frac{\sqrt{\beta_2^2 - 3\beta_3\beta_1}}{3\beta_3} \tag{12}$$

which means that

$$x_{extrema} = x_N \pm \delta \tag{13}$$

and that

$$\delta = \frac{\sqrt{\beta_2^2 - 3\beta_3\beta_1}}{3\beta_3} \tag{14}$$

https://doi.org/10.1017/psy.2024.2 Published online by Cambridge University Press

We can enforce positive distance values by defining $\delta = \sqrt{\delta^2}$ where

$$\delta^2 = \frac{\beta_2^2 - 3\beta_3\beta_1}{9\beta_3^2} \tag{15}$$

Finally, by defining h as the difference between y obtained at the extrema ($y_{x_N \pm \delta}$) and the inflection point (y_N)

$$\begin{aligned} h &= y_{x_N+\delta} - y_N \\ &= [\beta_0 + \beta_1(x_N + \delta) + \beta_2(x_N + \delta)^2 + \beta_3(x_N + \delta)^3] - \\ &\qquad [\beta_0 + \beta_1 x_N + \beta_2 x_N^2 + \beta_3 x_N^3] \end{aligned} \tag{16}$$

Simplifying this expression (see Supplemental for full details) results in

$$h = -2\beta_3\delta^3 \tag{17}$$

One final meaningful parameter we might define is the slope of the cubic function at the inflection point, β_N. Here solve the first derivative of Equation 6 at the inflection point and substitute

$$\begin{aligned} \frac{\mathrm{d}f}{\mathrm{d}x} &= \beta_1 + 2\beta_2 x_N + 3\beta_3 x_N^2 \\ &= (3\beta_3 x_N^2 - 3\beta_3\delta^2) + 2(-3\beta_3 x_N)x_N + 3\beta_3 x_N^2 \\ \beta_N &= -3\beta_3\delta^2 \end{aligned} \tag{18}$$

A schematic of these parameters is displayed in Figure 1. Note that the sign of h (or alternatively β_N) determines whether the cubic has an overall increasing ($h < 0$) or decreasing ($h > 0$) function (the direction of change locally between the extrema is opposite of change globally).

As we saw with the quadratic, we have a great deal of flexibility in terms of which nonlinear parameters we wish to use, so long as we can derive these transformation functions. Importantly, however, between any two given sets of parameters, the transformation functions are unique and deterministic because we must maintain the same functional form for the equations defined by the two parameter sets to be equivalent.

https://doi.org/10.1017/psy.2024.2 Published online by Cambridge University Press

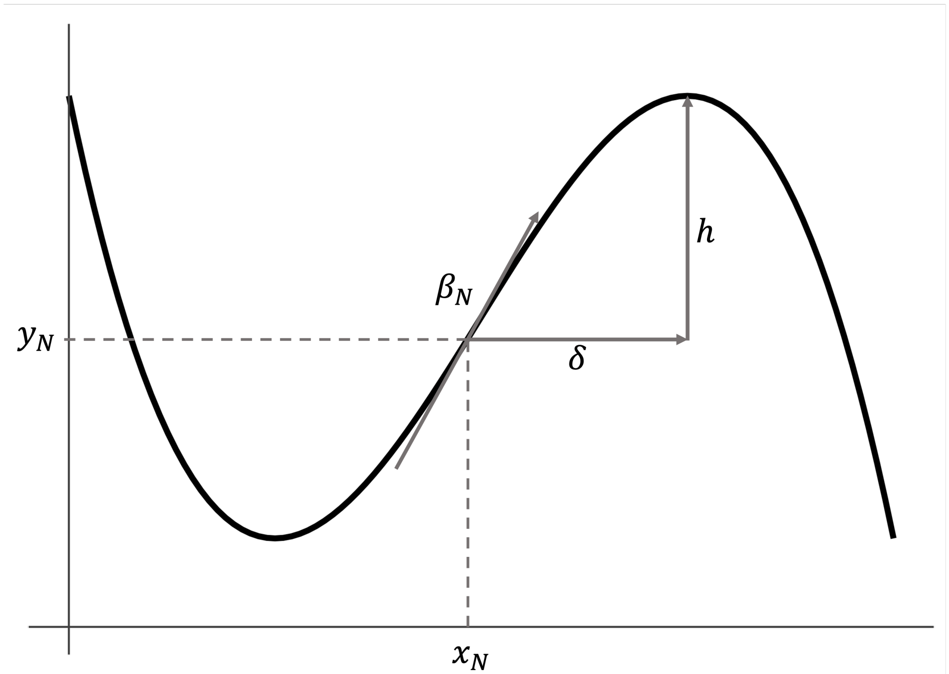

Figure 1. The interpretable parameters are superimposed on an idealized cubic function. Location parameters (x_N and y_N) are indicated with dashed lines, stretch parameters (δ and h) are indicated by single-headed arrows over the relevant distance, and the slope of the tangent at the inflection point (β_N) is indicated by a single-headed arrow tracing the tangent line at that point.

3.2. Defining the Nonlinear Model

New we can substitute the expressions for the meaningful parameters (x_N, y_N, δ, h) into the linear parameter model (Equation 6) to derive the nonlinear expression for an interpretable cubic.

$$\begin{aligned} y &= \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 \\ &= \left[y_N + 3\beta_3 x_N \delta^2 - \beta_3 x_N^3\right] + \left[3\beta_3 x_N^2 - 3\beta_3 \delta^2\right] x + \left[-3\beta_3 x_N\right] x^2 + \beta_3 x^3 \\ &= y_N + \beta_3 \left[\left(x - x_N\right)^3 - 3\delta^2 \left(x - x_N\right)\right] \\ &= y_N + \left(\frac{-h}{2\delta^3}\right) \left[\left(x - x_N\right)^3 - 3\delta^2 \left(x - x_N\right)\right] \end{aligned} \tag{19}$$

which results in

$$y = y_N - \left(\frac{h}{2}\right)\left[\left(\frac{x - x_N}{\delta}\right)^3 - 3\left(\frac{x - x_N}{\delta}\right)\right] \tag{20}$$

Alternatively, substituting Equation 18 instead of Equation 17 results in

$$y = y_N - \left(\frac{\beta_N \delta}{3}\right)\left[\left(\frac{x - x_N}{\delta}\right)^3 - 3\left(\frac{x - x_N}{\delta}\right)\right] \tag{21}$$

For complete details on how we arrive at these expressions, see the Supplemental Material. Note

https://doi.org/10.1017/psy.2024.2 Published online by Cambridge University Press

that these expressions are equivalent to the linear parameter model (Equation 6), and should fit the same functional form. We can use a real data example to see how this model fits in empirical settings.

3.2.1. Data Example 1: Negative Affect and Aging

To briefly highlight how Equation 20 produces meaningful inferences, we can turn to an empirical example. Here I extracted aggregated scatter plot data from Teachman (2006) (see Figure 2) and fit a cubic model using Equation 6 and Equation 20 to model age-related differences in negative affect in adults, ages 17-93.

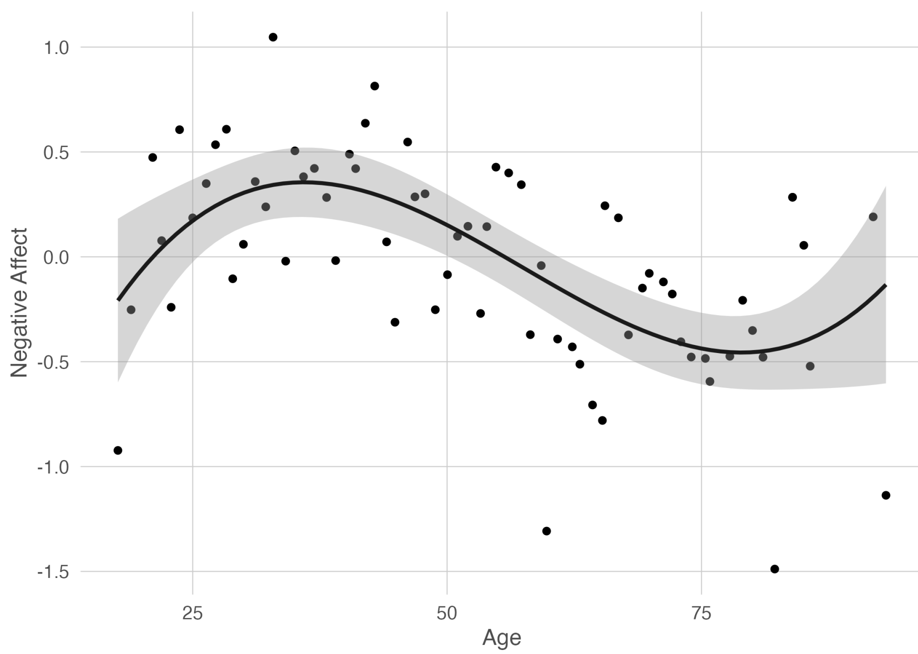

Figure 2. Cubic age-related changes in Negative Affect across the adult lifespan. The fitted cubic relationship is plotted over the raw data points.

Results of the two models are presented in Table 1. While both models have the same fit to the data (see log-likelihood and BIC values; the standard R^2 is only available in the linear parameter model), it is immediately apparent that the parameter values in the Nonlinear Parameter Model have intuitive meaning – the inflection point in negative affect is around 57 years-of-age and decreases by approximately 0.4 units across the 21 years between either extrema ($x_N \pm \delta$) and the inflection point. By contrast, the values of the Linear Parameter Model are virtually meaningless from a substantive standpoint without plotting. To emphasize, neither model is *wrong* – indeed they are identical in the model-implied curve – however, the nonlinear parameter model is more useful if our goal is to interpret meaningful points in the age-related change with precision (rather than the frequent practice of "eyeballing" a plot) and crucially, the appropriate level of uncertainty (via the standard errors).

https://doi.org/10.1017/psy.2024.2 Published online by Cambridge University Press

Table 1

Fitting Linear and Nonlinear Parameter Cubic Models

	Linear Parameter Model		Nonlinear Parameter Model
β_0	-2.288^{**} (0.802)	x_N	57.428^{***} (2.162)
β_1	0.174^{**} (0.052)	y_N	-0.050 (0.051)
β_2	-0.004^{***} (0.001)	δ	21.508^{***} (1.711)
β_3	$2.04 \times 10^{-5**}$ (6.10×10^{-6})	h	-0.406^{***} (0.068)
Num.Obs.	69		69
R^2	0.355		
ℓ	-33.286		-33.286
Num.Params.	5		5
BIC	87.7		87.7

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$; ℓ is the log-likelihood, k is the number of model parameters.

3.3. Multiphase Cubic Model for S-Shaped Trajectories

While the cubic model in its full form offers an ability to derive meaningful conclusions about developmental patterns, cubic polynomials can also be used to capture plateaus (e.g., Somerville et al., 2013) because of the saddle-point of the cubic. However, there is often less made inferentially about subsequent acceleration after that plateau, although any full cubic will continue off into infinity in principle. If we want better approximate developmental plateauing, or other S-shaped functional forms (e.g., logistic, Gompertz), then a multiphase (or piecewise) version of the cubic may offer an attractive alternative. McNeish et al. (2021) showed that two reparameterized quadratics (Cudeck & du Toit, 2002) can be linked together at the inflection point of a hypothetical S-shaped functional form in a multiphase polynomial model (Cudeck & Klebe, 2002; Flora, 2008). Additional pieces are specified such that once these linked quadratics reach their model-implied vertices (α_x), they stay fixed at the y value obtained at the vertex (α_y, see McNeish et al., 2021, for full details).

