Multivariate Behavioral Research

ISSN: 0027-3171 (Print) 1532-7906 (Online) Journal homepage: http://www.tandfonline.com/loi/hmbr20

Nonlinear Growth Models as Measurement Models: A Second-Order Growth Curve Model for Measuring Potential

Daniel McNeish & Denis Dumas

MULTIVARIATE BEHAVIORAL RESEARCH
2016, VOL. 0, NO. 0, 1–25
http://dx.doi.org/10.1080/00273171.2016.1253451

Nonlinear Growth Models as Measurement Models: A Second-Order Growth Curve Model for Measuring Potential

Daniel McNeish[a] and Denis Dumas[b]

[a]Department of Human Development and Quantitative Methodology, University of Maryland; [b]Department of Human Development and Psychoeducational Studies, Howard University

ABSTRACT
Recent methodological work has highlighted the promise of nonlinear growth models for addressing substantive questions in the behavioral sciences. In this article, we outline a second-order nonlinear growth model in order to measure a critical notion in development and education: potential. Here, potential is conceptualized as having three components—ability, capacity, and availability—where ability is the amount of skill a student is estimated to have at a given timepoint, capacity is the maximum amount of ability a student is predicted to be able to develop asymptotically, and availability is the difference between capacity and ability at any particular timepoint. We argue that single timepoint measures are typically insufficient for discerning information about potential, and we therefore describe a general framework that incorporates a growth model into the measurement model to capture these three components. Then, we provide an illustrative example using the public-use Early Childhood Longitudinal Study–Kindergarten data set using a Michaelis-Menten growth function (reparameterized from its common application in biochemistry) to demonstrate our proposed model as applied to measuring potential within an educational context. The advantage of this approach compared to currently utilized methods is discussed as are future directions and limitations.

KEYWORDS
Non-linear growth; second-order model; dynamic measurement model

Since the beginning of formalized scientific investigations of the human mind, psychologists have been interested in examining not only their participants' current skills or abilities but also their potential for developing those abilities (Galton, 1869; Spearman, 1927; Thorndike, 1913). For example, cognitive and developmental psychologists have long been interested in ascertaining students' potential for developing mental abilities such as memory, creativity, or intelligence (Cattell, 1987; Torrance, 1962). Moreover, understanding student academic potential is a principal goal of psychological measurement (Jones & Thissen, 2007).

This abiding interest in understanding human abilities and potential has led to much methodological innovation, both for the reliable measurement of abilities and skills (Borsboom, 2006; Borsboom, Mellenbergh, & van Heerden, 2003; Reise & Waller, 2009; Sijtsma, 2012) and for modeling the growth of those abilities longitudinally (Bliese & Ployhart, 2002; Duncan & Duncan, 2004; Hancock, Kuo, & Lawrence, 2001; Lawrence & Hancock, 1998; Muthén & Khoo, 1998). However, abilities measured at a single point in time, no matter how reliably measured or sophisticatedly modeled, are not synonymous with potential (Harradine, Coleman, & Winn, 2014; Schick & Phillipson, 2009). Furthermore, there is a tendency to rely on so-called population-averaged or marginal methods that neglect to acknowledge the sizeable variability that exists between individuals (Day & Dragoni, 2015; L. T. Rose, Rouhani, & Fischer, 2013; T. Rose, 2016). Therefore, even given the existing body of rigorous methodological work, psychologists do not currently have access to a developed modeling framework that provides the necessary parameters for fully understanding potential, never mind the ability to accurately measure it.

Components of potential

In the following subsections, we will overview three separate components of potential. As a broad overview, Figure 1 presents these three components graphically. Capacity is a time-invariant upper asymptote that represents the maximum value a person can be expected to attain. Ability and availability are time-dependent quantities that assess progress toward reaching capacity. Ability is a measure of how much of the capacity has been realized at a certain timepoint while availability is a measure of how much of the capacity has yet to be realized. These three components are expanded upon next.

CONTACT Daniel McNeish dmcneish@email.unc.edu University of North Carolina, Chapel Hill, 100 E. Franklin Street Suite 200, Chapel Hill, NC 27599.

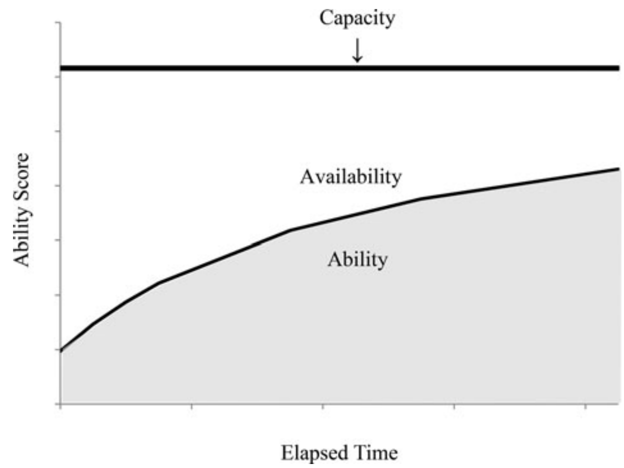

Figure 1. Theoretical depiction of components of potential. The space below the line is realized ability; the space above the line is unrealized availability; the horizontal line at the top is the capacity.

Ability

Ability refers to an individual's measured amount of skill or proficiency in a particular task at particular time (Mislevy, 1984; Spearman, 1927). Commonly employed single-timepoint psychometric models are capable of capturing a participant's current ability at a particular testing occasion and, as such, only offer limited information about that participant's potential for future growth in that ability (i.e., past performance is related to, but far from synonymous with, future performance). In Figure 1, these methods are akin to zooming in on ability at one particular point in time—an estimated ability at that timepoint may be accurate, but broader information about other aspects such as capacity and availability are lost.

To inform this important issue, building off of previous research (e.g., Baltes & Kliegl, 1992; Feuerstein, Rand, & Hoffman, 1979; Vygotsky, 1994), Sternberg et al. (2002) drew the distinction between *developed* abilities (which are measured statically and at a particular time) and *developing* abilities (which, like potential, largely exist in the future) after an initial test has been completed. From this understanding of traditional ability measures as static and representing developed abilities, it logically follows that, in order to derive some understanding of a participant's potential, multiple testing occasions, separated by opportunities for the participant to learn or develop their ability, are necessary (Babad & Budoff, 1974; Baltes & Kliegl, 1992; Budoff, Meskin, & Harrison, 1971). Therefore, existing psychological research that attempts to understand potential typically incorporates multiple measurement occasions (Budoff & Friedman, 1964; Fischer, 1980; Hamilton & Budoff, 1974; Plucker, 1999; Sternberg et al., 2002).

For example, dynamic assessment (DA; e.g., Tzuriel, 2001; Feuerstein et al., 1979) is one methodology that uses closely spaced measurement occasions intermingled with learning opportunities to tap a participant's potential for growth in an ability.

Interested readers are referred to Tiekstra, Minnaert, and Hessels (2016) for the most recent review of some aspects of DA methodology. Perhaps most germane to the focus of this article, Embretson (1987) developed a psychometric framework for measuring potential in the DA tradition, utilizing component latent trait models (Embretson, 1984) as well as structural models with predictive paths from DA pre- and posttests to irrelevant and transfer cognitive tasks. Although Embretson's approach was compelling, in order to be meaningfully applied to substantive questions, researchers were required to use tasks for which each cognitive component was known—something that is unfortunately rare in educational and psychological science (Grossnickle, Dumas, Alexander, & Baggetta, 2016). Moreover, it required at least two other tasks (i.e., irrelevant and transfer) to be given to participants on top of the already time-intensive DA process. For that reason, Embretson's (1987) psychometric approach to DA, like most DA methods, is suited for small samples, with whom the researcher spends considerable time, and not necessarily for large-scale data. Moreover, possibly because of the complexity of the approach, it has never to our knowledge been applied in its entirety to substantive questions concerning academic potential.

In contrast to DA methods, larger scale and long-term longitudinal studies have also been utilized to plot the growth of ability and infer individual differences in potential based on rate of growth (e.g., Plucker, 1999; Schaie, 2005). Such longitudinal methods may be desirable within the psychological and educational research fields, where large-scale longitudinal data sets concerning student learning are currently proliferating and becoming increasingly accessible to substantive researchers. In this way, ability measurements across time allow for some degree of inference to be made about how much ability an individual could develop given their current learning trajectory in a certain amount of future time. In our conceptualization of potential, these inferences are attempting to project information about ability on to another critical aspect of potential: capacity.

Capacity

Capacity, in our conceptualization, refers to the maximum amount of ability that an individual could develop given the individual's current and past ability levels (i.e., an upper asymptote on ability). It is interesting to note that capacity has been of explicit interest to psychologists since Binet and Simon's (1916) seminal early work in the field of intelligence. However, Binet and Simon's single-timepoint assessment methods did not allow them to tap capacity itself, but only time-dependent ability levels (Feuerstein et al., 1979; Sternberg et al., 2002). Later, Fischer (1980) posited his skill theory, which, like DA, emphasized measurement across multiple timepoints in order to ascertain a participant's optimal performance.

Still more recently, Baltes and colleagues (e.g., Baltes & Kliegl, 1992) developed their *testing-the-limits* method, which, similar to DA, used multiple testing occasions interspersed with task-specific training to ascertain how much of a given ability a participant was capable of developing over the course of the training. Baltes and Kliegl (1992), who were active in gerontology research, used this method to contrast what they called the *cognitive plasticity* of younger and older adults—namely, that older adults were less capable of improving their performance on a recall test than were younger adults, even with many training sessions. Perhaps most relevant to the current article, Baltes and Kleigl (1992) did not fit a longitudinal model to their data, but instead simply plotted group averages across time. When plotted, the time-specific averages for both groups unmistakably formed a nonlinear curve in which participants learned rapidly in the beginning, and then learned more slowly later on: the classic inverted-J "learning curve" first described by Ebbinghaus (1985) more than a century before. With these curves plotted, Baltes and Kliegl (1992) used participants' final ability level—collected until the curve began to plateau and resemble an asymptote—as their measure of capacity.

As previously mentioned, ability measures across time are necessary but not sufficient for ascertaining meaningful information about an individual's capacity. An effective method for accurately estimating a participant's capacity from the trajectory of ability growth is also required. In previous work on potential, ability growth over time is most typically considered to grow as a linear function of time due to the closeness of or scarcity of measurement occasions, meaning that the parameterization of conventional models is often inappropriate for determining the upper asymptote of one's ability because a linear growth model postulates continuous growth even as time approaches infinity (Cudeck & du Toit, 2002; Cudeck & Harring, 2007; Grimm, Ram, & Hamagami, 2011; Harring, Weiss, & Hsu, 2012; Preacher & Hancock, 2015; Ram & Grimm, 2007). Conceptually, in this way, an upper asymptote is necessary for quantifying capacity because it allows for a point at which growth in the measured ability terminates.

It is important to note that the DA framework used in the psychological literature to understand capacity often assumes a linear function over time to capture the relation between ability and capacity (Peterson & Gillam, 2015; Tzuriel, 2001). Although this may present a reasonable approximation for ability growth over small time scales, most abilities in which psychologists are interested are extremely unlikely to grow linearly indefinitely, and eventually approach an upper limit. Even if an individual never exactly reaches his or her capacity, the conceptualization of this upper asymptote of a developing ability function is critical for understanding potential, because without this estimate of maximum learning capacity, the calculation of as-yet-untapped potential—or available growth—is not possible.