We can use a similar approach here, taking advantage of the interpretable parameters I derived in Equation 20. If we take the case of a 4-parameter logistic function as an exemplar of an S-shaped curve, this function is defined for outcome y for person i at time t by

$$y_{ti} = A_{lower} + \frac{A_{upper} - A_{lower}}{1 + \left(\frac{x_{ti}}{x_N}\right)^{-Hill}} \tag{22}$$

Where A_{upper} and A_{lower} are the upper and lower asymptote, x_N is the x-location of the inflection point, and $Hill$ is related to the steepness of change at the inflection point. These parameters conceptually

https://doi.org/10.1017/psy.2024.2 Published online by Cambridge University Press

map on to several of the interpretable cubic parameters (Equation 20 or Equation 21) nicely, with x_N capturing the inflection point, β_N capturing the rate of change at that point, and the upper and lower asymptotes approximated by $y_N \pm h$. I use "approximate" advisedly because the multiphase cubic actually obtains the value of $y_N \pm h$, while Equation 22 only ever infinitely approaches A_{upper} and A_{lower}. We can define the multiphase cubic function for the same y_{ti} as

$$y_{ti} = \begin{cases} y_N - h, & \text{if } x_{ti} \leq x_N - \delta \\ y_N - \left(\frac{h}{2}\right)\left[\left(\frac{x_{ti}-x_N}{\delta}\right)^3 - 3\left(\frac{x_{ti}-x_N}{\delta}\right)\right], & \text{if } (x_N - \delta) < x_{ti} < (x_N + \delta) \\ y_N + h, & \text{if } x_{ti} \geq x_N + \delta \end{cases} \tag{23}$$

or alternatively

$$y_{ti} = \begin{cases} y_N - \frac{2}{3}\beta_N \delta, & \text{if } x_{ti} \leq x_N - \delta \\ y_N - \left(\frac{\beta_N \delta}{3}\right)\left[\left(\frac{x_{ti}-x_N}{\delta}\right)^3 - 3\left(\frac{x_{ti}-x_N}{\delta}\right)\right], & \text{if } (x_N - \delta) < x_i < (x_N + \delta) \\ y_N + \frac{2}{3}\beta_N \delta, & \text{if } x_{ti} \geq x_N + \delta \end{cases} \tag{24}$$

A schematic of the multiphase model with the relevant features highlighted can be seen in Figure 3.

3.3.1. Data Example 2: Pubertal Development

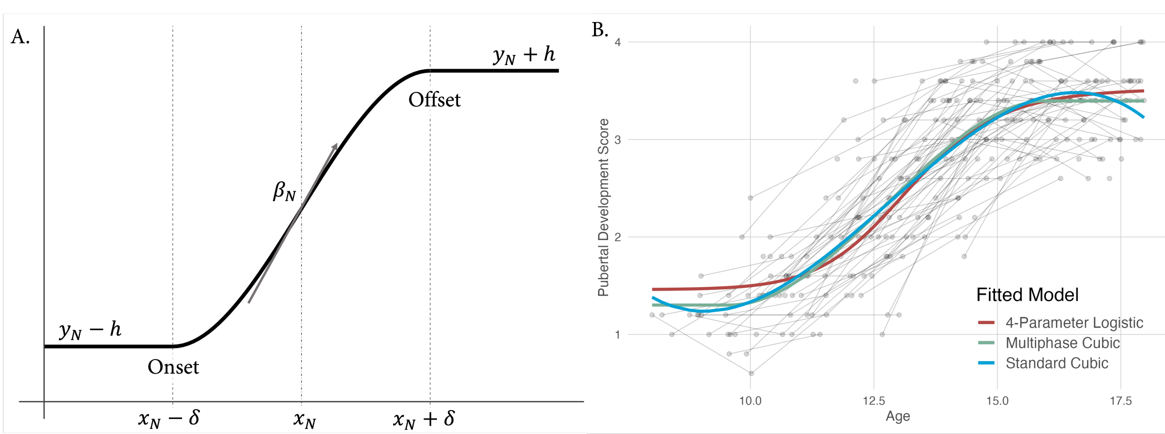

Figure 3. Multiphase Cubic Model. A) Using a multiphase cubic function (Equation 23 or Equation 24), we can approximate an S-shaped function with three components. The component between onset and offset is defined by the cubic function, while outside this range is defined by $y_N \pm h$. B) Alternative models were fit to pubertal developmental data, including a 4-parameter logistic (red), the multiphase cubic (green) and standard linear parameter cubic model (blue). The logistic and multiphase models do not enforce continued acceleration at the edges of the curve – an advantage over the standard cubic. While the logistic model continues increasing asymptotically, the multiphase cubic models stability outside of the cubic extrema.

To illustrate this model, I drew pubertal data from an accelerated longitudinal study of adolescent

https://doi.org/10.1017/psy.2024.2 Published online by Cambridge University Press

development (BrainTime; Braams et al., 2015; McCormick et al., 2021). Adolescents were assessed up to three times at two-year intervals and self-reported on their level pubertal development (Petersen et al., 1988). Prior work (Braams et al., 2015) determined that a standard cubic model was the best fit to these data compared with linear and quadratic polynomial alternatives. Here, I fit three random-intercept growth models to these data: 1) a 4-parameter logistic curve (4-PL; Equation 22), 2) the multiphase cubic model (Equation 23), and 3) a standard linear parameter version of the cubic (Equation 6). I modeled A_{lower}, y_N, and β_0 respectively as random terms to equate the complexity of the random effects structure for all three models (code and full results for all three models are available in the Supplemental Material).

The primary fixed effects of interest in each model are presented in Table 2. As reflected in Figure 3, the three models achieve reasonably similar results in terms of fitting a curve to the data. However, note that the BIC indicates that the 4-PL and multiphase-cubic models both fit better than the standard cubic function, due in large part to their different edge behavior (they asymptote or offset rather than continuing to accelerate). However, from a model interpretation stand-point, the 4-PL and multiphase cubic have clearly interpretable parameters that match substantively interesting features of the developmental trajectory. The estimates between these two models largely draw the same conclusions about both the levels of y at the asymptotes and where the inflection point is located. I additionally fit Equation 24 to compute β_N (0.491, SE = 0.034, p < 0.001), which is expressed directly as the slope of the tangent at the inflection point unlike the $Hill$ parameter, which controls the overall shape of the curve in the 4-PL model but is not in easily-interpretable units.

Table 2

Alternative S-Shaped Trajectories

4-Parameter Logistic		Multiphase Cubic		Standard Cubic	
A_{lower}	1.460*** (0.107)	x_N	12.780*** (0.142)	β_0	18.865*** (3.305)
A_{upper}	3.524*** (0.069)	y_N	2.348*** (0.054)	β_1	−4.757*** (0.767)
x_N	13.210*** (0.150)	δ	3.198*** (0.298)	β_2	0.406*** (0.058)
$Hill$	14.196*** (2.172)	h	1.047*** (0.054)	β_3	−0.011*** (0.001)
Marg. ℓ	−180.43		−167.51		−186.86
k	6		6		6
BIC	394.54		368.70		407.40

Note: * p < 0.05, ** p < 0.01, *** p < 0.001; ℓ is the marginal (Marg.) log-likelihood, k is the number of model parameters.

There are some additional theoretical advantages to the multiphase cubic compared with the 4-PL model. Unlike the asymptotic nature of the logistic, the multiphase cubic here and quadratic outlined

https://doi.org/10.1017/psy.2024.2 Published online by Cambridge University Press

by McNeish et al. (2021) obtain the minimum and maximum model-implied values. For outcomes measured on some natural scale – as opposed to the probability scale where asymptotic behavior is desirable – these multiphase models likely represent more realistic developmental conditions, especially when the scale has a natural or measurement boundary (i.e., floor or ceiling ; Feng et al., 2019). For instance, pubertal development is not an infinitely-occurring process – that is pre- and post-puberty are meaningful terms. Other processes such as cortical thinning (Fuhrmann et al., 2022; Mills et al., 2016; Tamnes et al., 2017) also reach some (at least local) minimum before other processes take over into the adult phase of development. Here the multiphase cubic could even be extended, with additional growth components modeled in the offset period, or to model processes of punctuated equilibrium, where change occurs in bursts followed by periods of stability, by linking several multiphase cubic functions. This approach is familiar in generalized additive (mixed) models where cubic splines have featured prominently (e.g., Wood, 2000). However, unlike GAMs, which focus on description and prediction, the multiphase cubic model maintains a primary focus on an explanatory model with interpretable parameters. Nevertheless, GAMs can be potentially valuable tools for conducting sensitivity analyses on interpretable parameter models to ensure that data-driven curves descriptively resemble the *a priori* functional forms fit to the data. Substantial departure from these imposed forms can reveal whether theory-driven models are indeed suitable to model the data at hand. For the various empirical examples, this GAM-based sensitivity approach was used to confirm the suitability of the functional forms used in the analysis (see the Supplemental Code for these results).

4 Linear Estimation, Nonlinear Inference (LENI)

Thus far, I have outlined a history of efforts to improve the interpretability of models by deriving alternative forms with theoretically meaningful parameters, including applications of the reparameterized quadratic (Cudeck & du Toit, 2002; Grimm et al., 2013; McNeish et al., 2021; Preacher & Hancock, 2015) and how to build on these principles to derive new quantities of interest. I then extended these ideas to the cubic model, deriving a set of new expressions with 5 interpretable parameters (Equation 20 and Equation 21), and a multiphase version of the cubic to model S-shaped curves. However, despite clear theoretical advantages and substantive interest in things like developmental "peaks" in the literature, these alternative models have remained largely restricted to advanced applications. In the remainder of this manuscript, I will lay out current challenges for the adoption of meaningful parameter models, and then offer a set of analytic approaches to address key limitations. The techniques discussed – which involve linear estimation with nonlinear parameter interpretation (LENI) – apply to both standard regression analysis, as well as major classes of longitudinal modeling, including

https://doi.org/10.1017/psy.2024.2 Published online by Cambridge University Press

mixed-effect (multilevel) and structural equation growth models (McCormick et al., 2023; McNeish & Matta, 2018).

4.1. Limitations of Nonlinear Models

While there are strong theoretical arguments for the use of interpretable parameter models, a number of technical challenges have historically been barriers to their widespread adoption. Many of these challenges stem from the relative difficulty of estimation in nonlinear parameter models compared to their linear parameter alternatives. First, linear parameter models are defined across all possible values of the parameters, whereas many nonlinear equations are undefined at specific values. For instance, the alternative quadratic (Cudeck & du Toit, 2002), (Equation 2) is undefined at $\alpha_x = 0$, and the alternative cubic (Equation 20) is undefined at $\delta = 0$. As parameter estimates approach these values, the model-implied values of the outcome become more unstable. By contrast, in the linear parameter models (Equation 1; Equation 6), any parameter $\beta_p = 0$ merely indicates an absence of that component of the polynomial function. This issue is not universal for alternative expressions, as the additional alternative quadratic derived in Equation 4 is defined at all parameter values, but it is a common problem with nonlinear equations.

An additional disadvantage of nonlinear alternative models is a lack of clearly hierarchical structure to the parameters of the model, both for the fixed effects and for the ordering of random effects (McNeish et al., 2021). For instance, as mentioned above, when a linear parameter $\beta_p = 0$, this is informative about the complexity of a curve. If $\beta_3 = 0$, then the cubic function devolves back to a quadratic one, and so forth. By contrast, the nonlinear alternative (Equation 20) cannot be reduced to a quadratic by setting any single parameter to 0, nor can we do nested model comparisons (i.e., likelihood ratio tests) to determine the optimal polynomial complexity of the curve. The lack of hierarchical structure additionally complicates model specification when the random effects structure needs to be constrained to achieve convergence (McNeish & Bauer, 2022). In the linear parameter model, we would typically constrain random effects from the highest-order (τ_{β_3}) to the lowest (τ_{β_1}). However, in the interpretable parameter model, the ordering of which random effects to constrain is less clear. The parameter transformation equations might give us some sense of a reasonable ordering – for instance, y_N is the only interpretable parameter that is a function of the intercept (β_0), and x_N is the only parameter that is **not** a function of β_1. This might suggest that x_N be constrained first, and y_N always be modeled in a random effects model, but this ordering is much less clear than in the linear parameter case.

Finally, two related challenges stem from the relative difficulty of estimating nonlinear parameter models compared to linear parameter alternatives. Even in the relatively simple regression case, nonlin-

https://doi.org/10.1017/psy.2024.2 Published online by Cambridge University Press

ear models lack a closed-form solution and require iterative fitting procedures (Fox & Weisberg, 2011). These estimation challenges only increase when fitting mixed-effects or structural equation growth models, where likelihoods in nonlinear models can be more poorly-behaved (Blozis, 2007) and more likely to result in local solutions, in addition to the challenges of improper solutions (e.g., $\delta \to 0$). Due in part to these challenges, nonlinear parameter models are often not easily accessible for applied researchers within widely-available software packages, requiring custom syntax and difficulty implementing these non-standard model equations. Paired with less wide-spread knowledge of alternative interpretable models, this creates a negative feedback loop, where software providers are not incentivized to develop resources for models that users are unlikely to fit.

The linear estimation with nonlinear inference (LENI) approach laid out in the following sections addresses each of these challenges, drawing on the strengths of estimation and simplicity in the linear parameter model but without sacrificing the interpretability of the nonlinear alternative models.

4.2. LENI for Fixed Effects

The idea of deriving some meaningful quantity from an estimated linear parameter model is not without precedence. Substantive applications have occasionally used the formulations for α_x and δ to identify the peaks (e.g., Braams et al., 2015; LeWinn et al., 2017) , especially when comparing trajectories across groups (often between males and females; e.g., Giedd et al., 2015) . However, these formulations only return a point estimate without an associated standard error. Nor does it allow for conditional effects of covariates of interest predicting the interpretable parameters. In the following sections, I lay out a general approach for deriving the meaningful parameters of the nonlinear alternative models discussed throughout, entirely from the results of a linear parameter model for the same functional form (quadratic or cubic). Note that for the fixed effects, these transformations apply equally to standard regression, mixed-effects (or multilevel) models, and structural equation models. For SEM growth models specifically, we can return to other, more direct, LENI approaches at the end of this treatment.

For considering fixed effects, we will estimate the following linear parameter models. Here we will consider a standard regression model with no random effects, but these approaches generalize to models with random effects structures without alteration (as shown in later sections). Thus for person i,

$$
\begin{aligned}
y_i &= \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \varepsilon_i \\
y_i &= \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \beta_3 x_i^3 + \varepsilon_i \qquad \text{where, } \varepsilon_i \sim N(0, \sigma^2)
\end{aligned}
\tag{25}
$$

https://doi.org/10.1017/psy.2024.2 Published online by Cambridge University Press

4.2.1. Point Estimates

To obtain point estimates for the interpretable parameters, we need only to apply the relationships derived previously by Cudeck & du Toit (2002) and here in Deriving Alternative Cubic Models. Namely, for the alternative quadratic model, the expressions for the interpretable parameters are as follows

$$\alpha_0 = \beta_0 \qquad \alpha_x = \frac{-\beta_1}{2\beta_2} \qquad \alpha_y = \beta_0 - \frac{\beta_1^2}{4\beta_2} \qquad \gamma = -\frac{\beta_1^2}{4\beta_2} \qquad \alpha_c = \beta_2 \tag{26}$$

and for the alternative cubic model, the expressions are

$$\begin{gathered} x_N = \frac{-\beta_2}{3\beta_3} \qquad y_N = \beta_0 - \frac{\beta_1\beta_2}{3\beta_3} + \frac{2\beta_2^3}{27\beta_3^2} \qquad \delta = \frac{\sqrt{\beta_2^2 - 3\beta_3\beta_1}}{3\beta_3} \\ h = -2\beta_3\delta^3 \qquad \beta_N = -3\beta_3\delta^2 \end{gathered} \tag{27}$$

Nothing more need be done to obtain these point estimates beyond some algebra. One note of caution, however, is that we can in theory compute more interpretable parameter point estimates than we estimate in the linear parameter model. On one hand, the interpretable parameters are not independent (there is an especially tight relationship between γ and α_y, and between h and β_N for instance), so we would just be repackaging the same information and thus only support the same conclusions. However, on the other hand, I would recommend that best practice would be to focus interpretation only on a set of parameters that would be estimated in a single nonlinear model (e.g., h or β_N, not both) to avoid over-extracting the results. Selecting which set of interpretable parameters to extract should be guided by the conclusions that researchers wish to draw, much like model selection proceeds generally.

4.2.2. Standard Errors

For these point estimates to be useful inferentially, we need a LENI approach to deriving the standard errors for the interpretable parameters. For an unknown quantity equal to a function of two quantities (θ_1 and θ_2) with known values and uncertainty (i.e., fixed effects and standard errors), the unknown uncertainty, $\mathrm{Var}(f(\theta_1, \theta_2))$, can be approximated by the quadratic expression of partial

https://doi.org/10.1017/psy.2024.2 Published online by Cambridge University Press

derivatives:[1]

$$\mathrm{Var}(f(\theta_1,\theta_2)) \approx \left(\frac{\partial f(\theta_1,\theta_2)}{\partial\theta_1}\right)^2 \mathrm{Var}(\theta_1) + 2\left(\frac{\partial f(\theta_1,\theta_2)}{\partial\theta_1}\right)\left(\frac{\partial f(\theta_1,\theta_2)}{\partial\theta_2}\right)\mathrm{Cov}(\theta_1,\theta_2) + \left(\frac{\partial f(\theta_1,\theta_2)}{\partial\theta_2}\right)^2 \mathrm{Var}(\theta_2) \tag{28}$$

Thus for x_N, we can take Equation 8 and express the expected variance of the parameter as

$$\mathrm{Var}(x_N) \approx \left(\frac{-1}{3\beta_3}\right)^2 \mathrm{Var}(\beta_2) + 2\left(\frac{-1}{3\beta_3}\right)\left(\frac{\beta_2}{3\beta_3^2}\right)\mathrm{Cov}(\beta_2,\beta_3) + \left(\frac{\beta_2}{3\beta_3^2}\right)^2 \mathrm{Var}(\beta_3) \tag{29}$$

Taking the square root of $\mathrm{Var}(x_N)$ yields the standard error, which can be used to compute a p-value or confidence interval as desired.

However, many of the nonlinear parameters are significantly more complex functions of the linear parameters. As such, this scalar equation approach quickly becomes tedious and error-prone. Alternatively, we can take a matrix approach and pre- and post-multiply the asymptotic covariance matrix of the linear parameter model – $\mathrm{ACOV}(\boldsymbol{\beta})$ – by the Jacobian of partial derivatives with respect to each linear parameter for a given nonlinear parameter transformation expression (e.g., $\mathbf{J}_{x_N}$).[2]

$$\mathrm{Var}(x_N) \approx \mathbf{J}'_{x_N}\ \mathrm{ACOV}(\boldsymbol{\beta})\ \mathbf{J}_{x_N} \tag{30}$$

If we expand the Jacobian with additional columns of partial derivatives corresponding to each nonlinear transformation, we can obtain the entire asymptotic covariance matrix of Equation 20 rather than only the variance of each parameter individually.

$$\mathrm{ACOV}(f(x_N, y_N, \delta, h)) \approx \mathbf{J}'_{f(x_N,y_N,\delta,h)}\ \mathrm{ACOV}(\boldsymbol{\beta})\ \mathbf{J}_{f(x_N,y_N,\delta,h)} \tag{31}$$

The square root of the diagonal of the resulting matrix is the vector of standard errors for the interpretable nonlinear parameters.

While this analytic approach works very well at approximating the ACOV matrix that would have resulted from directly fitting the nonlinear model (the error of approximation is very small; see Supplemental Code for complete details), we could alternatively use a bootstrapping approach to build an empirical standard error and confidence interval. This empirical alternative uses straightforward applications of the bootstrap and thus I will arrogate the details of such a procedure to the Supplemental Code for interested readers. We will see that the bootstrapping approach is more useful for

[1] Note that here I will use $\mathrm{VAR}(\theta)$ and $\mathrm{COV}(\theta)$ notation to denote the asymptotic variance-covariance terms. In contrast, I will use τ_θ or ψ_θ notation for the random effect variance-covariance terms in future sections.

[2] Note that the partial derivatives with respect to parameters that do not appear in the transformation expression (e.g., β_0 and β_1 for x_N) are simply 0.

https://doi.org/10.1017/psy.2024.2 Published online by Cambridge University Press

other components of the LENI approach in future sections.

4.2.3. Simulation Example 1: Fixed Effects

Before moving on to the random effects derivations, I first demonstrate the performance of these derivations in a simulated example of both the quadratic and cubic models. Here data were simulated from the nonlinear equations Equation 2 and Equation 20. The parameters of the quadratic are $\alpha_0 = 1$, $\alpha_x = 2$, and $\alpha_y = 8$, resulting in a concave quadratic function with a vertex at $x, y = (2, 8)$. The parameters of the cubic are $x_N = 0$, $y_N = 10$, $\delta = 3$, and $h = -2$, resulting in a cubic function where the local maximum occurs before the inflection point, the local minimum occurs after, and the vertex is at $(x, y) = (0, 10)$. For both models, I simulated data for 250 individuals 1000 times. I then fit the linear parameter model (Linear Estimates), generated LENI Estimates through the transformations outlined in the prior sections, and then fit the nonlinear parameter model directly (Nonlinear Estimates) with the true parameter values as starting values to avoid estimation issues. The mean parameter values across all iterations for each approach can be compared in Table 3. Notably, the LENI and Nonlinear point estimates and standard errors are nearly identical, with correspondence out to the 8th or 9th decimal place (see Supplemental for full precision details).

4.3. LENI for Random Effects

While the fixed effects are often the focus for empirical studies, we can also develop a set of transformations for the random effects. Random effects allow for individual variability in parameters of interest, and form the basis for more complex models, including conditional models (Biesanz et al., 2004; Curran et al., 2004; Raudenbush & Bryk, 2002) and models with distal outcomes (McCormick et al., 2024). Here the target is to transform the covariance matrix for the random effects obtained in the linear parameter version of the model into the covariance matrix we would have obtained from fitting the nonlinear parameter model directly. I will also show how to obtain standard errors associated with these (co)variance estimates.

To assess the generality of the LENI approach, I consider a maximal random effects model for both the quadratic and cubic model - that is all random effect variances (e.g., τ_{a_x}, τ_δ) and covariances (e.g., τ_{α_0,α_y}, $\tau_{y_N,h}$). Substantive applications might restrict the full covariance matrix (**T** in MLM, or in SEM growth models), usually for reasons of under-identification, either theoretically due to too few individual repeated measures or empirically due to a non-positive definite full matrix. Here the LENI approach takes advantage of the clear hierarchical structure of the linear parameter model, and we can restrict the random effects variances in sequence from highest-order (τ_{β_3}) to lowest (τ_{β_1}) as indicated.

https://doi.org/10.1017/psy.2024.2 Published online by Cambridge University Press

Table 3

LENI Approach to Fixed Effects Estimation

Pop. θ			Linear Estimates		LENI Estimates		Nonlinear Estimates
Quadratic Model							
α_0	1	β_0	1.020 (0.468)	α_0	1.020 (0.468)	α_0	1.020 (0.468)
α_x	2	β_1	6.987 (0.476)	α_x	1.997 (0.049)	α_x	1.997 (0.049)
α_y	8	β_2	−1.749 (0.111)	α_y	8.002 (0.446)	α_y	8.002 (0.446)
R^2	0.5		0.503				
Marg. ℓ			−738.80				−738.80
k			4				4
BIC			1499.68				1499.68
Cubic Model							
x_N	0	β_0	10.00 (0.139)	x_N	−0.003 (0.115)	x_N	−0.003 (0.115)
y_N	10	β_1	−1.004 (0.080)	y_N	10.00 (0.094)	y_N	10.00 (0.094)
δ	3	β_2	2.64×10^{-4} (0.013)	δ	3.008 (0.105)	δ	3.008 (0.105)
h	−2	β_3	0.037 (0.005)	h	−2.013 (0.130)	h	−2.013 (0.130)
R^2	0.5		0.506				
Marg. ℓ			−447.04				−447.04
k			5				5
BIC			921.69				921.69

Note: Parameter estimates and standard errors (in parentheses) are the mean values across 1000 iterations of data generation and model fitting. Pop θ indicates the generating value for each parameter. Linear Estimates indicate the fitted values from the linear parameter model. LENI Estimates indicate the transformed estimates of the nonlinear parameter model based on the Linear Estimates. Nonlinear Estimates indicate the fitted values from directly estimating the nonlinear parameter model. R^2 is the proportion of variance explained (only available in the linear parameter model), Marg. ℓ is the marginal log-likelihood, k is the number of model parameters, and BIC is the Bayesian Information Criterion.