Availability

We define *availability* as the difference between an individual's current ability and maximum capacity—*untapped potential* in lay terms. In this way, *availability* refers to unrealized potential growth. Therefore, *availability*, as defined here, bears theoretical resemblance to the Vygotsky's conceptual *zone of proximal development* (ZPD), drawn out over the time period of measurement (Vygotsky, 1994). It is important to note that DA has been used in the past to gain insight into the ZPD (Minik, 1987; 2005; Poehner & Lantolf, 2013; Sternberg et al., 2002), but such studies have been limited to a relatively small number of participants because of the resource-intensive nature of DA. It should be noted here that, in the present modeling framework, availability is a direct function of ability and capacity (i.e., availability = capacity – ability and summarizes the absolute distance between the two).

As a colloquial example, if an athletic coach says that a player has "potential," it may mean that the player has already developed skills and can presently contribute positively to the team (high ability), that the player has a high ceiling for developing relevant athletic skills (high capacity), or that the player could develop skills in the future but has yet to do so presently (high availability), or some combination of those three possibilities.

Modeling the components of potential

While technical details concerning the specific modeling framework adopted in this article are presented later, it is relevant here to discuss generally what characteristics a model would need to be able to produce estimates of these three components of potential: ability, capacity, and availability.

In a sense,measuring ability as a developed construct, at a single timepoint, has been the focus of the general field of psychometrics for the last century (Mislevy, 1984;

Spearman, 1927). Therefore, at this point, while incremental innovation is possible in the field, major shifts are not needed in order to be able to measure ability. However, as previously described, the developing components of potential—capacity and availability—cannot be measured with traditional single-timepoint psychometric models. Therefore, any model that would be able to tap these components of potential must necessarily be longitudinal in nature. Moreover, because capacity specifically refers to the upper limit of ability growth, a suitable model must be not only longitudinal, but also nonlinear with a functional form that approaches an upper asymptote.

In the most general sense, in order to account for differing participant capacities, such a model would need a specific mechanism by which participants could differ in their growth rate at different timepoints so that participants would also differ meaningfully in their upper asymptote of ability growth (capacity). Further, it would be useful for ability and capacity to be measured on the same scale as one another so as to be directly comparable. Indeed, this common scale is absolutely necessary if participant availability is to be calculated. Moreover, because capacity and ability have a temporal relation (i.e., participants' capacity always lies in the unmeasured future from their measured current ability), an ideal model would give some information about how far along the temporal path to capacity a given participant is at a given timepoint. For example, it would be useful to know what proportion of a given participant's estimated capacity had been reached at a given measured timepoint, or at what timepoints certain a priori designated proportions (e.g., half) of capacity are reached. This property of a model is especially important given that estimated capacities that lie very far in the future would necessarily be less precise (e.g., larger standard error estimates), so such a temporally scaled parameter would provide information needed by researchers and practitioners to make substantive inferences about students.

Finally, in much of the cognitive and behavior sciences, and especially in educational assessment practice, focus is placed on subject-specific information as opposed to marginal or average information. In the most basic sense, if a model for potential were to be able to be applied in the educational setting, it must provide student-specific information. Therefore, an ideal model for potential would be capable of including random effects on each of its parameters, allowing for all participants to have each of the parameters relevant to them "measured" and saved for future analysis. In this way, a model for potential would be capable of estimating subject-specific ability and availability at each timepoint and a unique subject-specific time-invariant capacity asymptote.

Review of nonlinear growth models with interpretable asymptotes

Although psychologists have largely conceded that a successful model for potential requires repeated measures, few previous attempts at modeling potential have considered incorporating the extensive literature on growth models, and nonlinear growth models in particular. Emerging research has shown that models can be parameterized conveniently so that an upper asymptote (reminiscent of capacity) is directly estimable and interpretable (Preacher & Hancock, 2015). If the outcome variable were an ability measured at several timepoints, this upper asymptote could be helpful for measuring capacity in a manner that other models for potential have been unable to achieve to date.

Several different types of models can be parameterized so that an upper asymptote is a parameter that is directly estimated from the model. These models generally model growth in two broad classes of methods: "S-shaped" trajectories or "inverted J-shaped" trajectories (Browne, 1993; Grimm & Ram, 2009; Grimm et al., 2011; Harring, Kohli, Silverman, & Speece, 2012; Ram & Grimm, 2007). Depending on the type of research question, the age of the participants, and the phenomena under investigation, different types of trajectories may be more desirable. Table 1 provides a nonexhaustive overview of some defining characteristics between commonly employed trajectories with each class. The following subsections will further differentiate between the different classes of trajectories as well as the particular functions within each class.

S-shaped growth trajectories

S-shaped (a.k.a., generalized logistic or sigmoid) growth trajectories may be helpful for scenarios where growth is slower near both extremes of the observation window and more rapid in the middle. A possible example might be measuring growth in second language acquisition where ability may grow slowly as individuals struggle initially to master the basic rules and vocabulary, then develop very quickly through conversational fluency as a sufficiently large vocabulary is developed, and finally grow slowly as they wrestle with idioms or other nonliteral expressions that come naturally to native speakers but are not intuitive to others (Ortega & Iberri-Shea, 2005). In addition to having an upper asymptote, S-shaped trajectories are often parameterized in terms of a lower asymptote, rate parameter (how quickly growth is occurring), an inflection point, and a rate-of-approach parameter. The inflection point in an S-shaped trajectory is the point at which the rate of growth no longer is increasing and the

Table 1. Summary of select nonlinear growth curves with interpretable upper asymptotes.

	Inverted J-shaped trajectories		
	Michaelis-Menten	Generalized Michaelis-Menten	von Bertalanffy
Mean trajectory	$\alpha_0 + \frac{(\alpha_U - \alpha_0)t}{\alpha_R + t}$	$\alpha_0{}^c + \frac{(\alpha_U{}^c - \alpha_0{}^c)t^c}{\alpha_R{}^c + t^c}$	$\alpha_A + (\alpha_0 - \alpha_U)\exp(-\alpha_R t)$
Inflection point	None	$\alpha_U[(c-1)/(c+1)]^{1/c}$	None
Inflection point symmetry	—	No	—
Benchmark parameter	α_R is the midpoint	α_R is the midpoint	None
Special cases contained within	None	Michaelis-Menten	Negative exponential
	S-Shaped trajectories		
	Gompertz†	Richards	Logistic
Mean trajectory	$\alpha_U \exp\{\ln(\frac{\alpha_0}{\alpha_U})\exp(-\alpha_R t)\}$	$\alpha_L + \frac{\alpha_U}{(1+\kappa e^{-\alpha_R(t-\varsigma)})^{1/\kappa}}$	$\alpha_L + \frac{\alpha_U}{1+e^{-\alpha_R(t-\varsigma)}}$
Inflection point	Fixed	ς	ς
Inflection point symmetry	No, 37% before inflection point	Only as a special case	Yes, 50% on either side of ς
Benchmark parameter	None	ς is the inflection point	ς is the inflection point
Special cases contained within	None	von Bertalanffy, Logistic, Gompertz	None

Note. t = time; α_0 = initial value parameter defined as the value of the outcome at $t = 0$; α_U = upper asymptote parameter defined as the outcome value when $t \to \infty$; α_R = rate parameter, which determines how quickly the curve approaches the asymptote; ς = an inflection point parameter, which locates when the rate of change begins to decrease rather than increase; κ = speed of approach parameter, which determines how quickly in the observation window the inflection point occurs; c = a dimension parameter that creates an inflection point (there is no inflection when $c = 1$); e is Euler's number, which is approximately 2.718.
†The Gompertz mean trajectory in the table is written using form used by Laird (1964), which uses the parameterization common to J-shaped trajectories (rate, asymptote, initial value). To better observe the relation between the Gompertz and Richards curves, the mean trajectory could alternatively be written as $\alpha_U e^{e^{-\alpha_R(t-\alpha_L)}}$ (Wood et al., 2015).

outcome variable continues to grow but at a decreasing rate. The rate-of-approach parameter determines how quickly in the observation window an inflection point occurs (i.e., it "pushes" the inflection point ahead or back in time). Very general curves such as the Schnute curve (Schnute, 1981) or the Richards curve (Richards, 1959; which is itself a special case of the Schnute curve) contain each of these parameters although not all types of S-shaped trajectories necessarily contain each parameter. Popular curves (e.g., logistic) can be obtained as special cases of the more general form by constraining certain parameters to specific values. For example, the popular logistic curve can be obtained by constraining the inflection point to occur such that 50% of growth occurs before the inflection point and 50% occurs after (Lopez et al., 2000). It should be noted here that, because the modeling of the components of potential, as defined here, is novel, S-shaped curves have not yet been utilized in the literature for that specific purpose. However, we present them here as one methodological option for researchers that may be useful given the hypothesized growth trajectories of an ability (i.e., abilities whose growth is theorized to have an inflection point or meaningful lower asymptote).

Inverted J-shaped growth trajectories

Inverted J-shaped trajectories also feature an interpretable upper asymptote but feature a different pattern of growth. Namely, the most rapid growth occurs early on and then tapers off at later timepoints, eventually reaching the upper asymptote. This type of growth may be more suitable, for example, when modeling younger participants' vocabulary growth over time (Harring et al., 2012). Unlike S-shaped curves, inverted J-shaped curves often do not have a lower asymptote, and the classical forms of these curves do not have inflection points (although they can be adapted to have such a feature; Lopez et al., 2000). Because the Michaelis-Menten model will be featured prominently in later sections, we provide more detail on it in the next subsection.

Michaelis-Menten

Commonly used in biochemistry, the Michaelis-Menten model[1] (Michaelis & Menten, 1913) also models inverted J-shaped growth. Although typically used to model reaction rates as a function of substrate concentration in enzyme kinetics, the Michaelis-Menten model is not commonly implemented in the behavioral sciences (although see Harring et al., 2012 for a conditionally linear variation fit to educational psychology data). The rate parameter in the Michaelis-Menten curve possesses a benchmark interpretation—the rate parameter is defined as the timepoint in the observation window at which the outcome variable is halfway between the initial value and the asymptote. As explained later in the example used in

[1] In biochemical applications, technically a Michaelis-Menten equation uses theoretical values whereas a Monod equation (Monod, 1949) takes the same form but uses empirical estimates to determine the parameters of the curve. Given that previous behavioral science applications have typically referred to this type of curve as a Michaelis-Menten curve despite the fact that the parameters are determined empirically (e.g., Codd & Cudeck, 2014; Cudeck & Harring, 2007; Cudeck & du Toit; 2002; Harring et al., 2012), we will follow this same convention. We are indebted to an anonymous reviewer for alerting us to the differentiation between the two types of equations in the original area of application.

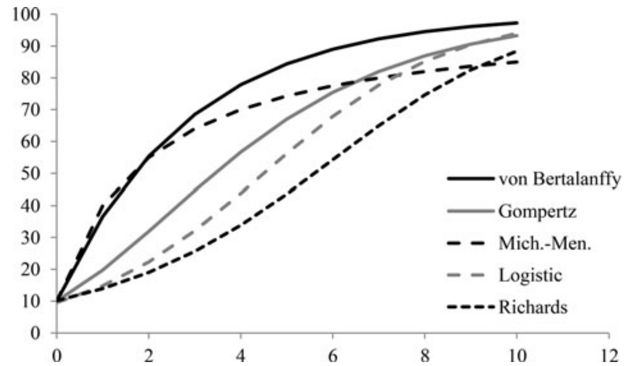

Figure 2. Demonstrative comparison of trajectories of Gompertz, Von Bertalanffy, Logistic, Richards, and Michaelis-Menten Curves. All curves have initial values of 10 and asymptotes of 100. The Gompertz, von Bertalanffy, Logistic, and Richards curves have rate parameters equal to 0.35 and the Michaelis-Menten model has a rate parameter equal to 2. The Richards curve has a speed of approach parameter of 1.5, which pushes the inflection point later in time compared to the Logistic curve (whose speed of approach parameter is fixed to 1).

this article, this parameter contains valuable substantive information that makes the Michaelis-Menten function potentially more suitable for educationally related research questions than similar inverted J-shaped curves such as the von Bertalanffy. Similar to the von Bertalanffy curve, the Michaelis-Menten curve may be most appropriate for research scenarios where the most rapid growth occurs at early ages, before tapering off at later timepoints (e.g., Harring et al., 2012).