4.3.1. Variance Estimates

In a beautiful bit of symmetry, we can apply the same approach for obtaining the standard errors that outlined above to obtain the variance-covariance matrix of the random effects for the nonlinear parameter model ($\mathbf{T}_{f(x_N,y_N,\delta,h)}$). That is, we can use the Jacobian of partial derivatives of each transformation function with respect to the linear parameters. However, instead of the ACOV($\boldsymbol{\beta}$) matrix as in Equation 31, here we will use $\mathbf{T}_{\boldsymbol{\beta}}$ – the variance-covariance matrix for the linear parameter model – instead.

$$\mathbf{T}_{f(x_N,y_N,\delta,h)} \approx \mathbf{J}'_{f(x_N,y_N,\delta,h)}\ \mathbf{T}_{\boldsymbol{\beta}}\ \mathbf{J}_{f(x_N,y_N,\delta,h)} \tag{32}$$

Unlike the standard error approach, the resulting $\mathbf{T}_{f(x_N,y_N,\delta,h)}$ matrix can be used directly, and

https://doi.org/10.1017/psy.2024.2 Published online by Cambridge University Press

the off-diagonal covariances (e.g., τ_{x_N,y_N}) or standardized correlations are often of direct theoretical interest.

4.3.2. Variance Standard Errors

While conceptually similar to deriving the standard errors of the fixed effects, deriving analytic expressions for the standard errors of the variance components in $\mathbf{T}_{f(x_N,y_N,\delta,h)}$ require us to consider additional transformations. Namely that the point estimates expressions of the variance components that we must compute the Jacobian partial derivatives on are now the quadratic expressions given by Equation 32.

While the full set of transformations is outlined in the Supplemental Material, prior work has shown that analytic expressions are unlikely to provide optimal estimates for these standard errors in practice because of the asymmetric nature of variance estimates, a point proven by the simulation results. As is the case with variance-covariance parameters in standard linear parameter models (Bolker, 2016), it is preferable to generate bootstrap confidence interval estimates instead and apply the point estimate transformations from Equation 32 to obtain confidence intervals on the nonlinear parameter model estimates.

4.3.3. Simulation Example 2: Random Effects

To test the LENI approach to random effects, I simulated data for 1000 replicated samples from the nonlinear equations with saturated random effects structures (quadratic: $N = 500$, $t = 4$; cubic: $N = 750$, $t = 6$). Like before, I present results from the linear parameter model (Linear Estimates), transformed results (LENI Estimates), and from a nonlinear parameter model directly (Nonlinear Estimates). In contrast to the tight correspondence between LENI and Nonlinear point estimates (Table 3), the estimates for the variance (τ) and correlation (ρ) parameters differ to a greater extent from one another, and from the population-generating parameters. While the results for the quadratic model appear reasonable, the estimates for the cubic model are much less accurate for either LENI or Nonlinear approaches, with wide standard errors, reflecting the general difficulty of fitting such high-dimensional random effects models (Table 4). Follow-up investigation of the results suggests that this is due to poor recovery of the Linear Parameter covariance matrix elements, where the standardized bias in the quadratic ($M_{abs.val.} = 0.535$, $range_{abs.val.} = 0.166 - 1.164$) and cubic ($M_{abs.val.} = 0.681$, $range_{abs.val.} = 0.049 - 1.839$) models were high. The substantially better fit of the Nonlinear model, as indicated by the lower average BIC, is likely driven by two factors: 1) the population generating model is from the nonlinear equation while the LENI estimates are approximations, and 2) population values were input into the Nonlinear model as starting values because of convergence issues, meaning that

https://doi.org/10.1017/psy.2024.2 Published online by Cambridge University Press

these are best-case estimates of the Nonlinear model. Despite these discrepancies in the random effects results, the LENI estimates for the fixed effects in these models still perform well (see Supplemental for full details). This suggests that if the fixed effects are of key theoretic interest, the LENI results can still perform well, while if the covariances/correlations are key to testing the substantive theory, alternative methods should be utilized. One would be to directly fit the Nonlinear model (although recovery is likely to still be poor in the cubic model), or alternatively to utilize the structural equation model approach outlined in a later section.

4.4. LENI for Including Predictors of Interest

4.4.1. Conditional Effects

The point estimates, variance-covariance terms, and associated standard errors of the interpretable parameters offer important information about the developmental process under consideration, however, by themselves, they are largely descriptive of the pattern of change over time. Many important developmental questions involve testing predictors of growth parameters (Bauer & Curran, 2005; Curran et al., 2004), and we can derive a LENI approach to this as we did with the fixed effects of the unconditional growth model.

https://doi.org/10.1017/psy.2024.2 Published online by Cambridge University Press

Table 4

LENI Approach to Random Effects Estimation

Pop. θ		Linear Estimates		LENI Estimates		Nonlinear Estimates	
Quadratic Model							
τ_{α_0}	1	τ_{β_0}	1.044 (0.307)	τ_{α_0}	1.044 (0.307)	τ_{α_0}	0.934 (0.331)
τ_{α_x}	0.1	τ_{β_1}	1.557 (0.577)	τ_{α_x}	0.157 (0.057)	τ_{α_x}	0.092 (0.023)
τ_{α_y}	0.25	τ_{β_2}	0.232 (0.077)	τ_{α_y}	0.378 (0.175)	τ_{α_y}	0.246 (0.121)
ρ_{α_0,α_x}	0	ρ_{β_0,β_1}	−0.730 (0.124)	ρ_{α_0,α_x}	0.003 (0.198)	ρ_{α_0,α_x}	−0.092 (0.264)
ρ_{α_0,α_y}	0.3	ρ_{β_0,β_2}	0.497 (0.163)	ρ_{α_0,α_y}	0.252 (0.252)	ρ_{α_0,α_y}	0.320 (0.241)
ρ_{α_x,α_y}	−0.2	ρ_{β_1,β_2}	−0.914 (0.044)	ρ_{α_x,α_y}	−0.356 (0.248)	ρ_{α_x,α_y}	−0.355 (0.313)
R^2	0.5		0.506				
Marg. ℓ			−4469.58				−4440.13
k			10				10
BIC			9015.03				8942.41
Cubic Model							
τ_{x_N}	0.2	τ_{β_0}	3.696 (1.342)	τ_{x_N}	0.253 (0.057)	τ_{x_N}	0.195 (0.053)
τ_{y_N}	0.25	τ_{β_1}	0.455 (0.166)	τ_{y_N}	1.047 (0.553)	τ_{y_N}	0.702 (0.310)
τ_{δ}	0.2	τ_{β_2}	0.016 (0.003)	τ_{δ}	0.269 (0.045)	τ_{δ}	0.202 (0.027)
τ_h	0.5	τ_{β_3}	0.001 (2.41×10^{-4})	τ_h	2.265 (0.884)	τ_h	1.259 (0.569)
ρ_{x_N,y_N}	0.4	ρ_{β_0,β_1}	0.076 (0.244)	ρ_{x_N,y_N}	0.231 (0.273)	ρ_{x_N,y_N}	0.099 (0.327)
$\rho_{x_N,\delta}$	0	ρ_{β_0,β_2}	−0.855 (0.079)	$\rho_{x_N,\delta}$	−0.004 (0.126)	$\rho_{x_N,\delta}$	−0.006 (0.118)
$\rho_{x_N,h}$	−0.2	ρ_{β_0,β_3}	−0.049 (0.180)	$\rho_{x_N,h}$	−0.049 (0.218)	$\rho_{x_N,h}$	−0.119 (0.331)
$\rho_{y_N,\delta}$	−0.1	ρ_{β_1,β_2}	−0.041 (0.200)	$\rho_{y_N,\delta}$	−0.042 (0.207)	$\rho_{y_N,\delta}$	0.046 (0.273)
$\rho_{y_N,h}$	0.35	ρ_{β_1,β_3}	−0.819 (0.064)	$\rho_{y_N,h}$	0.044 (0.318)	$\rho_{y_N,h}$	0.223 (0.278)
$\rho_{\delta,h}$	0.15	ρ_{β_2,β_3}	0.020 (0.144)	$\rho_{\delta,h}$	0.243 (0.190)	$\rho_{\delta,h}$	0.128 (0.254)
R^2	0.5		0.514				
Marg. ℓ			−16 199.67				−16 131.59
k			15				15
BIC			32 498.64				32 362.48

Note: Parameter estimates and standard errors (in parentheses) are the mean values across 1000 iterations of data generation and model fitting. Pop θ indicates the generating value for each parameter. Linear Estimates indicate the fitted values from the linear parameter model. LENI Estimates indicate the transformed estimates of the nonlinear parameter model based on the Linear Estimates. Nonlinear Estimates indicate the fitted values from directly estimating the nonlinear parameter model. τ parameters represent variances, ρ parameters represent correlations, R^2 is the proportion of variance explained (only available in the linear parameter model), Marg. ℓ is the marginal log-likelihood, k is the number of model parameters, and BIC is the Bayesian Information Criterion.

Here we need to distinguish between time-varying and time-invariant covariates (Curran & Bauer, 2011; McNeish & Matta, 2019). The relevant difference here is that time-varying covariates predict the repeated measures outcome directly, while time-invariant covariates predict the repeated measures indirectly through the growth parameters. The central insight here is that this indirect prediction

https://doi.org/10.1017/psy.2024.2 Published online by Cambridge University Press

results in interaction terms in the growth equation. As such, we can use basic principles of interactions in regression analysis (Aiken & West, 1991) to derive point estimates and standard errors for conditional effects. Consider a binary time-invariant covariate w_i which is included in the cubic equation, predicting all growth parameters. For person i at time t, the resulting growth model would have the form

$$\begin{aligned} y_{ti} = \beta_0 + \beta_1 x_{ti} + \beta_2 x_{ti}^2 + \beta_3 x_{ti}^3 + \\ \beta_4 w_i + \beta_5 w_i x_{ti} + \beta_6 w_i x_{ti}^2 + \beta_7 w_i x_{ti}^3 \end{aligned} \tag{33}$$

We can therefore use what we know about the expected value of this equation at different levels of w_i to define the conditional effect on the nonlinear parameter. For instance, for $w_i = 0$, the expected value of x_N is

$$x_{N|(w=0)} = \frac{-\beta_2}{3\beta_3} \tag{34}$$

while for $w_i = 1$ then the expected value is

$$x_{N|(w=1)} = \frac{-(\beta_2 + \beta_6)}{3(\beta_3 + \beta_7)} \tag{35}$$

because β_6 is the expected change in β_2 when w_i is shifted one unit, and β_7 is the expected change in β_3 across the same change in w_i.

Taking the difference of Equation 35 and Equation 34 gives the change in x_N per unit change in w_i (denoted here as $\pi_{x_N,w}$).

$$\pi_{x_N,w} = \frac{\Delta x_N}{\Delta w = 1} = \frac{-(\beta_2 + \beta_6)}{3(\beta_3 + \beta_7)} - \frac{-\beta_2}{3\beta_3} \tag{36}$$

The conditional effects of each of the interpretable nonlinear parameters can be derived in this fashion, and the standard errors can be obtained using the same Jacobian or bootstrapping approaches outlined in Standard Errors.

4.4.2. Simulation Example 3: Conditional Effects

I simulated a conditional version of each nonlinear equation, where w_i was used to predict each growth term. The covariate was simulated w_i as both a binary ($w_i \sim Bernoulli(0.4)$) and a continuous ($w_i \sim N(0, 0.25^2)$) covariate in separate simulations for generality, but the LENI computations of the relevant conditional effects apply across predictor types. Below, the results of the binary simulations are reported, but the continuous results can be seen in the Supplemental. Like in the unconditional model, the LENI approach does an excellent job of approximating the Nonlinear results in the fixed effects for both the quadratic and cubic models (Table 5).

https://doi.org/10.1017/psy.2024.2 Published online by Cambridge University Press

Table 5

LENI Approach to Conditional Effects Estimation

Pop. θ		Linear Estimates		LENI Estimates		Nonlinear Estimates	
Quadratic Model							
α_0	5	β_0	5.008 (0.133)	α_0	5.008 (0.133)	α_0	4.988 (0.146)
α_x	2	β_1	-4.014 (0.189)	α_x	2.002 (0.043)	α_x	2.000 (0.042)
α_y	1	β_2	1.004 (0.059)	α_y	0.991 (0.105)	α_y	0.993 (0.100)
$\pi_{\alpha_0,w}$	-0.5	β_3	-0.506 (0.211)	$\pi_{\alpha_0,w}$	-0.506 (0.211)	$\pi_{\alpha_0,w}$	-0.505 (0.232)
$\pi_{\alpha_x,w}$	0.25	β_4	0.907 (0.299)	$\pi_{\alpha_x,w}$	0.258 (0.109)	$\pi_{\alpha_x,w}$	0.260 (0.107)
$\pi_{\alpha_y,w}$	0	β_5	-0.313 (0.094)	$\pi_{\alpha_y,w}$	0.005 (0.164)	$\pi_{\alpha_y,w}$	0.006 (0.154)
R^2	0.5		0.467				
Marg. ℓ			-2007.62				-1990.89
k			8				8
BIC			4070.51				4025.94
Cubic Model							
x_N	0	β_0	10.01 (0.279)	x_N	-0.001 (0.047)	x_N	-0.006 (0.047)
y_N	10	β_1	3.751 (0.104)	y_N	10.00 (0.192)	y_N	10.13 (0.189)
δ	4	β_2	-1.96×10^{-4} (0.011)	δ	4.000 (0.031)	δ	3.998 (0.031)
h	10	β_3	-0.078 (0.003)	h	10.01 (0.262)	h	10.00 (0.262)
$\pi_{x_N,w}$	0	β_4	0.237 (0.441)	$\pi_{x_N,w}$	0.001 (0.073)	$\pi_{x_N,w}$	0.001 (0.073)
$\pi_{y_N,w}$	0.25	β_5	-0.019 (0.164)	$\pi_{y_N,w}$	0.245 (0.305)	$\pi_{y_N,w}$	0.243 (0.301)
$\pi_{\delta,w}$	-0.1	β_6	4.05×10^{-4} (0.017)	$\pi_{\delta,w}$	-0.098 (0.047)	$\pi_{\delta,w}$	-0.098 (0.047)
$\pi_{h,w}$	-0.3	β_7	-0.004 (0.004)	$\pi_{h,w}$	-0.294 (0.413)	$\pi_{h,w}$	-0.298 (0.414)
R^2	0.5		0.504				
Marg. ℓ			-8089.88				-8059.37
k			10				10
BIC			16 239.68				16 178.66

Note: Parameter estimates and standard errors (in parentheses) are the mean values across 1000 iterations of data generation and model fitting. Pop θ indicates the generating value for each parameter. Linear Estimates indicate the fitted values from the linear parameter model. LENI estimates indicate the transformed estimates of the nonlinear parameter model based on the Linear Estimates. Nonlinear Estimates indicate the fitted values from directly estimating the nonlinear parameter model. π parameters represent conditional effects of w_i on the nonlinar parameters, R^2 is the proportion of variance explained (only available in the linear parameter model), Marg. ℓ is the marginal log-likelihood, k is the number of model parameters, and BIC is the Bayesian Information Criterion.

4.5. LENI Real Data Example

As an empirical demonstration of the LENI approach, I drew network modularity data from the BrainTime sample (McCormick et al., 2021) and used self-reported sex (female = 0; male = 1) as a

https://doi.org/10.1017/psy.2024.2 Published online by Cambridge University Press

predictor of the curve components. I then fit the following conditional random-intercept model

$$\begin{aligned} modularity_{ti} =& \gamma_{00} + \gamma_{10}\ age_{ti} + \gamma_{20}\ age_{ti}^2 + \gamma_{01}\ male_i + \\ & \gamma_{11}\ male_i\ age_{ti} + \gamma_{21}\ male_i\ age_{ti}^2 + u_{0i} + r_{ti} \end{aligned} \tag{37}$$

with both main effects and interactions with the predictor. I also considered models hierarchically – a benefit of the linear estimation component of LENI – with linear and quadratic random effects, but these models were singular. Using the defined transformations, I then examined the implied nonlinear results and interpreted sex-specific trajectories of brain network organization.

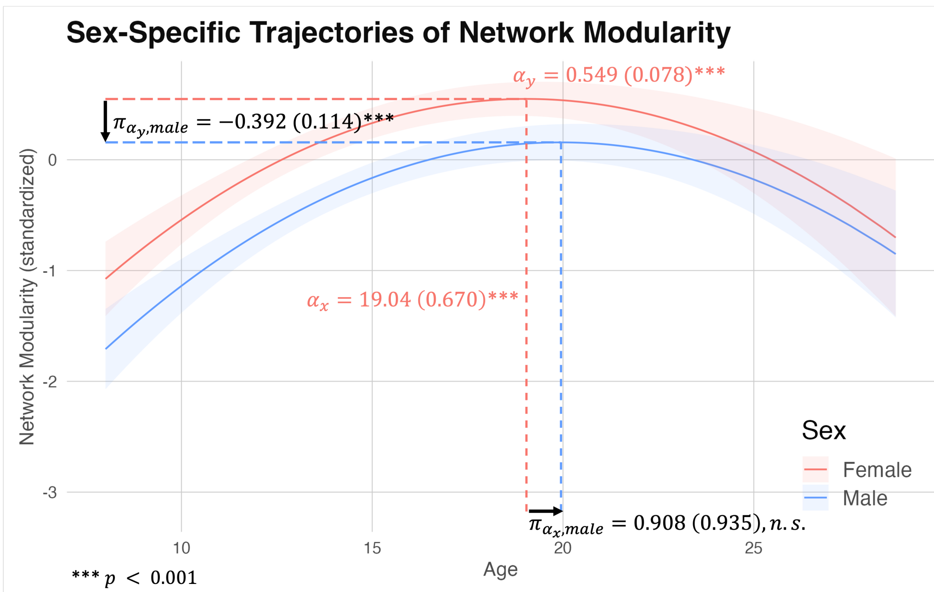

Figure 4. Sex-Specific Trajectories of Brain Network Organization. While the implied trajectory for male adolescents (blue) appears to have a delayed peak compared with female adolescents (red, α_x), the inference test on the nonlinear parameter ($\pi_{\alpha_x,male}$) derived from the LENI approach shows that they are statistically indistinguishable.

Plotting the data, we can see that the trajectory for network modularity in female adolescents (red) peaks at 19.04 years-of-age ($SE = 0.670$) with male adolescents (blue) appearing to reach this vertex somewhat later (implied at 19.95 years) and at lower peak levels (Figure 4). While common practice has been to simply assert this difference as meaningful (i.e., "boys show delayed development compared with girls"), the LENI approach allows us to build a direct statistic test for sex-specific differences in meaningful parameters of these trajectories. The LENI results suggest that while female adolescents do show a higher peak value in modularity at the vertex (α_y) than male adolescents ($\pi_{\alpha_y,male}$ = -0.392, $SE = 0.114$, t = -3.429; $p < 0.001$), there is no significant sex difference in the age at which the vertex is reached ($\pi_{\alpha_x,male} = 0.908$, $SE = 0.935$, $t = 0.971$, $p = 0.332$). This demonstrates the importance of parameterizing statistical models to test meaningful hypotheses with appropriate uncertainty (here the

https://doi.org/10.1017/psy.2024.2 Published online by Cambridge University Press

standard errors) rather than relying solely on point estimates to draw inferences about developmental theories (e.g., sex-specific delays in maturation).

4.6. Linearized Structural Equation Models

In the prior sections, I laid out a series of transformations that allow us to transform the results from a linearly-estimated model to approximate the results of directly estimating the nonlinear model – with attendant advantages related to estimation, measures of R^2, and hierarchical specification of the random effects structure. I have highlighted these transformations thus far primarily using mixed-effects (multilevel) models on time-unstructured data for maximum generality, although these methods could easily be applied in more time-structured cohort data without issue. Indeed, the set of LENI transformations could be equally useful in structural equation growth models like the latent curve (Meredith & Tisak, 1990) and latent change score models (McArdle et al., 2009).

However, adopting the SEM framework allows us to extend the LENI conceptual framework in a more interesting way by estimating a *linearized version* of the nonlinear model directly within the latent variable software (Blozis, 2004; Browne, 1993; Preacher & Hancock, 2012, 2015).[3] That is, rather than estimating the familiar linear parameter model and applying post-hoc transformations to obtain the nonlinear inferences, we can specify a linearized SEM to allow for direct estimation of the nonlinear parameters within a linear estimator. Prior work has used this approach to model a wide array of potential nonlinear functions within a linear SEM framework, including logistic curves (Choi et al., 2009), multiphase (piecewise) models with random knots (Feng et al., 2019; Preacher & Hancock, 2015), half-life with negative exponentials (Blozis, 2004; Preacher & Hancock, 2015), and time-to-criterion models (Johnson & Hancock, 2019), among others.

The general procedure for fitting linearized SEM has been detailed in full by prior work (Blozis, 2004, 2007; Feng et al., 2019; Preacher & Hancock, 2012, 2015), so I briefly review the relevant modeling steps here and then move into the specific use-cases of the models discussed thus far. First (1), we need to define the nonlinear equation – either a reparameterization such as the alternative quadratic or cubic, or a natively nonlinear function (e.g., logistic, negative exponential) – with a set of parameters that correspond to theoretically meaningful quantities. Then (2), to make the nonlinear function compatible with the linearly-estimated SEM, we linearize the function through a first-order Taylor series approximation by taking the partial derivative of the nonlinear function with respect to each parameter at the mean value of all other parameters. Once we have these partial derivatives (3), we use a structured latent curve model (SLCM; Blozis, 2004; Browne & Du Toit, 1991; Browne, 1993) approach where we set the factor loadings of each latent variable – which now represent the meaningful

[3]It should be noted that this form of linearization is also possible within a Bayesian mixed-effects framework due to the full flexibility of that set of modeling approaches.

https://doi.org/10.1017/psy.2024.2 Published online by Cambridge University Press

nonlinear parameters – as the partial derivative with respect to that same parameter, and specify each observed repeated measure intercept as the mean of the target nonlinear function at that value of time. Finally (4), with this model specification, we can estimate the linearized model with standard SEM software and obtain direct estimates and standard errors for the meaningful nonlinear parameters.