Figure 2 compares the trajectories of five different curves featured in Table 1 (we do not differentiate between the Michaelis-Menten and Generalized Michaelis-Menten in Figure 2) for hypothetical data that begins at 10 and has a maximum at 100 with comparable rate parameters. From Figure 2, it can be seen that the S-shaped Gompertz, Logistic, and Richards curves grow slower initially but maintain steadier growth prior to approaching the asymptote. This behavior is caused by the presence of an inflection point (implicitly or explicitly) in these models. Both inverted J-shaped curves start similarly, but the von Bertalanffy continues to grow more steeply while the Michaelis-Menten function begins to plateau at an earlier timepoint. The behavior of these curves is prototypical of those lacking inflection points.

Overview of modeling strategy

Before fitting the nonlinear growth model, a measurement model is needed for the construct for which potential is of interest because abilities in psychological research can rarely be adequately measured with manifest variables (Ferrer, Balluerka, & Widaman, 2008; Geiser, Keller, & Lockhart, 2013). Then, this measurement model will be applied to individuals at several different timepoints, which necessitates that longitudinal invariance be upheld. Afterward, the best-fitting, most appropriate, and/or most interpretable nonlinear growth trajectory is selected and the growth in the ability factor scores is modeled with what is commonly referred to as a second-order growth model (Hancock et al., 2001; Meredith & Tisak, 1990) or curve-of-factors model (McArdle, 1988). Further, to address criticisms of previous models that focus on marginal relations rather than subject-specific relations (Molenaar & Campbell, 2009; L. T. Rose et al., 2013), each of the growth model coefficients has a random effect, allowing for a fully subject-specific model such that each individual has a unique initial value, upper asymptote, and growth rate. Estimates of these parameters allow each of our identified components of potential—ability, capacity, and availability—to be quantified for each individual. Ability and capacity are directly estimated from the model, and availability (a relative summary of capacity and ability) is calculated by subtracting the estimated ability from the subject-specific capacity estimate. It should be noted that placing a random effect on the rate parameter of the model results in some random effects having a multiplicative interpretation rather than an additive interpretation[2] (Grimm & Ram, 2009). This leaves two possible solutions—fit the model as a nonlinear mixed effects model (NLME; e.g., in SAS PROC NLMIXED) or as a structured latent curve model (SLCM; e.g., in M*plus*).

Although these two approaches are often considered to be more-or-less interchangeable, recent articles by Blozis and Harring (2016a; 2016b) comprehensively demonstrate the perceptible differences in the two model specifications. To briefly summarize, NLMEs require that every subject-specific curve follow the same functional form as the mean curve. Conversely, subject-specific curves in SLCMs do not need to have the same functional form as the mean curve so long as the sum of the individual curves equals the mean curve. This stems from the differing estimation methods used for each model—NLMEs use quadrature methods, which leave the nonlinear portions of the model intact and attempt to approximate an

[2] To explain the difference between additive and multiplicative random effects further, consider the von Bertalanffy model, $\alpha_U + (\alpha_0 - \alpha_U)\exp(-\alpha_R t)$. If random effects are placed on α_0 and α_U, then the random effects would be *additive* because the random effects for α_U and α_0 only interact via addition or subtraction. This results in the model being linear in the random effects, which can be estimated in either a structural equation model or a regression framework without issue (in the SEM framework, this would be referred to as a *conditionally linear model*; Harring et al., 2012). If a random effect is also placed on α_R, the random effects then are multiplied together, meaning that the random effects are *multiplicative* and that the model is nonlinear in the random effects. These types of models can only be estimated in their true form in a regression framework (Grimm & Ram, 2009; albeit with computational intensive methods). To estimate the model in an SEM, the model must first be linearized, which changes the interpretation of some of the parameters (Blozis & Harring, 2016a; 2016b). The same principle applies for the Gompertz and Michaelis-Menten models.

analytically intractable integral in the likelihood function. This prescribes that the subject-specific curves also follow the general form of the mean curve. On the other hand, SLCMs are required when fitting models in a structural equation model framework when the conditionally linear form of nonlinear growth models proposed by Meredith and Tisak (1990) is not upheld (Blozis & Harring, 2016a). In this case, SLCMs linearize the nonlinear portions of the model using first-order Taylor series expansions, which are taken with respect to the parameters in the mean curve such that the subject-specific curves are evaluated according to time (Blozis, 2004; Blozis & Harring, 2016a).

Because the inherent interest of the model is in the subject-specific curves (i.e., the components of potential for each individual), we will employ the NLME specification to ensure that the estimates adhere to a subject-specific interpretation rather than the marginalized interpretation inherent with the SLCM specification (although SLCM estimates will be used for starting values). As will be discussed shortly, retaining the subject-specific information will make the estimation quite complex.

Proposed model

First-order measurement model

First, a first-order measurement model is developed as would be carried out if one were measuring a traditionally developed construct (e.g., mathematics or reading ability) at multiple timepoints such that

$$\mathbf{y}_{it} = \boldsymbol{\tau}_t + \boldsymbol{\Lambda}_t \boldsymbol{\eta}_{it} + \boldsymbol{\varepsilon}_{it}, \tag{1}$$

where $\mathbf{y}_{it}$ is a $p \times 1$ vector of manifest indicator variables at time t ($t = 0, 1, 2, \ldots, T-1$) for the ith individual ($i = 1, 2, \ldots, N$); $\boldsymbol{\tau}_t$ is a $p \times 1$ vector of manifest variable intercepts at time t; $\boldsymbol{\Lambda}_t$ is a $p \times q$ matrix of factor loadings at time t; $\boldsymbol{\eta}_{it}$ is a $q \times 1$ vector of latent variable values (factor scores) for the ith person at time t; and $\boldsymbol{\varepsilon}_{it}$ is a $p \times 1$ vector of residuals for the ith individual at time t where $\boldsymbol{\eta} \sim MVN(\boldsymbol{\nu}, \boldsymbol{\psi})$ and $\boldsymbol{\varepsilon} \sim MVN(\mathbf{0}, \Theta)$ for N the total sample size, T the number of timepoints, p the number of manifest variables, and q the number of latent variables.

Note that the p observed means in the $\mathbf{y}$ vector are a function of p elements in $\boldsymbol{\tau}$ but also q latent means in $\boldsymbol{\eta}$. This means that the mean structure is overparameterized as $p + q$ elements would need to be estimated from p observed values. A common solution to this problem is to define the mean vector of $\boldsymbol{\eta}$ to be null (i.e., all factors have means constrained to 0). However, when featured in part of a second-order growth model, the first-order factors must have a nonzero mean structure so that growth can be effectively modeled (otherwise the mean at each timepoint will be zero and growth will be null). To circumvent the overparameterized mean structure, we constrain the mean of one of the indicator variables (contained in $\boldsymbol{\tau}$) to be 0 at each timepoint instead. This implies that the mean of the latent variables will then be on the scale of this selected indicator, which will allow a growth curve to be fit to the first-order factors (whose means now can be estimated).

Longitudinal invariance

Because the model involves repeated measures, q and p will necessarily be equal for each value of t, and the same manifest and latent variables must be present at each time t in order to achieve configural longitudinal invariance, which occurs when the time-specific measurement model is identical for each time t (Horn, McArdle, & Mason, 1983; Meredith, 1993). Both weak and strong factorial invariance must also be established such that the factor loadings and intercept means are invariant across time to ensure that the measurement of the latent variables is consistent across time (McArdle, 1988; Meredith & Horn, 2001); notationally, in the presence of strong longitudinal invariance, $\boldsymbol{\Lambda}$ and $\boldsymbol{\tau}$ will no longer carry a t subscript because one set of estimates will apply across all timepoints.

Second-order growth model

As noted earlier, the trajectory of the growth model applied to the first-order latent variables depends on many facets (e.g., outcome variable, participant's age, the domain of research). A Michaelis-Menten-type trajectory is used in the subsequent illustrative example, and we describe the second-order subject-specific extension of the Michaelis-Menten equation in Table 1. Notationally, the second-order growth portion of the model based upon a Michaelis-Menten trajectory can be written as

$$\eta_{it} = \beta_{0i} + \frac{(\beta_{Ui} - \beta_{0i})t}{\beta_{Ri} + t} + d_{it} \tag{2}$$

and

$$\beta_{0i} = \alpha_0 + \sum_{k_1=1}^{K_1} \gamma_{0k_1} X_{k_1} + \zeta_{0i}$$

$$\beta_{Ui} = \alpha_U + \sum_{k_2=1}^{K_2} \gamma_{Uk_2} X_{k_2} + \zeta_{Ui}$$

$$\beta_{Ri} = \alpha_R + \sum_{k_3=1}^{K_3} \gamma_{Rk_3} X_{k_3} + \zeta_{Ri} \tag{3}$$

where

$$\begin{aligned} \mathbf{d}_i &\sim MVN(\mathbf{0}, \mathbf{R}) \\ \boldsymbol{\zeta}_i &\sim MVN(\mathbf{0}, \mathbf{G}) \end{aligned} \tag{4}$$

where $\mathbf{R}$ is a $q \times q$ covariance matrix for the residuals; $\mathbf{G}$ is an $s \times s$ covariance matrix for the random effects where s is equal to the number of growth factor effects in the model; the ζ parameters are subject-specific random effects; the α parameters are the growth factor means; the γ parameters are coefficients for predictors of the growth factors that are constant for all individuals in the sample; and X are time-invariant predictors that need not be the same for β_{0i}, β_{Ui}, or β_{Ri}. Note that the outcome variable in Equation (2) is η_{it}, the value of the latent variables from the first-order measurement model in Equation (1), not an observed variable.

Recall from the previous section that $\boldsymbol{\eta} \sim MVN(\boldsymbol{\nu}, \boldsymbol{\psi})$. According to Equations (2) through (4), both $\mathbf{R}$ and $\mathbf{G}$ model $\boldsymbol{\psi}$, which means that constraints must be placed on $\mathbf{R}$ and/or $\mathbf{G}$ so that the covariance submodel is not overparameterized (i.e., because $\mathbf{G}$ is a covariance matrix of latent variables, $\mathbf{R}$ and $\mathbf{G}$ cannot both be unstructured because this would result in $[s(s+1)/2]$ more parameters than degrees of freedom available in the covariance submodel). It is often desirable to allow $\mathbf{G}$ to be unstructured, so constraints are often applied to $\mathbf{R}$ such that off-diagonal terms are equal to 0 (e.g., diagonal structures), off-diagonal terms are constrained to equality (e.g., compound symmetric structures), or subdiagonals are constrained to equality (e.g., Toeplitz or autoregressive structures).