The resulting SLCM (Browne, 1993) models both average and individual-level change, although with some notable differences in interpretation compared with both the standard latent curve model (Meredith & Tisak, 1990) as well as fully nonlinear mixed-effects growth models. In the standard latent curve model, individual trajectories must follow the same functional form as the average trajectory, whereas the SLCM relaxes this constraint to allow for increased flexibility to fit unique patterns of change that do not follow the average (Blozis & Harring, 2017). However, like the standard latent curve model, SCLMs show dynamic consistency – that is the model for the population mean response is equal to the average of the individual-level effects – in contrast with fully non-linear models where this equivalence is not imposed (for an in-depth treatment of these issues and how they impact model interpretation, see Blozis & Harring, 2016; Harring & Blozis, 2016).

While differing in this way from fully nonlinear mixed-effects models, adopting the SLCM framework allows us to take advantage of the full flexibility of the SEM to model additional complexities to the core linearized function, including covariates (Curran et al., 2004; Preacher & Hancock, 2015), distal outcomes (McCormick et al., 2023, 2024), and approaches for parameter moderation (Bauer, 2017). We can next turn to the alternative polynomials as examples of this process.

4.6.1. Alternative Polynomial Models

Here I will focus on the alternative cubic, as a linearized version of the alternative quadratic model (Cudeck & du Toit, 2002) has been demonstrated previously (Preacher & Hancock, 2015), however, the code for both models can be found in the Supplemental Material. We can use either Equation 20 or Equation 21 as the target nonlinear equation from, so I will define the partial derivatives for all 5 nonlinear parameters (note only 4 are modeled at any given time). The partial derivatives for Equation 20 are

$$\begin{aligned}
\frac{\partial f(x_N, y_N, \delta, h, x_{ti})}{\partial x_N} &= -\frac{3h\left(\delta^2 - (x_{ti} - x_N)^2\right)}{2\delta^3} \\
\frac{\partial f(x_N, y_N, \delta, h, x_{ti})}{\partial y_N} &= 1 \\
\frac{\partial f(x_N, y_N, \delta, h, x_{ti})}{\partial \delta} &= -\frac{3h\left(x_N - x_{ti}\right)\left((x_{ti} - x_N)^2 - \delta^2\right)}{2\delta^4} \\
\frac{\partial f(x_N, y_N, \delta, h, x_{ti})}{\partial h} &= \frac{\left(x_N - x_{ti}\right)\left((x_{ti} - x_N)^2 - 3\delta^2\right)}{2\delta^3}
\end{aligned} \tag{38}$$

https://doi.org/10.1017/psy.2024.2 Published online by Cambridge University Press

and for Equation 21 are

$$
\begin{aligned}
\frac{\partial f(x_N, y_N, \delta, \beta_N, x_{ti})}{\partial x_N} &= -\frac{\beta_N \left(\delta^2 - (x_{ti} - x_N)^2\right)}{\delta^3} \\
\frac{\partial f(x_N, y_N, \delta, \beta_N, x_{ti})}{\partial y_N} &= 1 \\
\frac{\partial f(x_N, y_N, \delta, \beta_N, x_{ti})}{\partial \delta} &= -\frac{2\beta_N (x_N - x_{ti})^3}{3\delta^3} \\
\frac{\partial f(x_N, y_N, \delta, \beta_N, x_{ti})}{\partial \beta_N} &= \frac{(x_N - x_{ti}) \left((x_{ti} - x_N)^2 - 3\delta^2\right)}{3\delta^2}
\end{aligned}
\tag{39}
$$

Once we have these partial derivatives, we can set each factor loading to the relevant expression at each value of x_{ti}[4] and then define the intercepts of the repeated measures as the mean of the target nonlinear function at that value of x_{ti}. The corresponding path diagram for both the alternative quadratic (A) and cubic (B) are presented in Figure 5. Note that this specification allows us to directly model the means, variances, and covariances of the interpretable parameters. Unlikely standard latent curve model specifications, the estimated parameters of the latent variables are used to define the numerical value of the factor loadings (remember that the partial derivatives are evaluated at the means of the other parameters).

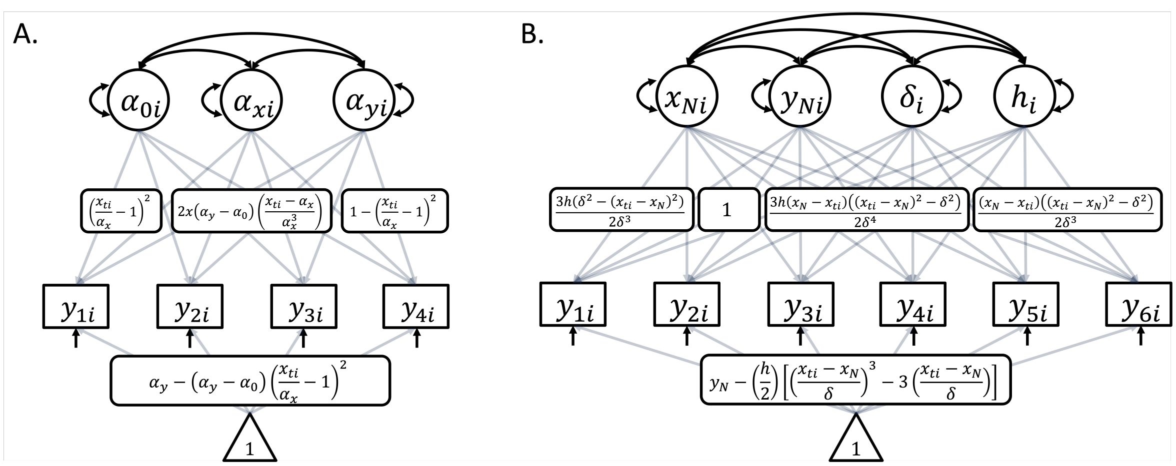

Figure 5. Linearized SEM. We can directly model the nonlinear parameters (e.g., α_x and δ) as random latent variables through a process of linearization where we set the factor loadings to partial derivatives of the target nonlinear function with respect to each modeled parameter and set the intercepts of the repeated measures to the mean of the target function.

To avoid unnecessary repetition with prior sections, I will leave the full parameter comparison to the Supplemental Material, but note that all the linearized models successfully capture the interpretable

[4]While I do not consider this extension here, the presence of the i subscript here allows for the possibility of modeling definition variables (Mehta & West, 2000) for time-unstructured data.

https://doi.org/10.1017/psy.2024.2 Published online by Cambridge University Press

model parameters and have near-identical fit to the standard linear versions of the polynomial LCMs ($\Delta\chi^2_{quadratic} = -7.713 \times 10^{-5}$; $\Delta\chi^2_{cubic,\,h} = 8.414 \times 10^{-7}$; $\Delta\chi^2_{cubic,\,\beta_N} = -1.126 \times 10^{-6}$).

4.6.2. Multiphase Polynomial Models

Finally, I applied the linearization approach to the multiphase (or piecewise) cubic model outlined previously. Feng et al. (2019) outlined a related approach to deal with floor and ceiling effects in modeled variables using a 3-phase linear model. Their approach reparameterized the standard multiphase linear model into a single equation by taking the median of the three linear functions – see Figure 1 and Equations 1-3 in Feng et al. (2019) for complete details. Unfortunately, while this approach successfully models the multiphase trajectory for monotonic functions, like the 3-phase linear model, for nonmonotonic functions like the quadratic and cubic, it does not appropriately define the onset and offset of the phases (see the Supplemental Material for examples of the challenges).

To accomplish the same idea, I parameterized a multiphase function where rather than taking the median of the *functions* of each phase, I instead took the median of three quantities – $(x_N - \delta)$, x_{ti}, and $(x_N + \delta)$ – which *does* change monotonically as a function of x_{ti} (Figure 6 B). Thus when x_{ti} is $< (x_N - \delta)$ or $> (x_N + \delta)$, its value is effectively fixed at those boundary points. This allows cubic change to only occur within the onset-offset boundaries (Figure 6 C). As highlighted in Feng et al. (2019), the median of three monotonic functions can be computed as

$$y = median(g_1, g_2, g_3) = sum(g_1, g_2, g_3) - min(g_1, g_2) - max(g_2, g_3) \tag{40}$$

where the minimum and maximum of two quantities are

$$\begin{aligned} min(a,b) &= \frac{1}{2}\left(a + b - \sqrt{(a-b)^2}\right) \\ max(a,b) &= \frac{1}{2}\left(a + b + \sqrt{(a-b)^2}\right) \end{aligned} \tag{41}$$

By substituting Equation 23 or Equation 24 into these expressions, I derived the following forms for the multiphase cubic (see Supplemental Material for full derivation details)

https://doi.org/10.1017/psy.2024.2 Published online by Cambridge University Press

$$
\begin{aligned}
y_{ti} &= y_N - \left(\frac{h}{2}\right)\left[\left(\frac{med\,(x_N-\delta, x_{ti}, x_N+\delta)-x_N}{\delta}\right)^3 - 3\left(\frac{med\,(x_N-\delta, x_{ti}, x_N+\delta)-x_N}{\delta}\right)\right] \\
&= y_N - \left(\frac{h}{2}\right)\left[\left(\frac{\frac{1}{2}\left(\sqrt{(x_N-\delta-x_{ti})^2}-\sqrt{(x_{ti}-x_N-\delta)^2}\right)}{\delta}\right)^3 - 3\left(\frac{\frac{1}{2}\left(\sqrt{(x_N-\delta-x_{ti})^2}-\sqrt{(x_{ti}-x_N-\delta)^2}\right)}{\delta}\right)\right] \\
&= y_N - \left(\frac{\beta_N\delta}{3}\right)\left[\left(\frac{med\,(x_N-\delta, x_{ti}, x_N+\delta)-x_N}{\delta}\right)^3 - 3\left(\frac{med\,(x_N-\delta, x_{ti}, x_N+\delta)-x_N}{\delta}\right)\right] \\
&= y_N - \left(\frac{\beta_N\delta}{3}\right)\left[\left(\frac{\frac{1}{2}\left(\sqrt{(x_N-\delta-x_{ti})^2}-\sqrt{(x_{ti}-x_N-\delta)^2}\right)}{\delta}\right)^3 - 3\left(\frac{\frac{1}{2}\left(\sqrt{(x_N-\delta-x_{ti})^2}-\sqrt{(x_{ti}-x_N-\delta)^2}\right)}{\delta}\right)\right]
\end{aligned}
\tag{42}
$$

I took these equations and linearized them using the same procedure as above (see Figure 6 A for a representative path diagram). I then generated cortical thickness trajectory data based on Fuhrmann et al. (2022), with 100 cases observed across 12 waves. The linearized SEM ($\chi^2_{model} = 84.299$, $df = 75$, $p = 0.217$) captured the parameters of the growth trajectory, with an average onset of change at 12 ($x_N - \delta = 12.028$, $SE = 0.201$), an inflection point at 15 ($x_N = 15.049$, $SE = 0.105$) with a negative instantaneous slope (β_N = -0.337, $SE = 0.022$), and an offset of change at 18 years-of-age ($x_N + \delta$ = 18.069, $SE = 0.164$; see Figure 6 D for model-implied individual and group trajectories; see the Supplemental Material for full code and output). This example highlights the extreme flexibility of the linearization approach for linear estimation with nonlinear inference (LENI) – any target nonlinear function can be transformed using the approach outlined above and accommodated within standard linear estimation software.

https://doi.org/10.1017/psy.2024.2 Published online by Cambridge University Press