Calculating the components of potential

A participant's capacity is simply equal to the subject-specific value of $\beta_{Ui} = \alpha_U + \zeta_{Ui}$; estimated ability at time t is equal to η_{it}; and availability is not directly estimated from the model but can be calculated simply from $\beta_{Ui} - \eta_{it} = \alpha_U + \zeta_{Ui} - \eta_{it}$, the difference between capacity (which is time invariant) and ability at time t. It should be noted that the Michealis-Menten function, being J-shaped, features an upper asymptote but not a lower asymptote. For the present conceptualization of the components of potential, only an upper asymptote is needed. β_{Ri} helps to illustrate how an individual's ability estimates progress over time. Specifically, ζ_{Ri} allows for an individual's growth trajectory to deviate from the average trajectory in ways that allow for the identification of participants who are fast or slow learners, early or late bloomers, and so forth. These distinctions are substantively relevant and will be briefly elucidated upon in the illustrative example provided in the next section using the public use 1999 Early Childhood Longitudinal Study–Kindergarten (ECLS-K; Tourranegeau et al., 2009).

Illustrative example using ECLS-K

Data and research question

ECLS-K 1999 tracked students in the 1998–1999 kindergarten cohort and collected data at seven timepoints: fall and spring of kindergarten, fall and spring of Grade 1, spring of Grade 3, spring of Grade 5, and spring of Grade 8. The publically available data set contains 21,054 students and several thousand variables including direct cognitive assessments, teacher reports, parent reports, and a host of questionnaires and demographic and background variables. For reasons described shortly, we restrict our analysis to students who have complete data. In this analysis, we used vertically scaled reading and math assessments at each timepoint of interest.

Means and standard deviations of each of the manifest variables utilized in this analysis are included in Table 2. In addition, bivariate correlations among all of these variables are included in Table 3. As can be seen, all bivariate correlations were positive and statistically significant, with the lowest correlations being approximately .20 and the strongest being approximately .80. All variables are continuous in nature.

First, we created an Academic Ability latent variable at each timepoint that loaded on these ECLS-K assessment data (vertically scaled mathematics and readings scores). We then established longitudinal invariance for the first-order model and fit a second-order growth model to the Academic Ability latent variables and reported the estimates. Fitting with the illustrative purpose of this model and to maintain clarity of exposition, we will fit an unconditional model without predictors. Example plots demonstrating the utility of the model output are also discussed.

Table 2. Descriptive statistics of manifest variables.

Testing occasion	Domain	Mean	*SD*
Kindergarten fall	Reading	36.70	10.32
	Math	28.16	9.45
Kindergarten spring	Reading	48.90	14.48
	Math	39.31	12.16
1st grade fall	Reading	51.37	14.95
	Math	45.99	14.00
1st grade spring	Reading	82.13	23.87
	Math	65.12	17.79
3rd grade spring	Reading	132.80	26.73
	Math	103.68	23.62
5th grade spring	Reading	156.00	24.81
	Math	128.27	22.92
8th grade spring	Reading	175.03	25.29
	Math	144.80	20.70

Table 3. Bivariate correlation matrix of manifest variables.

		Kindergarten fall		Kindergarten spring		1st grade fall		1st grade spring		3rd grade spring		5th grade spring		8th grade spring	
Testing occasion	Domain	Reading	Math	Reading	Math	Reading	Math	Reading	Math	Reading	Math	Reading	Math	Reading	Math
Kindergarten fall	Reading	1.00													
	Math	0.70	1.00												
Kindergarten spring	Reading	0.82	0.64	1.00											
	Math	0.62	0.82	0.65	1.00										
1st grade fall	Reading	0.79	0.63	0.89	0.61	1.00									
	Math	0.53	0.68	0.54	0.76	0.62	1.00								
1st grade spring	Reading	0.67	0.63	0.75	0.62	0.80	0.57	1.00							
	Math	0.52	0.71	0.54	0.76	0.54	0.74	0.63	1.00						
3rd grade spring	Reading	0.52	0.58	0.56	0.59	0.58	0.57	0.70	0.61	1.00					
	Math	0.48	0.66	0.50	0.71	0.51	0.68	0.59	0.77	0.70	1.00				
5th grade spring	Reading	0.47	0.55	0.51	0.57	0.53	0.55	0.66	0.60	0.84	0.70	1.00			
	Math	0.46	0.61	0.48	0.66	0.47	0.65	0.57	0.73	0.67	0.86	0.72	1.00		
8th grade spring	Reading	0.42	0.49	0.43	0.50	0.45	0.50	0.55	0.56	0.74	0.64	0.78	0.67	1.00	
	Math	0.42	0.55	0.44	0.60	0.44	0.60	0.53	0.66	0.66	0.80	0.70	0.85	0.72	1.00

Note. All bivariate correlations significant at $p < .001$.

Analysis details

We restricted our sample to the 2,134 students (10% of the total sample) who had complete data.[3] We used PROC NLMIXED in SAS 9.3 using nonadaptive Gaussian quadrature with four quadrature points and conjugate gradient optimization to estimate the model. Although more quadrature points are often recommended (e.g., Lesaffre & Spiessens, 2001; Littell, Milliken, Stroup, Wolfinger, Schabenberger, 2006), the 10 random effects in the model (seven first-order latent variables, three subject-specific growth factors) and the large sample size created an extremely difficult estimation problem—even with only four quadrature points, the model took a few hundred hours to converge.[4] Using linearization estimation methods (such as FIRO) was also not an option to expedite estimation because the model can only be specified through the GENERAL distribution in PROC NLMIXED and FIRO only is permissible for the NORMAL distribution (Littell et al., 2006). Pinheiro and Chao (2006) showed that the difference in estimates between three and seven quadrature points (a more commonly recommended number) is fairly mild; their simulations showed that fixed effect parameters differed by about 2%–3% while variance parameters differed by 5%–6%. Starting values were obtained by taking the final estimates from either SLCM specification as run in M*plus* or from the first-order model fit by itself, depending on the parameter.

Because the empirical variance grew as time proceeded (see standard deviations in Table 2), it made sense to model the residual errors as a heterogeneous diagonal (all diagonal elements are uniquely estimated but off-diagonal are constrained to 0) rather than a homogeneous diagonal (all diagonal terms are constrained to be equal and off-diagonal elements are constrained to 0), which is the only preprogrammed structure contained in PROC NLMIXED. Therefore, we followed demonstrations from Harring and Blozis (2014) that show how to use the GENERAL distribution in PROC NLMIXED to program a user-defined log-likelihood that allows for more complex error structures that are appropriate for these data. The appendix of this article includes SAS NLMIXED code for this analysis as well as M*plus* code for running the model using the alternative SLCM specification (which was not used, except for obtaining starting values for some parameters, for previously stated reasons).

The complete model

The following subsections will provide targeted discussions that focus on specific portions of the model. Before delving into these more nuanced subsections, we will first present the complete second-order nonlinear mixed-effects model specification for the ECLS-K data. Previous sections of this article mentioned the SLCM specification, and although we will not present this in text, full details are included in Appendix A for interested readers.

3 PROC NLMIXED is able to accommodate missing data with full information maximum likelihood; however, the extremely large percentage of missing data coupled with the complexity of the model led to convergence issues. The assumption of missing-at-random missing values was also tenuous because the demonstration model is unconditional. This analysis is for illustrative purposes rather than for making substantive claims, so we restricted the same to stabilize the estimates.

4 We had tried to use seven and ten quadrature points as well, but SAS on our university's statistical software server had insufficient memory to carry out the number of computations needed to estimate even one iteration. If data sets are much larger, additional random effects are required, or the growth trajectory has additional parameters that require random effects, traditional personal computing environments may not be equipped to estimate such models. Fewer quadrature points could be used (e.g., the popular Laplace approximation is essentially equal to adaptive Gaussian quadrature with a single quadrature point and maintains desirable statistical properties in large samples), but researchers should note that using fewer quadrature points leads to less stable and less precise estimates. We also ran our model with two quadrature points and convergence was reached in about 3.5 hours; estimates were noticeably different but still in the same general vicinity.

The first-order model consists of seven latent variables (one for each timepoint), each with two indicators (vertically scaled mathematics and reading scores). The latent variables are assigned scale using the referent indicator method by setting the loading from each latent variable to the mathematics score to 1.0. In equation form, this can be written as

$$\begin{bmatrix} MathFallK \\ ReadFallk \\ \vdots \\ MathSpringG8 \\ ReadSpringG8 \end{bmatrix} = \begin{bmatrix} 0 \\ \tau_2 \\ \vdots \\ 0 \\ \tau_{14} \end{bmatrix}$$
$$+ \begin{bmatrix} 1 & & \\ \lambda_{2,1} & & \\ & \ddots & \\ & & 1 \\ & & \lambda_{14,7} \end{bmatrix} \begin{bmatrix} Academic\,Ability_1 \\ Academic\,Ability_2 \\ \vdots \\ Academic\,Ability_6 \\ Academic\,Ability_7 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_{13} \\ \varepsilon_{14} \end{bmatrix}$$
$$\textbf{Assessment Scores}_{it} = \boldsymbol{\tau}_t + \boldsymbol{\Lambda}_t \; (\textbf{Academic Ability})_{it} + \boldsymbol{\varepsilon}_{it} \tag{5}$$

where

$$\textbf{Academic Ability} \sim MVN\left(\begin{bmatrix} \nu_1 \\ \vdots \\ \nu_7 \end{bmatrix}, \begin{bmatrix} \psi_{11} & \cdots & \psi_{17} \\ \vdots & \ddots & \vdots \\ \psi_{71} & \cdots & \psi_{77} \end{bmatrix} \right) \tag{6}$$

$$\boldsymbol{\varepsilon} \sim MVN\left(\begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} \theta_{1,1} & 0 & \cdots & 0 \\ 0 & \theta_{2,2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \theta_{14,14} \end{bmatrix} \right). \tag{7}$$

Readers may note that there are only two indicators per factor in the preceding equations, which is typically not ideal. In this context, two indicators should generally not be problematic for two reasons. First, the sample size with this data set is rather large. A simulation by Marsh, Hau, Balla, and Grayson (1998) showed that two indicator models performed essentially identically to six indicator models with sample sizes reminiscent of those in this data, especially if indicators were parceled, which is the case in the ECLS-K data because the test scores at each year are taken from multiple items. Second, in the framework devised by Little, Lindenberger, and Nesselroade (1999), these two indicators would be classified as being low on selection diversity because they are fairly related to each other and are strongly related to the intended latent variable (academic ability). In their simulation, Little et al. (1999) found that models with two indicators and moderate selection diversity performed more desirably than six-indicator models with high selection diversity.

After establishing longitudinal invariance, the estimated loadings and manifest intercepts in Equation (5) (the λ and τ parameters) are constrained to be equal to one another (described in more detail in the following). Then, the second-order growth model is applied to the first-order factors (when moving from the first-order model only to the complete model, the factor covariance is removed from Equation [6] and is explicitly modeled with the second-order model). As will be discussed in more detail in the following, the Michaelis-Menten growth model fit best, meaning that the second-order model can be written as

$$Academic\,Ability_{it} = \beta_{0i} + \frac{(\beta_{Ui} - \beta_{0i})t}{\beta_{Ri} + t} + d_{it} \tag{8}$$

where

$$\begin{aligned} \beta_{0i} &= \alpha_0 + \zeta_{0i} \\ \beta_{Ui} &= \alpha_U + \zeta_{Ui} \\ \beta_{Ri} &= \alpha_R + \zeta_{Ri} \end{aligned} \tag{9}$$

and

$$\boldsymbol{\zeta}_i \sim MVN\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} g_{00} & g_{0U} & g_{0R} \\ g_{U0} & g_{UU} & g_{UR} \\ g_{R0} & g_{RU} & g_{RR} \end{bmatrix} \right) \tag{10}$$

$$\mathbf{d}_i \sim MVN(\mathbf{0}_7\,,\; diag(\sigma_1{}^2 \;\; \sigma_2{}^2 \;\; \sigma_3{}^2 \;\; \sigma_4{}^2 \;\; \sigma_5{}^2 \;\; \sigma_6{}^2 \;\; \sigma_7{}^2\,)). \tag{11}$$

The model-implied mean structure and model-implied covariance structure for the model are worked out in Appendix B for interested readers.

The first-order model and longitudinal invariance

We will not delve into details about the assessment of longitudinal invariance because the readership of *Multivariate Behavioral Research* is likely familiar with the process. Because of the large sample size (Brannick, 1995; Meade, Johnson, & Braddy, 2008), we used the Cheung and Rensvold (2002) method to assess weak and strong invariance. This method takes the difference between the fit indices from a constrained and unconstrained model, and ΔRMSEA values less than or equal to 0.01 and ΔCFI values greater than or equal to -0.01 suggest invariance. The test for weak longitudinal invariance resulted in a ΔRMSEA of 0.006 and ΔCFI of -0.004, and a test for strong longitudinal invariance resulted in a ΔRMSEA of -0.003 and ΔCFI of -0.007. Because the vertically scaled reading scores and vertically scaled math scores are on the same scale and because the first-order latent variables were given scale by setting the mean of the math scores to 0, the estimated intercepts of the reading scores were quite close to 0 (because they are the relative distance between math and reading scores at a given timepoint). We tested a model where the reading score intercepts were also constrained to 0, and the fit was not appreciably worse than if the intercepts were estimated (ΔRMSEA = 0.005, and ΔCFI = −0.004). The overall fit of the model after all constraints were applied was RMSEA = 0.147, 90% CI [0.144, 0.148)], SRMR = 0.071, CFI = 0.857.

According to these findings and the magnitude of the loadings (when standardized loadings exceed 0.70 the Hu and Bentler, 1999, cutoffs are too strict so they are not necessarily a fair benchmark; Cole & Preacher, 2014; Hancock & Mueller, 2011; Heene et al., 2011; Kang, McNeish, & Hancock, 2016; Miles & Shevlin, 2007; Saris, Satorra, & van der Veld, 2009; Savalei, 2012), the model demonstrates good fit based on the SRMR and borderline good fit based on the CFI. The minimum fit function chi-square was significant [χ^2 (69) = 3,250, $p < .01$] although the large sample size may render this statistic overpowered for practical purposes (e.g., Hu & Bentler, 1998).

Although we did not originally intend to include residual covariances, the fit of the model could be greatly improved if residual covariances are added between all the manifest math variables and between all the manifest reading variables, RMSEA = 0.055, 90% CI [0.053, 0.056], SRMR = 0.045, CFI = 0.989. There is disagreement about how defensible residual covariances are (Cole, Ciesla, & Steiger, 2007), especially when their inclusion is empirical rather than theoretical (Grilli & Varriale, 2014). Wolfinger (1996) recommends to use the simplest structure that provides reasonable fit. Given that this example is illustrative and that we do not want to potentially oversell the fit of the model by including residual covariances on empirical grounds, we proceed using the heterogeneous diagonal structure because we find the overall fit to still be reasonable, especially given the strong loadings.

Second-order nonlinear growth model and model estimates

Choosing a model

After longitudinal invariance has been established, then a second-order growth model can be placed upon the first-order factors (Hancock et al., 2001; McArdle, 1988). As with all types of growth models, selecting the appropriate functional form for the second-order growth model is vital to obtaining trustworthy estimates and inferences from the model. There are multiple guidelines in the literature for selecting a nonlinear growth function, given the phenomena being modeled. A recent article by Wood

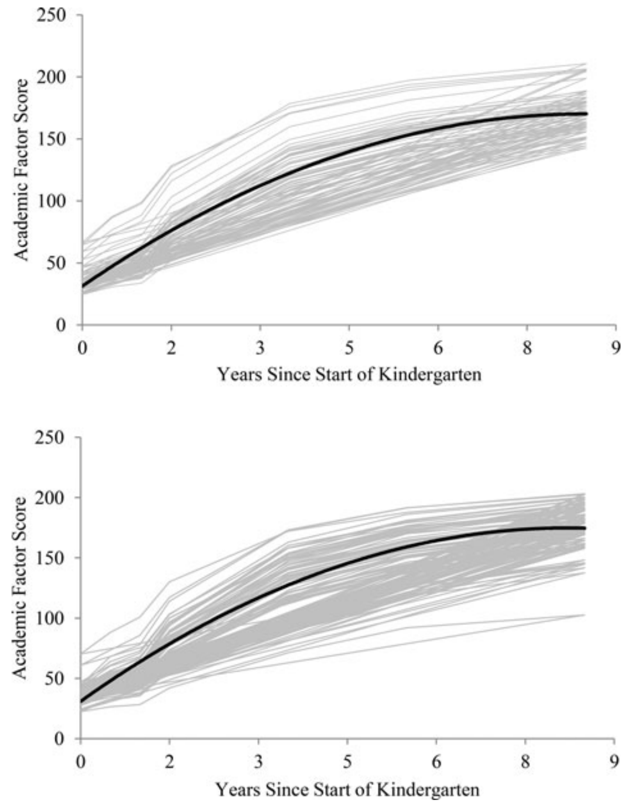

Figure 3. Trajectory plots for first-order factor scores for two random samples of 50 individuals with a superimposed sample mean trajectory.

et al. (2015) contains an overview of the advantageous and disadvantageous aspects of several different models and suggests an empirically based approach where model-data fit is a primary consideration. Alternatively, guidelines by Cudeck and Harring (2007) provide three theoretically driven criteria: (a) the ability of the model to fit well to the data, (b) relevant substantive interpretability of the model parameters, and (c) whether the behavior of the function complements the data (i.e., whether the properties of the function match the data with respect to aspects such as inflection points, asymptotes, etc.). Cudeck and Harring (2007) also note that "graphical displays are irreplaceable in judging whether a function is appropriate," and for large samples, they recommended plotting a subset of the data for inspection.

We will blend the recommendations from Wood et al. (2015) and Cudeck and Harring (2007) by choosing a small number of models that align well with our theoretical modeling goals and will then compare the empirical fit of these models. Keeping with best recommended practice, Figure 3 shows the empirical trajectory plots for the growth in the latent Academic Ability factor from the first-order model over time for two random samples of 50 students. Fitting with how development in younger children is conceived (e.g., Craik & Bialystok, 2006; Goswami, 2011), the growth of the Academic Ability factor over time appears to be nonlinear such that it increases at a decreasing rate and the trajectory does not appear to have a noticeable inflection point or relevant lower asymptote. From the properties of the growth models discussed previously, the Michaelis-Menten and von Bertalanffy curves seem to fit this type of trajectory most closely theoretically and based upon the empirical plot. We will fit models using these curves to obtain a more formal empirical comparison of model fit. Although the plots do not seem to behave consistent with a trajectory that includes an inflection point or lower asymptote, we will also include the Gompertz curve in our comparison because it is possible that an inflection point may exist but it may not be visually apparent from the samples plotted in Figure 3.

Therefore, to determine which type of nonlinear trajectory may be most appropriate for the first-order factor scores, we fit the model without subject-specific growth parameters (i.e., no random effects) using the Gompertz, von Bertalanffy, and Michaelis-Menten trajectories, similar to the process outlined by Harring et al. (2012). With the ECLS-K data, the Michaelis-Menten trajectory yielded a lower Bayesian information criterion (BIC; 202,518) compared to the Gompertz (206,786) or von Bertalanffy (210,845) trajectories. Also important to note, as mentioned earlier, the Michaelis-Menten model provides the most substantively interpretable parameters because the rate parameter is interpreted as the point at which growth is halfway between the intercept and asymptote, which gives more rich substantive information than either the von Bertalanffy or Gompertz models. We will therefore focus on the Michaelis-Menten function for the remainder of this section because it satisfies our substantive needs and fits well empirically.[5] In addition, now that the second-order portion of the model is of interest, we switch software to SAS so that we can fit the NLME model for the aforementioned reasons.

Model results

Keeping with the illustrative purpose of this example, we will fit the full unconditional growth model such that there are no predictors for the growth factors. However, we will return to the possibility of predictors on growth parameters in the discussion section.

The results of the model are presented in two tables where information pertaining primarily to the first-order measurement model (unstandardized loadings, standardized loadings, the total factor variance, the manifest residual variance, the total manifest variance, and the percent of the manifest variance that is explained by the model) are reported in Table 4, and information pertaining primarily to the second-order growth model (first-order factor error variances, first-order factor variance explained, growth parameters fixed effects and variance components, and random-effect correlations) are reported in Table 5.

The initial value (α_0; estimate $= 22.82$) and capacity (α_F; estimate $= 249.47$) estimates are on the scale of the latent variable, which is assigned via constraints on indicator variables. This makes these values more informative for relative comparisons rather than in an absolute sense (i.e., a capacity of 250 in isolation does not give a great deal of information because the scale is interpreted somewhat arbitrarily depending on which indicator variable's mean is constrained for scale. A value of 250 does become meaningful if compared to another individual's value). Initial value ($g_{00} = 45.12$) and capacity ($g_{UU} = 2{,}498.02$) have significant variability across individuals. The mean of the rate parameter ($\alpha_R = 9.07$) has a direct interpretation on the time scale: in the unconditional model, on average, the midpoint between students' initial value and capacity occurs about 9 years after starting kindergarten. Also note the rather modest correlations of the growth parameter random effects—individuals' initial values in kindergarten are correlated with their capacity at 0.24. Individuals who had higher capacity estimates tended to reach their midpoint of growth slightly later in their development as evidenced by a correlation of 0.17 (i.e., higher rate parameters are associated with late bloomers). Initial value and the rate parameter were weakly negatively related with a correlation of -0.15. It is important to note that such correlations among parameter estimates may be highly meaningful for substantive research on the development of academic abilities.

Regarding the fit of the overall model, one drawback of the NLME specification is that there is no consensus for how to judge the global fit of the model (Tang, Slud, & Pfeiffer, 2014; in the SLCM context, this can be done easily with standard SEM fit criteria such as SRMR, RMSEA, CFI, etc.). Common model fit metrics for these models include information criteria, significance tests of fixed-effect parameters, variance explained metrics, or graphical plots (Pan & Lin, 2005). Table 5 reports the BIC and the variance explained for the various variables in the model.

[5] As noted by an anonymous reviewer, more general models that feature the Michaelis-Menten model as a special case may also be reasonable choices for this data. To maintain a streamlined narrative, we proceed in the main text by only discussing the results for the Michaelis-Menten model. However, we also fit a Morgan-Mercer-Flodin model to the data, which features Michaelis-Menten as a special case. The Morgan-Mercer-Flodin parameter estimates were quite similar to the Michaelis-Menten estimates, and the additional parameters in the Morgan-Mercer-Flodin were close to the values the values that would reduce the model to the Michaelis-Menten model. Interested readers are referred to Appendix C for a more detailed discussion of the Morgan-Mercer-Flodin model and the results of fitting this model to the ECLS-K data.

Table 4. Parameter estimates for measurement model factor loadings, total variances, and explained variance at the manifest level.