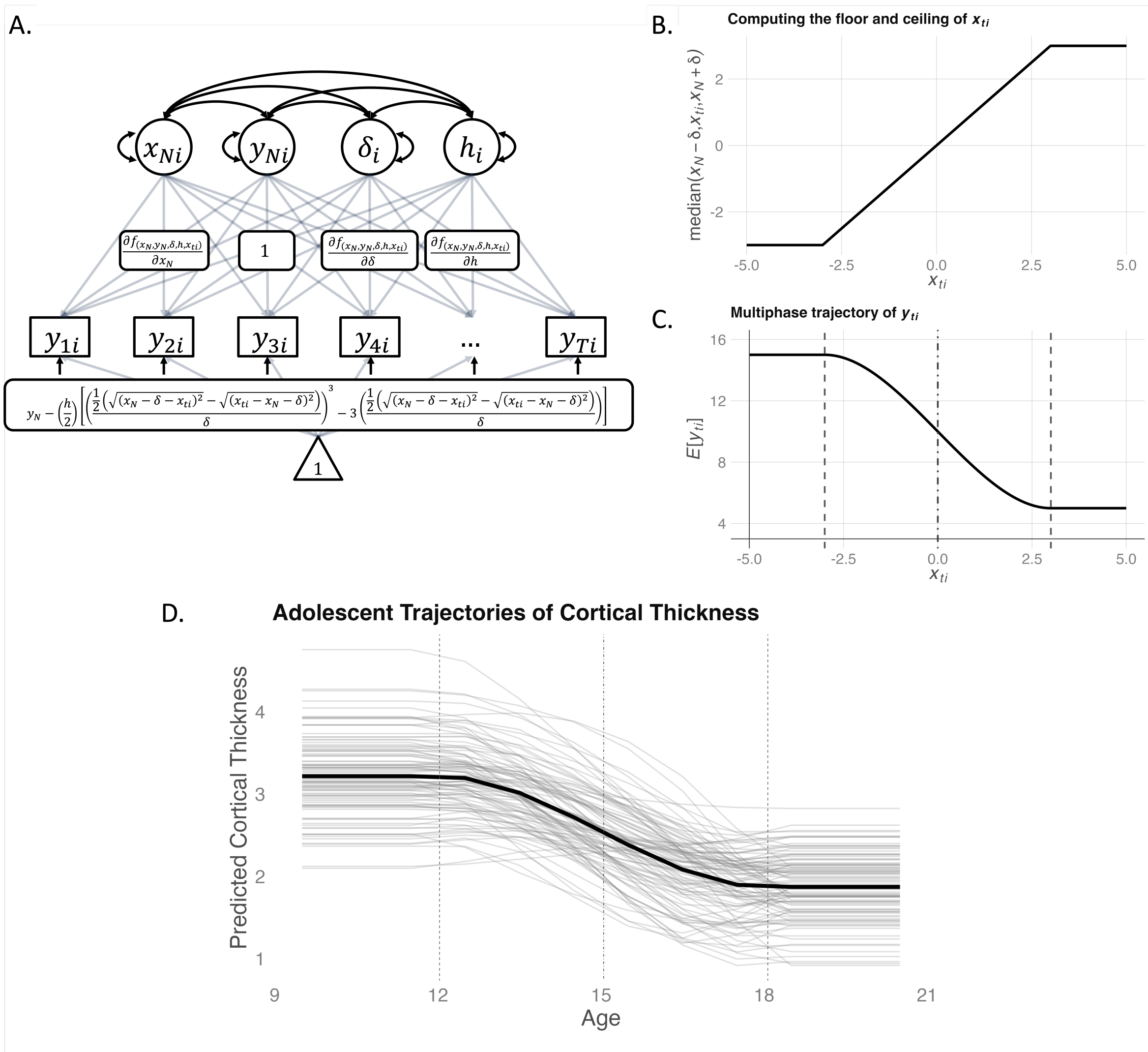

Figure 6. Multiphase Linearized SEM. A) A canonical path diagram of the linearized SEM model for the multiphase cubic. B) Because the cubic function is non-monotonic, we need to instead take the median of $(x_N - \delta)$, x_{ti}, and $(x_N + \delta)$, which is monotonic as a function of x_{ti}. C) The boundaries formed for the predictor x_{ti} in (B) allow for the proper specification of the multiphase cubic function for y_{ti}. D) The multiphase cubic model was fit to generated trajectories of cortical thinning (Fuhrmann et al., 2022) which successfully recovered individual (grey) and group (black) trajectories.

5 Recommendations for Applied Researchers

For applied research, the LENI suite of approaches offers exciting new opportunities for testing novel theoretical hypotheses with an ultimate eye toward the development of generative models for change over time. Hypotheses must be generated and tested on meaningful and interpretable developmental quantities (Preacher & Hancock, 2015), and not narrowly restricted to linear parameter models that are the default in major software implementations. The potential of LENI is for researchers to be able to fit the model they *want* to test theoretically, while avoiding some of the issues inherent in nonlinear

https://doi.org/10.1017/psy.2024.2 Published online by Cambridge University Press

model estimation. Here I demonstrated both mixed-effects and structural equation model options for random-effects growth modeling, allowing researchers maximal flexibility in specifying the nonlinear model within broader modeling traditions that exist in different fields. For instance, both the mixed-effect and SEM LENI approaches allow for the additional inclusion of time-varying covariates through direct prediction of the within-person repeated measure outcome, as well as additional time-invariant predictors and distal outcomes at the between-person (factor or random effect) level.[5]

Selecting between the various LENI approaches outlined may seem daunting, but ultimately the choice is likely far simpler than it appears. While there are many differences between mixed-effect and structural equation modeling approaches to longitudinal data analysis (McCormick et al., 2023; McNeish & Matta, 2018) – some important in the way that they model the repeated measures and others simply due to conventions or idiosyncratic discipline preferences – there are no additional issues related to the LENI set of approaches which would preference either class of modeling. For instance, time-unstructured data is more easily accommodated in mixed-effect approaches while multivariate models are more naturally fit within structural equation models, but this applies equally to linear and nonlinear/linearized versions of the two frameworks. As such, researchers should select the particular approach that is already suitable for testing their theoretical question and for accommodating the structure of the data that they have. LENI is here to help in that endeavor - not to impose additional restrictions.

To aid in the application of LENI approaches in substantive research, I have developed an *R* package (*leni*; [**Blinded**]) which allows users to convert the output of linear regression and mixed-effects models into nonlinear estimates using the transformations highlighted here, as well as to generate *lavaan* (Rosseel, 2012) syntax for linearized SEMs.

6 Conclusions

The linear estimation with nonlinear inference (LENI) approach is a broad framework that allows for the modeling of nonlinear parameters which represent theoretically interesting quantities while taking advantages of the well-behaved properties of linear parameter models for estimation. My goal was to offer a comprehensive introduction to 1) the motivation and approach for defining nonlinear models with interpretable parameters, 2) defining a set of transformation functions to convert linear mixed-effects models into nonlinear output, and 3) direct estimation of nonlinear parameters through a linearized SEM approach. This foundation of the LENI approach for modeling growth offers fertile ground for additional research and methodological development, with additional avenues for work

[5]Although distal outcomes present some additional complexity with mixed-effects models because they require a two-step procedure in a frequentist modeling framework.

https://doi.org/10.1017/psy.2024.2 Published online by Cambridge University Press

on small-sample behavior, methods for increasing the reliability of random effect/factor (co)variances, optimal Bayesian approaches (e.g. transformation robustness when applied at the individual draw level), and the role of time coding for estimation and interpretation of the nonlinear parameters. Combining the theoretical perspective of the LENI approach – focusing on meaningful features of change over time – with its computational efficiencies in approximating complex nonlinear equations shows great promise for advancing developmental science and the analysis of longitudinal data broadly.

https://doi.org/10.1017/psy.2024.2 Published online by Cambridge University Press

Acknowledgments

[**Blinded**] was supported with funds from the NWO (Nederlandse Organisatie voor Wetenschappelijk Onderzoek), Domain Social Sciences and Humanities (SSH) Sector Plan: Resilience in Youth, and Jacobs Foundation Fellowship ([**Blinded**]). Special thanks go to [**Blinded**] for critical feedback on the manuscript.

Competing Interests

The author declares no competing interests.

https://doi.org/10.1017/psy.2024.2 Published online by Cambridge University Press

References

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Sage Publications, Inc.

Bauer, D. J. (2017). A more general model for testing measurement invariance and differential item functioning. *Psychological Methods*, *22*(3), 507–526.

Bauer, D. J., & Curran, P. J. (2005). Probing interactions in fixed and multilevel regression: Inferential and graphical techniques. *Multivariate Behavioral Research*, *40*(3), 373–400. https://doi.org/10.1207/s15327906mbr4003_5

Biesanz, J. C., Deeb-Sossa, N., Papadakis, A. A., Bollen, K. A., & Curran, P. J. (2004). The role of coding time in estimating and interpreting growth curve models. *Psychological Methods*, *9*(1), 30–52. https://doi.org/10.1037/1082-989X.9.1.30

Blozis, S. A. (2004). Structured latent curve models for the study of change in multivariate repeated measures. *Psychological Methods*, *9*(3), 334–353. https://doi.org/10.1037/1082-989X.9.3.334

Blozis, S. A. (2007). On fitting nonlinear latent curve models to multiple variables measured longitudinally. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*(2), 179–201. https://doi.org/10.1080/10705510709336743

Blozis, S. A., & Harring, J. R. (2016). On the Estimation of Nonlinear Mixed-Effects Models and Latent Curve Models for Longitudinal Data. *Structural Equation Modeling: A Multidisciplinary Journal*, *23*(6), 904–920. https://doi.org/10.1080/10705511.2016.1190932

Blozis, S. A., & Harring, J. R. (2017). Understanding Individual-level Change Through the Basis Functions of a Latent Curve Model. *Sociological Methods & Research*, *46*(4), 793–820. https://doi.org/10.1177/0049124115605341

Bolker, B. (2016). Wald errors of variance. https://rpubs.com/bbolker/waldvar

Braams, B. R., van Duijvenvoorde, A. C. K., Peper, J. S., & Crone, E. A. (2015). Longitudinal changes in adolescent risk-taking: A comprehensive study of neural responses to rewards, pubertal development, and risk-taking behavior. *Journal of Neuroscience*, *35*(18), 7226–7238. https://doi.org/10.1523/JNEUROSCI.4764-14.2015

Browne, M. W., & Du Toit, S. H. C. (1991). Models for learning data. American Psychological Association. https://doi.org/10.1037/10099-004

Browne, M. W. (1993, January 1). Structured latent curve models [DOI: 10.1016/B978-0-444-81531-6.50016-7]. In C. M. Cuadras & C. R. Rao (Eds.). North-Holland. https://doi.org/10.1016/B978-0-444-81531-6.50016-7

Chaiken, S. R. (1994). The inspection time not studied: Processing speed ability unrelated to psychometric intelligence. *Intelligence*, *19*(3), 295–313. https://doi.org/10.1016/0160-2896(94)90003-5

Choi, J., Harring, J. R., & Hancock, G. R. (2009). Latent growth modeling for logistic response functions. *Multivariate Behavioral Research*, *44*(5), 620–645. https://doi.org/10.1080/00273170903187657

Cudeck, R., & du Toit, S. H. C. (2002). A version of quadratic regression with interpretable parameters. *Multivariate Behavioral Research*, *37*(4), 501–519. https://doi.org/10.1207/s15327906mbr3704_04

Cudeck, R., & Klebe, K. J. (2002). Multiphase mixed-effects models for repeated measures data. *Psychological Methods*, *7*(1), 41–63. https://doi.org/10.1037/1082-989X.7.1.41

Curran, P. J., & Bauer, D. J. (2011). The disaggregation of within-person and between-person effects in longitudinal models of change. *Annual Review of Psychology*, *62*(1), 583–619. https://doi.org/10.1146/annurev.psych.093008.100356

Curran, P. J., Bauer, D. J., & Willoughby, M. T. (2004). Testing main effects and interactions in latent curve analysis. *Psychological Methods*, *9*(2), 220–237. https://doi.org/10.1037/1082-989X.9.2.220

Derefinko, K. J., Charnigo, R. J., Peters, J. R., Adams, Z. W., Milich, R., & Lynam, D. R. (2016). Substance use trajectories from early adolescence through the transition to college. *Journal of Studies on Alcohol and Drugs*, *77*(6), 924–935. https://doi.org/10.15288/jsad.2016.77.924

Eggleston, E., Laub, J., & Sampson, R. (2004). Methodological sensitivities to latent class analysis of long-term criminal trajectories. *Journal of Quantitative Criminology*, *20*, 1–26. https://doi.org/10.1023/B:JOQC.0000016696.02763.ce

Farrell, S., & Lewandowsky, S. (2018, February 22). *Computational modeling of cognition and behavior*. Cambridge University Press.

Feng, Y., Hancock, G. R., & Harring, J. R. (2019). Latent growth models with floors, ceilings, and random knots. *Multivariate Behavioral Research*, *54*(5), 751–770. https://doi.org/10.1080/00273171.2019.1580556

Flora, D. B. (2008). Specifying piecewise latent trajectory models for longitudinal data. *Structural Equation Modeling: A Multidisciplinary Journal*, *15*(3), 513–533. https://doi.org/10.1080/10705510802154349

Fox, J., & Weisberg, S. (2011). Nonlinear regression and nonlinear least squares in R. SAGE.

Fuhrmann, D., Madsen, K. S., Johansen, L. B., Baaré, W. F. C., & Kievit, R. A. (2022). The midpoint of cortical thinning between late childhood and early adulthood differs between individuals and brain regions: Evidence from longitudinal modelling in a 12-wave neuroimaging sample. *NeuroImage*, *261*, 119507. https://doi.org/10.1016/j.neuroimage.2022.119507

Giedd, J. N., Raznahan, A., Alexander-Bloch, A., Schmitt, E., Gogtay, N., & Rapoport, J. L. (2015). Child psychiatry branch of the national institute of mental health longitudinal structural magnetic resonance imaging study of human brain development. *Neuropsychopharmacology*, *40*(1), 43–49. https://doi.org/10.1038/npp.2014.236

Grimm, K. J., Zhang, Z., Hamagami, F., & Mazzocco, M. (2013). Modeling Nonlinear Change via Latent Change and Latent Acceleration Frameworks: Examining Velocity and Acceleration of Growth Trajectories. *Multivariate Behavioral Research*, *48*(1), 117–143. https://doi.org/10.1080/00273171.2012.755111

Harring, J., & Blozis, S. (2016). A Note on Recurring Misconceptions When Fitting Nonlinear Mixed Models. *Multivariate Behavioral Research*, *51*. https://doi.org/10.1080/00273171.2016.1239522

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology*, *27*(2), 236–248. https://doi.org/10.1037/0012-1649.27.2.236

Johnson, T. L., & Hancock, G. R. (2019). Time to criterion latent growth models. *Psychological Methods*, *24*(6), 690–707. https://doi.org/10.1037/met0000214

Karriker-Jaffe, K. J., Foshee, V. A., Ennett, S. T., & Suchindran, C. (2008). The development of aggression during adolescence: Sex differences in trajectories of physical and social aggression among youth in rural areas. *Journal of abnormal child psychology*, *36*(8), 1227–1236. https://doi.org/10.1007/s10802-008-9245-5

Lenroot, R. K., Gogtay, N., Greenstein, D. K., Wells, E. M., Wallace, G. L., Clasen, L. S., Blumenthal, J. D., Lerch, J., Zijdenbos, A. P., Evans, A. C., Thompson, P. M., & Giedd, J. N. (2007). Sexual dimorphism of brain developmental trajectories during childhood and adolescence. *NeuroImage*, *36*(4), 1065–1073. https://doi.org/10.1016/j.neuroimage.2007.03.053

LeWinn, K. Z., Sheridan, M. A., Keyes, K. M., Hamilton, A., & McLaughlin, K. A. (2017). Sample composition alters associations between age and brain structure. *Nature Communications*, *8*(1), 874. https://doi.org/10.1038/s41467-017-00908-7

Mareschal, D., & Thomas, M. S. C. (2007). Computational modeling in developmental psychology. *IEEE Transactions on Evolutionary Computation*, *11*(2), 137–150. https://doi.org/10.1109/TEVC.2006.890232

McArdle, J. J., Grimm, K. J., Hamagami, F., Bowles, R. P., & Meredith, W. (2009). Modeling life-span growth curves of cognition using longitudinal data with multiple samples and changing scales of measurement. *Psychological Methods*, *14*(2), 126–149. https://doi.org/10.1037/a0015857

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

McCormick, E. M., Curran, P. J., & Hancock, G. R. (2024). Latent growth factors as predictors of distal outcomes. *Psychological Methods*. https://doi.org/10.1037/met0000663

McCormick, E. M., Peters, S., Crone, E. A., & Telzer, E. H. (2021). Longitudinal network re-organization across learning and development. *NeuroImage*, *229*, 117784. https://doi.org/10.1016/j.neuroimage.2021.117784

McNeish, D., & Bauer, D. J. (2022). Reducing incidence of nonpositive definite covariance matrices in mixed effect models. *Multivariate Behavioral Research*, *57*(2-3), 318–340. https://doi.org/10.1080/00273171.2020.1830019

McNeish, D., Bauer, D. J., Dumas, D., Clements, D. H., Cohen, J. R., Lin, W., Sarama, J., & Sheridan, M. A. (2021). Modeling individual differences in the timing of change onset and offset. *Psychological Methods*. https://doi.org/10.1037/met0000407

McNeish, D., & Dumas, D. (2017). Nonlinear Growth Models as Measurement Models: A Second-Order Growth Curve Model for Measuring Potential. *Multivariate Behavioral Research*, *52*(1), 61–85. https://doi.org/10.1080/00273171.2016.1253451

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

McNeish, D., & Matta, T. H. (2019). Flexible treatment of time-varying covariates with time unstructured data. *Structural Equation Modeling: A Multidisciplinary Journal*, *27*(2), 298–317. https://doi.org/10.1080/10705511.2019.1627213

Mehta, P. D., & West, S. G. (2000). Putting the individual back into individual growth curves. *Psychological Methods*, *5*(1), 23–43. https://doi.org/10.1037/1082-989X.5.1.23

Meredith, W., & Tisak, J. (1990). Latent curve analysis. *Psychometrika*, 107–122. https://doi.org/10.1007/BF02294746

Mills, K. L., Goddings, A.-L., Herting, M. M., Meuwese, R., Blakemore, S.-J., Crone, E. A., Dahl, R. E., Güroğlu, B., Raznahan, A., Sowell, E. R., & Tamnes, C. K. (2016). Structural brain development between childhood and adulthood: Convergence across four longitudinal samples. *NeuroImage*, *141*, 273–281. https://doi.org/10.1016/j.neuroimage.2016.07.044

Nickalls, R. (1993). A new approach to solving the cubic: Cardan's solution revealed. *The Mathematical Gazette*, *77*(480), 354–359. https://doi.org/10.2307/3619777

Nunes, A. S., Vakorin, V. A., Kozhemiako, N., Peatfield, N., Ribary, U., & Doesburg, S. M. (2020). Atypical age-related changes in cortical thickness in autism spectrum disorder. *Scientific Reports*, *10*(1), 11067. https://doi.org/10.1038/s41598-020-67507-3

Orben, A., Przybylski, A. K., Blakemore, S.-J., & Kievit, R. A. (2022). Windows of developmental sensitivity to social media. *Nature Communications*, *13*(1), 1649. https://doi.org/10.1038/s41467-022-29296-3

Patzelt, E. H., Hartley, C. A., & Gershman, S. J. (2018). Computational phenotyping: Using models to understand individual differences in personality, development, and mental illness. *Personality Neuroscience*, *1*, e18. https://doi.org/10.1017/pen.2018.14

Petersen, A. C., Crockett, L., Richards, M., & Boxer, A. (1988). A self-report measure of pubertal status: Reliability, validity, and initial norms. *Journal of Youth and Adolescence*, *17*(2), 117–133. https://doi.org/10.1007/BF01537962

Pfefferbaum, A., Kwon, D., Brumback, T., Thompson, W. K., Cummins, K., Tapert, S. F., Brown, S. A., Colrain, I. M., Baker, F. C., Prouty, D., De Bellis, M. D., Clark, D. B., Nagel, B. J., Chu, W., Park, S. H., Pohl, K. M., & Sullivan, E. V. (2018). Altered brain developmental trajectories in adolescents after initiating drinking. *American Journal of Psychiatry*, *175*(4), 370–380. https://doi.org/10.1176/appi.ajp.2017.17040469

Pleskac, T. J., Yu, S., Hopwood, C., & Liu, T. (2019). Mechanisms of deliberation during preferential choice: Perspectives from computational modeling and individual differences. *Decision*, *6*(1), 77–107. https://doi.org/10.1037/dec0000092

Preacher, K. J., & Hancock, G. R. (2012). On interpretable reparameterizations of linear and nonlinear latent growth curve models. IAP Information Age Publishing.

Preacher, K. J., & Hancock, G. R. (2015). Meaningful aspects of change as novel random coefficients: A general method for reparameterizing longitudinal models. *Psychological Methods*, *20*(1), 84–101. https://doi.org/10.1037/met0000028

Ram, N., & Grimm, K. (2007). Using simple and complex growth models to articulate developmental change: Matching theory to method. *International Journal of Behavioral Development*, *31*(4), 303–316. https://doi.org/10.1177/0165025407077751

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*. SAGE.

Rosseel, Y. (2012). Lavaan: An r package for structural equation modeling. *Journal of Statistical Software*, *48*, 1–36. https://doi.org/10.18637/jss.v048.i02

Shaw, P., Kabani, N. J., Lerch, J. P., Eckstrand, K., Lenroot, R., Gogtay, N., Greenstein, D., Clasen, L., Evans, A., Rapoport, J. L., Giedd, J. N., & Wise, S. P. (2008). Neurodevelopmental trajectories of the human cerebral cortex. *The Journal of Neuroscience*, *28*(14), 3586–3594. https://doi.org/10.1523/JNEUROSCI.5309-07.2008

Somerville, L. H., Jones, R. M., Ruberry, E. J., Dyke, J. P., Glover, G., & Casey, B. J. (2013). The medial prefrontal cortex and the emergence of self-conscious emotion in adolescence. *Psychological Science*, *24*(8), 1554–1562. https://doi.org/10.1177/0956797613475633

Tamnes, C. K., Herting, M. M., Goddings, A.-L., Meuwese, R., Blakemore, S.-J., Dahl, R. E., Güroğlu, B., Raznahan, A., Sowell, E. R., Crone, E. A., & Mills, K. L. (2017). Development of the cerebral cortex across adolescence: A multisample study of inter-related longitudinal changes in cortical volume, surface area, and thickness. *The Journal of Neuroscience*, *37*(12), 3402–3412. https://doi.org/10.1523/JNEUROSCI.3302-16.2017

Teachman, B. A. (2006). Aging and negative affect: The rise and fall and rise of anxiety and depression symptoms. *Psychology and Aging*, *21*(1), 201–207. https://doi.org/10.1037/0882-7974.21.1.201

Wilson, R. C., & Collins, A. G. (2019). Ten simple rules for the computational modeling of behavioral data (T. E. Behrens, Ed.). *eLife*, *8*, e49547. https://doi.org/10.7554/eLife.49547

Wood, S. N. (2000). Modelling and smoothing parameter estimation with multiple quadratic penalties. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *62*(2), 413–428. https://doi.org/10.1111/1467-9868.00240

Zhang, Z., McArdle, J. J., & Nesselroade, J. R. (2012). Growth rate models: Emphasizing growth rate analysis through growth curve modeling. *Journal of Applied Statistics*, *39*(6), 1241–1262. https://doi.org/10.1080/02664763.2011.644528