Measurement model							
Time	Indicator	Unstd. loading	Std. loading	Total fac. var.	Residual var.	Total var.	% Manifest var. exp.
Fall kindergarten	Math	1.00[a]	0.87	60.14	19.71	79.85	75
	Read	1.22	0.84		32.85	109.64	70
Spring kindergarten	Math	1.00[a]	0.85	103.91	41.33	145.24	72
	Read	1.22	0.82		66.56	199.24	67
Fall Grade 1	Math	1.00[a]	0.83	145.09	66.43	211.52	69
	Read	1.22	0.79		111.39	296.66	63
Spring Grade 1	Math	1.00[a]	0.84	254.05	110.50	364.55	70
	Read	1.22	0.79		197.60	522.00	62
Spring Grade 3	Math	1.00[a]	0.88	466.72	140.56	607.28	77
	Read	1.22	0.85		229.66	825.61	72
Spring Grade 5	Math	1.00[a]	0.87	443.00	147.46	590.46	75
	Read	1.22	0.86		207.29	772.96	73
Spring Grade 8	Math	1.00[a]	0.90	411.56	100.87	511.56	80
	Read	1.22	0.83		245.91	770.52	68

Note. Unstd. = unstandardized; Std. = standardized; Total fac. var. = total variance of the first-order factor; Residual var. = residual variance; % manifest var. exp. = percentage of the manifest variable that is explained by the model.
[a]The loading was constrained to this value to give the latent variables scale.

Although not synonymous with global model fit in the SEM sense, the percent of variance explained of both the manifest variables (63% to 80%) and the first-order factors (64% to 90%) was quite high.

Table 5. Variance explained in first-order factors and substantively important second-order parameter estimates for the nonlinear mixed-effects specification.

First-order factors			
Time	Residual variance	Explained variance	% Variance explained
Fall kindergarten	15.14	45.12	81
Spring kindergarten	19.67	84.24	80
Fall Grade 1	29.04	116.05	75
Spring Grade 1	63.48	190.58	64
Spring Grade 3	168.75	297.97	64
Spring Grade 5	116.59	326.41	74
Spring Grade 8	39.54	372.02	90

Second-order growth model		
Parameter name	Parameter	Est.
Initial value	α_0	22.82
Capacity	α_A	249.47
Rate	α_R	9.07
Var (Initial Value)	g_{00}	45.12
Var (Capacity)	g_{UU}	2498.02
Var (Rate)	g_{RR}	9.26
Corr (Initial Value, Capacity)	$r_{(g_{00}, g_{UU})}$	0.24
Corr (Initial Value, Rate)	$r_{(g_{00}, g_{RR})}$	− 0.15
Corr (Capacity, Rate)	$r_{(g_{UU}, g_{RR})}$	0.17

Model fit	
−2Log-likelihood	183,326
No. of parameters	31
BIC	183,563

Note. Correlations between the random effects are reported rather than the covariances for ease of interpretation. *Cov* (Initial Value, capacity) = 81.28; *Cov* (Initial Value, rate) = −3.05; *Cov* (Capacity, rate) = 26.43.

Demonstrative plots

Because the proposed model includes random effects, the most pertinent and rich information is contained within the subject-specific curves. In this section, we will demonstrate how the model can be used to detect the different components of potential at the individual level. We draw attention to three subject-specific curves exemplifying three different types of students: (a) an early bloomer, (b) a fast learner, and (c) a late bloomer. Note that these prototypes were not determined with model-based classification such as a mixture model and were selected on the basis of the authors' own inspection of the subject-specific parameter estimates. The labels given to these prototypes were also given by the authors according to our interpretation of the plots.

Early bloomer

Figure 4 shows an individual in the data who is a prototypical early bloomer, which we informally define as an individual in the bottom quintile on midpoint and capacity but in the top quintile on initial value. More descriptively, an early bloomer's ability is high at early timepoints but slowly regresses back toward the mean at later timepoints. This student has an initial value that is much higher than the average student (65.65 vs. 22.82, 99th percentile), but the midpoint is only 6.58 (20th percentile), meaning that half of the student's capacity is reached years before most other students in the data set. This student maintains above-average ability even up to eighth grade, but other students make up ground on this student in later years, and this student is only predicted to have a capacity of 207.10 (the 20th percentile).

Previous methodologies that consider shorter time spans or that primarily use single-timepoint ability estimates to gauge capacity might erroneously declare that

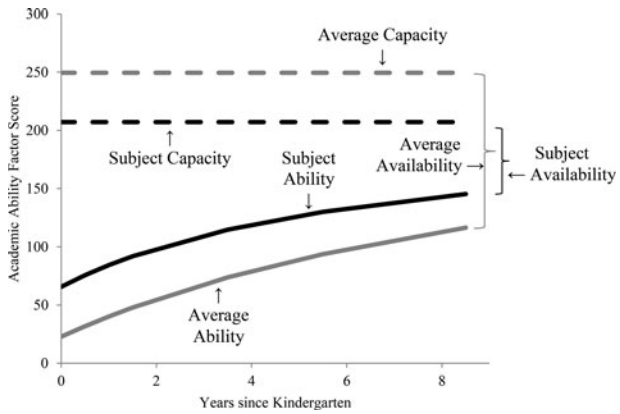

Figure 4. Comparison of an early bloomer's subject-specific potential curve (black) to the average curve (grey). The subject-specific parameter values are 65.65, 6.58, and 207.10 for the initial value, rate, and asymptote parameters, respectively. The average parameter values are 22.82, 9.07, and 249.47 for the initial value, rate, and asymptote parameters, respectively.

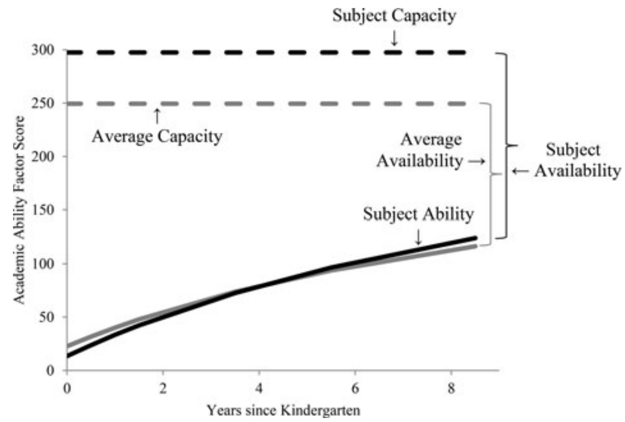

Figure 5. Comparison of fast learner's subject-specific potential curve (black) to the average curve (grey). The subject-specific parameter values are 13.57, 13.33, and 297.56 for the initial value, rate, and asymptote parameters, respectively. The average parameter values are 22.82, 9.07, and 249.47 for the initial value, rate, and asymptote parameters, respectively.

this student has high capacity on the basis of measures taken at a younger age. Rather, this student is predicted to develop below-average capacity despite the fact that his or her ability at younger ages might suggest otherwise. Using the proposed model, teachers or administrators should not be surprised if/when this student begins to regress toward the average in later years of education because the gap between this student and the average student can be seen to narrow as time progresses in Figure 4. Indeed, using previous methodologies, such regression to the mean might be a cause for concern, whereas in our proposed modeling framework it is expected behavior according to this student's estimated trajectory and something that may, more or less, be beyond the control of the educational system.[6]

Moreover, in a teacher evaluation scenario, a teacher with a student like this in class might be adversely and unfairly affected in the evaluation process because traditional models for teacher evaluation (e.g., value-added models) may expect teachers to maintain or increase the high-level performance that this student exhibited early on in his or her education. By contrast, because our proposed model predicts that this student will level off and regress toward his or her peers, teacher evaluations and expectations concerning this particular student using this model may taper so that expectations are more reasonable.

Fast learner

Figure 5 shows an individual in the data who is a prototypical fast learner, which we informally define as an individual in the bottom quintile on initial value but in the top quintile on capacity and midpoint. This student has an initial value that is much lower than the average (13.57 vs. 22.82, 9th percentile) but then grows much quicker than the average student and continues to grow with a predicted capacity that is fairly well above the average capacity (297.56 vs. 249.47, 83rd percentile).

To single-timepoint ability measures, the fast learner prototype appears about average if ability is measured once. However, this can be slightly misleading because the individual started below the other students but then makes up ground quickly and consistently, making performance appear average at an instantaneous point in time. As an analogy, consider a 1,500 meter race in which one of the faster runners stumbles from the block—if the places in the race are assessed 100 meters in, this faster runner may appear average or even below average. But the longer the race continues, the more opportunities the runner has to make up ground on the other participants in the race and conceivably win.

Late bloomer

Figure 6 shows a prototypical late bloomer, which we informally define as an individual in the bottom quintile on initial value, the top quintile on midpoint, and the middle quintile on capacity. This student had a smaller initial ability level compared to the average (13.60, 9th percentile) but has a late midpoint compared to the average (17.68, 99th percentile) and near average capacity (255.87, 56th percentile).

According to Figure 6, up to Grade 8, this student appears to be unambiguously below average if potential were assessed over a short interval. However, in actuality,

[6] Note that the unconditional model does not take other factors into account, so perhaps there is an external circumstance (e.g., problems at home, mental health issues) that the education system may be able to address to prevent this student from regressing.

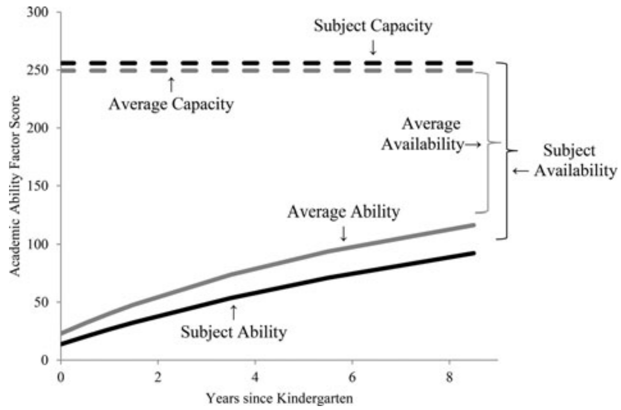

Figure 6. Comparison of a late bloomer's subject-specific potential curve (black) to the average curve (grey). The subject-specific parameter values are 13.60, 17.68, and 255.87 for the initial value, rate, and asymptote parameters, respectively. The average parameter values are 22.82, 9.07, and 249.47 for the initial value, rate, and asymptote parameters, respectively.

this student appears to be developing a capacity for Academic Ability that is quite in line with the average and is merely a very late bloomer that should be able to continue to show noticeable progress in subsequent years of education.

Furthermore, this student's high availability follows from the relatively low ability and average capacity, which marks him or her as a student who may not necessarily be meeting potential by eighth grade. Therefore, to assist in meeting capacity in the future, this student may be a possible candidate for instructional interventions, specialized educational tracking, or at least a watchful eye. Using typical measures based on ability capture at a single timepoint, the student may be unambiguously targeted for remedial or lower-track courses; however, on the basis of the estimated capacity, this may be detrimental to the student's development because the student should have no issue with material that other average students receive.

Discussion

In this investigation, we have presented a modeling framework that incorporates three components of potential: ability, capacity, and availability. Historical attempts to measure potential directly with traditional measurement models have been fleeting endeavors because these models are suitable for measuring developed constructs and encounter difficulties estimating developing constructs, such as a student's capacity. Therefore, psychologists have concluded that potential requires multiple measurements and have generally strayed away from traditional measurement models such as confirmatory factor analyses in favor of methods like dynamic assessment. However, as we have discussed, confirmatory factor analysis is still a relevant first-order method provided that it is accompanied by a second-order model to capture the developing nature of potential. Furthermore, although dynamic assessment has been shown to be effective for individual students in the clinical setting (Feuerstein et al., 1979; Tzuriel, 2001), it would be highly inefficient to implement on a large scale and cannot be applied to the wealth of secondary data sets that are commonly used to address questions in educational and behavioral sciences because researcher intervention is required.

Our proposed model simplifies the conceptualization and measurement of potential because psychologists do not have to consider ways to directly measure potential but rather only need to be able to measure the ability for which potential is of interest. For example, in our illustrative example where academic potential was of interest, we only needed to devise a measure for academic ability at one timepoint and repeat the measurements multiple times, which is a more manageable task compared to conceptualizing a direct, reliable measure for a latent academic potential construct. By virtue of the model parameterization, the second-order portion of the model quantifies the components of potential—capacity, ability, and availability (as a linear function of capacity and ability). In essence, if potential is of primary interest to a researcher, the second-order growth portion of our proposed model serves more so as a conceptual extension to the measurement model and provides estimates that can be readily interpreted by researchers. Put another way, the growth portion of the model becomes ingrained within and central to the measurement model.

Once the model is run, researchers can calculate and save each individual's estimate of capacity or availability and use them as outcome or explanatory variables in subsequent analysis in a two-step estimation process (i.e., Step 1: estimate the capacities/abilities/availabilities; Step 2: estimate a separate statistical model using capacities/abilities/availabilities as variables). For example, an interesting extension may be to compare estimates of capacity to standardized admission test scores (which measures ability at only a single timepoint) as predictors of success in college performance (for a recent example of a study that extracts latent variables first and then applies a complex model to them as a second stage of estimation, see Galatzer-Levy, Karstoft, Statnikov, & Shalev, 2014). Models could also be built using cognitive, motivational, or psychosocial variables as covariates to explain why some individuals more closely approach their capacity while others fall short or to address why some individuals learn more quickly than others. These estimates could also conceivably be used as control variables in models for teacher effectiveness.

Depending on the research interest, these types of questions could also be addressed using a single estimation step by adding covariates directly into the model, a process that some studies have shown to be more advantageous to recover latent abilities (e.g., Curran et al., 2014). As an extension of the single-stage estimation process, mixture components could also be added to the model in attempt to classify the trajectories of different subpopulations. This type of mixture could be specified to give different growth parameter estimates for a single model or could be specified so that different groups actually have different trajectories completely (e.g., subgroup 1 follows a Michaelis-Menten curve; subgroup 2 follows a Gompertz curve). An apparent area of application would be special education where some students' growth tends to differ from the general population. This is related to the suggestion of Wood et al. (2015) to test the dimensionality of the data to inspect whether the sample appears to follow a common trajectory.

We also point out that the proposed model is flexible enough to accommodate growth trajectories that may not necessarily be monotonic. For example, if one is studying an aging population or is interested in declining constructs, as is common in the field of gerontology (e.g., Yilmaz et al., 2016), our proposed model can accommodate this by changing the parameterization of the second-order model. For example, Cudeck and du Toit (2002) outlined a reparameterized quadratic growth model that estimates a maximum point but allows the curve to decline at later timepoints. Put more succinctly, our model is not constricted to only increasing functions and can be extended to other trajectories provided that researchers adapt the second-order model to the nature of the phenomenon under investigation.

Moreover, as one possible future application of this model to substantive questions within educational psychology, as previously mentioned, this modeling framework provides estimates (i.e., capacity and availability) that are theoretically related to Vygotsky's ZPD (Minick, 1987; 2005). Therefore, substantive researchers who are interested in Vygotsky's theory may be interested in utilizing this model to test hypotheses concerning the ZPD. Indeed, although Vygotsky's theory has proven to be useful in educational psychology (Daniels, 2005), the specific predictions of the theory remain largely untested from a quantitative standpoint. Therefore, this modeling framework may be useful in theory testing within the general fields of educational and psychological science.

Current measurement models (somewhat inappropriately) infer future values from current and/or past measurements, which can present problems when dealing with developing constructs. The proposed modeling framework is a direct realization of the many instances in the methodological literature where reparameterizing growth models creates more meaningfully interpretable parameters (Cameron, Grimm, Steele, Castro-Schilo, & Grissmer, 2015; Cudeck & du Toit, 2002; Harring et al., 2012; Molenaar & Campbell, 2009; Preacher & Hancock, 2015; Sterba, 2014) and extends previous calls from the literature to integrate measurement models and growth models into one larger simultaneous process (Bauer & Curran, 2015; Curran et al., 2014). We hope that our proposed framework can help to push the methodological literature toward more complete and more appropriate models for this type of phenomenon and to increase awareness of the fact that measurement and growth often belong hand in hand rather than being compartmentalized into two distinct types of modeling.

Limitations and further extensions

We are proposing a new framework for modeling academic potential and related developing constructs, and there are only so many details that can be covered in a single article. As a result, many more questions about the utility of this framework exist, and there are many limitations to the example we have used in the current study, especially considering the complexity of the model as it is susceptible to issues that plague measurement models, growth models, and nonlinear models, simultaneously.

First, as with other psychometric models, the ultimate functional form of the model is vital to the estimates the model produces. For example, in the context of unidimensional item-response theory, latent ability estimates from a 1PL, 2PL, or 3PL model will almost certainly differ with empirical data. With our proposed model, the same caveat certainly applies, and implementations of the model must be cautious that the proper functional form is applied. That is, capacity estimates from a Michaelis-Menten curve, a von Bertalanffy curve, or a Gompertz curve will likely differ, so there must be a strong rationale behind the choice of curve and why it is appropriate for the data. Furthermore, there are many other types of curves that may be relevant that we did not consider in this article. Wood et al. (2015) gave a highly readable and comprehensive discussion of different types of curves that we did not cover in much detail here (e.g., Richards curves, Schnute curves, Janoschek curves, and Morgan-Mercer-Flodin curves). We highly recommend this article to readers for an overview of the types of curves that may be relevant for psychological processes (Figure 3 and Table 3 in this article offer particularly useful summaries). We reiterate that the general framework we propose is flexible enough to accommodate a variety of curves, but it will ultimately be the responsibility of the researcher to select the appropriate curve for the process of interest.

Related to the previous point, the length of time in which individuals are followed and how closely ability approaches capacity at the final timepoint is always a highly important consideration. If ability at the final timepoint is still rather far from where ability approaches its upper asymptote, this may call into question whether one has chosen the appropriate growth processes—a question that cannot be answered empirically. For example, in the academic ability data that we used to illustrate our model, the ability at Grade 8 was roughly half of the capacity estimate, on average. It is plausible, for example, that there is an inflection point during the high school years where academic ability rapidly accelerates or decelerates as it approaches the capacity estimate. This type of change occurs beyond the observation window and therefore cannot be identified empirically. If this were the true trajectory, then the Michaelis-Menten curve that appeared to fit fairly well from kindergarten to Grade 8 may not actually be able to yield valid capacity estimates, and a more complex curve that allows for inflection points (such as the generalized Michaelis-Menten model from Table 1) would be required.

Coherent with the subject-specific angle taken throughout this article, as outlined by Lopez et al. (2000), these inflection points can also have random effects so that the location at which the growth trajectory changes varies for each person. If such inflection points did exist, the observed timepoint's proximity to the point of inflection would affect the ability of the model to detect the inflection. For example, in the ECLS-K illustrative example, there was a three-year gap between the last two measurement occasions, so if an inflection point exists at the fall of Grade 7 (which may be plausible because this point marks when some students transition from elementary school to middle school), then there may not be sufficient information to detect a change in the curve.

As a general problem inherent with growth models that aim to predict future events or behaviors, the further the observed timepoints are from the future event, the more uncertain the estimates will be. As applied to our model, the further the final ability estimate is from the time where ability developmentally tapers off toward the asymptote, the higher the standard error estimates on the capacity estimate will be. If the standard error estimates became too large, this could be highly problematic in practice because it would be difficult to defend using capacity estimates for decision-making purposes, for example, if there was significant overlap in the confidence limits among various individuals in the data. However, the answer to the question "how close is close enough?" between the final observed time and asymptotic behavior is beyond the scope of the current study, and future research would be necessary to determine at what point in the growth process the standard error estimates are sufficiently small for the estimates to be useful. Simulation work to address this issue could be a valuable future direction. Alternatively, it is possible for the growth trajectory to shift its functional form at a certain point in time (Banks, 1993). If there are no observed data near the point where asymptotic behavior begins to appear, then trajectories at later times are largely speculative.

Regarding the measurement model, the resulting estimates of capacity are only as good as the measurement model allows. With difficult-to-measure or more loosely defined constructs (e.g., creativity), deficient measurement in the first-order model cannot be rescued by the second-order model—the second-order model reflects the quantities of the first-order model. In addition, because of the number of interdependencies and the complexity of the model, poor measurement quality could lead to convergence problems or imprecise comparisons of different types of trajectories in the second-order portion of the model. For example, a more complex but necessary type of growth curve may fail to converge, leading to the selection of an overly simplistic curve, which may lead to less-nuanced inferences or less-granular conclusions about the true behavior of the phenomenon of interest compared to if one had larger samples and/or higher quality data. Alternatively, poor measurement of the first-order factors would lead to poor estimates of first-order factor scores, which could similarly adversely affect selection of the proper functional curve of the second-order growth model (as well as increase uncertainty in the form of larger standard error estimates). This of course begs the question of how reliable is reliable enough in the first-order measurement model. Our guess would be that the standardized loadings would need to be rather high, at least in the 0.70 to 0.80 range, to ensure that the first-order factors are measured well enough and to keep the uncertainly to a minimum. However, a definitive answer would likely require simulation work so that the population model could be known and misspecification of the second-order growth function could be easily tracked. Simulation studies would also be able to quantify the adverse effects of standard error estimates and second-order model selection with varying levels of reliability. Longitudinal invariance is also a minimum requirement for the model, which is not always achievable in practice.

Although the primary goal of this article was not to make substantive claims from the ECLS-K data, several extensions from this example spring to mind. First, although we decided to model the growth at the second-level with a Michaelis-Menten model, due to the computational resources that would be required, we did not fit the full Gompertz and von Bertalanffy curves (the model comparison was done with a reduced model). Related to the high computational cost of estimating the model, as with all NLMEs, starting values play a central role to the quality of the estimates as well as the speed of the estimation process (e.g., Lindstrom & Bates, 1990). We used approximations or submodels of the fully nonlinear model to provide starting values although one could also use graphical methods for such a purpose. In physical sciences, Eadie-Hofstee (Hofstee, 1959) diagrams plot partial derivatives (typically as a function of time in the context of changes over time) used to gauge the form of the relation while also serving to avoid the arduous Gaussian quadrature estimation. Although these methods may not suffice to replace this estimation in the behavioral sciences because there is much more emphasis on the subject-specific random effects, these graphical approaches are uncommon in behavioral sciences despite their utility and may be informative, for example, in narrowing down the type of candidate growth trajectories or in identifying the presence of possible inflection points (Banks, 1993, p. 88).

This study demonstrates the utility of nonlinear growth models for the psychometric estimation of developing constructs such as academic capacity. We have specifically forwarded the second-order Michaelis-Menten function as one highly useful J-shaped growth trajectory that appears, according to the results of this study, to be capable of providing meaningful estimates of each of the identified components of potential. However, while the methodology of the present study was capable of producing these estimates, it was not capable of empirically validating them because the ECLS-K data do not continue into participants' adulthood. For example, while academic capacity asymptotes were generally estimated to be reached around the students' early 20s, we did not have academic or wage data on these participants through early adulthood to test that model-based prediction. In the future, applying this modeling framework to longitudinal data collected during schooling and then validating the model estimates using data collected afterward to determine how well capacity corresponds to final observed behavior will be a critical test. However, such tests are domain specific and will vary depending on the area of application—this is not a mathematical issue concomitant with the model we propose. Indeed, such a future study is necessary for the continuing scientific process associated with understanding student potential, in which rigorously attempting to falsify model predictions must, and will, play a central role. As this line of inquiry progresses, we believe that our proposed modeling framework—perhaps with further methodological refinement—will allow for the empirical examination of the specific factors that influence an individual's ability growth and hence capacity and availability. In this way, student potential may be for the first time the subject of valid, ethical, and fruitful scientific attention.

Article information

Conflict of Interest Disclosures: Each author signed a form for disclosure of potential conflicts of interest. No authors reported any financial or other conflicts of interest in relation to the work described.

Ethical Principles: The authors affirm having followed professional ethical guidelines in preparing this work. These guidelines include obtaining informed consent from human participants, maintaining ethical treatment and respect for the rights of human or animal participants, and ensuring the privacy of participants and their data, such as ensuring that individual participants cannot be identified in reported results or from publicly available original or archival data.

Funding: This work was not supported by any outside funding.

Role of the Funders/Sponsors: None of the funders or sponsors of this research had any role in the design and conduct of the study; collection, management, analysis, and interpretation of data; preparation, review, or approval of the manuscript; or decision to submit the manuscript for publication.

Acknowledgments: The authors thank Brenden Bishop, Bob Cudeck, and Melissa McTernan for helpful comments on an early draft of this article. We also thank Phil Wood for exceptionally useful feedback on a presentation based on this article. We are indebted to Jeffrey Harring for guidance with nonlinear mixed-effects models and for providing invaluable insight on deriving model-implied structures. We also thank three anonymous reviewers and Peter Molenaar for improving the quality of the manuscript through the peer review process. The ideas and opinions expressed herein are those of the authors alone, and endorsement by the authors' institutions is not intended and should not be inferred.

References

Babad, E. Y., & Budoff, M. (1974). Sensitivity and validity of learning-potential measurement in three levels of ability. *Journal of Educational Psychology*, *66*, 439–447. doi:10.1037/h0036510

Baltes, P. B., & Kliegl, R. (1992). Further testing of limits of cognitive plasticity: Negative age differences in a mnemonic skill are robust. *Developmental Psychology*, *28*, 121–125. doi:10.1037/0012-1649.28.1.121

Banks, R. B. (1993). *Growth and diffusion phenomena: Mathematical frameworks and applications*. Berlin: Springer-Verlag.

Bauer, D., & Curran, P. (2015). The discrepancy between measurement and modeling in longitudinal data analysis. In J. R. Harring, L. M. Stapleton, & S. N. Beretvas (Eds.), *Advances in multilevel modeling for educational research: Addressing*

practical issues found in real-world applications (pp. 3–38). Charlotte, NC: Information Age.
Binet, A., & Simon, T. (1916). *The development of intelligence in children.* Baltimore, MD: Williams & Wilkins.
Bliese, P. D., & Ployhart, R. E. (2002). Growth modeling using random coefficient models: Model building, testing, and illustrations. *Organizational Research Methods*, *5*, 362–387. doi:10.1177/109442802237116
Blozis, S. A. (2004). Structured latent curve models for the study of change in multivariate repeated measures. *Psychological Methods*, *9*, 334–353. doi:10.1037/1082-989x.9.3.334
Blozis, S. A., & Harring, J. R. (2016a). Understanding individual-level change through the basis functions of a latent curve model. *Sociological Research & Methods.* Advanced online publication. doi:10.1177/0049124115605341
Blozis, S. A., & Harring, J. R. (2016b). On the estimation of nonlinear mixed-effects models and latent curve models for longitudinal data. *Structural Equation Modeling*, *23*, 904–920. doi:10.1080/10705511.2016.1190932
Blozis, S. A., Harring, J. R., & Mels, G. (2008). Using LISREL to fit nonlinear latent curve models. *Structural Equation Modeling, 15*, 346–369. doi:10.1080/10705510801922639
Borsboom, D. (2006). The attack of the psychometricians. *Psychometrika*, *71*, 425–440. doi:10.1007/s11336-006-1447-6
Borsboom, D., Mellenbergh, G. J., & Van Heerden, J. (2003). The theoretical status of latent variables. *Psychological Review*, *110*, 203–219. doi:10.1037/0033-295x.110.2.203
Brannick, M. T. (1995). Critical comments on applying covariance structure modeling. *Journal of Organizational Behavior*, *16*, 201–213. doi:10.1002/job.4030160303
Browne, M. W. (1993). Structured latent curve models. In C. M. Cuadras & C. R. Rao (Eds.), *Multivariate analysis: Future directions 2* (pp. 171–197). Amsterdam, The Netherlands: Elsevier Science.
Browne, M. W., & du Toit, S. H. C. (1991). Models for learning data. In L. M. Collins & J. L. Horn (Eds.), *Best methods for the analysis of change* (pp. 47–68). Washington, DC: American Psychological Association.
Budoff, M., & Friedman, M. (1964). "Learning potential" as an assessment approach to the adolescent mentally retarded. *Journal of Consulting Psychology*, *28*, 434–439. doi:10.1037/h0040631
Budoff, M., Meskin, J., & Harrison, R. H. (1971). Educational test of the learning-potential hypothesis. *American Journal of Mental Deficiency*, *76*, 159–169. doi:10.1093/acprof:oso/9780198206927.003.0002
Cameron, C. E., Grimm, K. J., Steele, J. S., Castro-Schilo, L., & Grissmer, D. W. (2015). Nonlinear Gompertz curve models of achievement gaps in mathematics and reading. *Journal of Educational Psychology*, *107*, 789–804. doi:10.1037/edu0000009
Cattell, R. B. (1987). *Intelligence: Its structure, growth and action*. New York, NY: Elsevier.
Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*, 233–255. doi:10.1207/s15328007sem0902_5
Codd, C. L., & Cudeck, R. (2014). Nonlinear random-effects mixture models for repeated measures. *Psychometrika*, *79*, 60–83. doi:10.1007/s11336-013-9358-9
Cole, D. A., Ciesla, J. A., & Steiger, J. H. (2007). The insidious effects of failing to include design-driven correlated residuals in latent-variable covariance structure analysis. *Psychological Methods*, *12*, 381–398. doi:10.1037/1082-989x.12.4.381
Cole, D. A., & Preacher, K. J. (2014). Manifest variable path analysis: Potentially serious and misleading consequences due to uncorrected measurement error. *Psychological Methods*, *19*, 300–315. doi:10.1037/a0033805
Craik, F. I. M., & Bialystok, E. (2006). Cognition through the lifespan: Mechanisms of change. *Trends in Cognitive Sciences*, *10*, 131–138. doi:10.1016/j.tics.2006.01.007
Cudeck, R., & du Toit, S. H. (2002). A version of quadratic regression with interpretable parameters. *Multivariate Behavioral Research*, *37*, 501–519. doi:10.1207/s15327906mbr3704_04
Cudeck, R., & Harring, J. R. (2007). Analysis of nonlinear patterns of change with random coefficient models. *Annual Review of Psychology*, *58*, 615–637. doi:10.1146/annurev.psych.58.110405.085520
Curran, P. J., McGinley, J. S., Bauer, D. J., Hussong, A. M., Burns, A., Chassin, L., … & Zucker, R. (2014). A moderated nonlinear factor model for the development of commensurate measures in integrative data analysis. *Multivariate Behavioral Research*, *49*, 214–231.
Daniels, H. (2005). Vygotsky and educational psychology: Some preliminary remarks. *Educational and Child Psychology*, *22*, 6–17.
Day, D. V., & Dragoni, L. (2015). Leadership development: An outcome-oriented review based on time and levels of analyses. *Annual Review Organizational Psychology and Organizational Behavior*, *2*, 133–156. doi:10.1146/annurev-orgpsych-032414-111328
Duncan, T. E., & Duncan, S. C. (2004). An introduction to latent growth curve modeling. *Behavior Therapy*, *35*, 333–363. doi:10.1016/s0005-7894(04)80042-x
Ebbinghaus, H., (1985). *Memory. A contribution to experimental psychology*. New York, NY: Dover.
Embretson, S. E. (1984). A general latent trait model for response processes. *Psychometrika*, *49*, 175–186. doi:10.1007/ BF02294171
Embretson, S. E. (1987). Toward development of a psychometric approach. In C. S. Lidz (Ed.), *Dynamic assessment: An interactional approach to evaluating learning potential* (pp. 141–170). New York, NY: Guilford Press.
Ferrer, E., Balluerka, N., & Widaman, K. F. (2008). Factorial invariance and the specification of second-order latent growth models. *Methodology*, *4*, 22–36. doi:10.1027/1614-2241.4.1.22
Feuerstein, R., Rand, Y., & Hoffman, M. B. (1979). *The dynamic assessment of retarded performers: The Learning Potential Assessment Device, theory, instruments, and techniques*. Baltimore, MD: University Park Press.
Fischer, K. W. (1980). A theory of cognitive development: The control and construction of hierarchies of skills. *Psychological Review*, *87*, 477–531. doi:10.1037/0033-295X.87.6.477
Galatzer-Levy, I. R., Karstoft, K. I., Statnikov, A., & Shalev, A. Y. (2014). Quantitative forecasting of PTSD from early trauma responses: A machine learning application. *Journal of Psychiatric Research*, *59*, 68–76. doi:10.1016/j.jpsychires.2014.08.017
Galton, S. F. (1869). *Hereditary genius*. London, UK: Macmillan.
Geiser, C., Keller, B. T., & Lockhart, G. (2013). First- versus second-order latent growth curve models: Some insights

from latent state-trait theory. *Structural Equation Modeling*, *20*, 479–503. doi:10.1080/10705511.2013.797832

Goswami, U. C. (2011). *The Wiley-Blackwell handbook of childhood cognitive development* (2nd ed.). Chichester, UK: Wiley-Blackwell.

Grilli, L., & Varriale, R. (2014). Specifying measurement error correlations in latent growth curve models with multiple indicators. *Methodology*, *10*, 117–125. doi:10.1027/1614-2241/a000082

Grimm, K. J., & Ram, N. (2009). Nonlinear growth models in Mplus and SAS. *Structural Equation Modeling*, *16*, 676–701. doi:10.1080/10705510903206055

Grimm, K. J., Ram, N., & Hamagami, F. (2011). Nonlinear growth curves in developmental research. *Child Development*, *82*, 1357–1371. doi:10.1111/j.1467-8624.2011.01630.x

Grossnickle, E. M., Dumas, D., Alexander, P. A., & Baggetta, P. (2016). Individual differences in the process of relational reasoning. *Learning and Instruction*, *42*, 141–159. doi:10.1016/j.learninstruc.2016.01.013

Hamilton, J. L., & Budoff, M. (1974). Learning potential among the moderately and severely mentally retarded. *Mental Retardation*, *12*, 33–36.

Hancock, G. R., Kuo, W. L., & Lawrence, F. R. (2001). An illustration of second-order latent growth models. *Structural Equation Modeling*, *8*, 470–489. doi:10.1207/s15328007sem0803_7

Hancock, G. R., & Mueller, R. O. (2011). The reliability paradox in assessing structural relations within covariance structure models. *Educational and Psychological Measurement*, *71*, 306–324. doi:10.1177/0013164410384856

Harradine, C. C., Coleman, M. R. B., & Winn, D.-M. C. (2014). Recognizing academic potential in students of color: Findings of U-STARS∼PLUS. *Gifted Child Quarterly*, *58*, 24–34. doi:10.1177/0016986213506040

Harring, J. R. (2009). A nonlinear mixed effects model for latent variables. *Journal of Educational and Behavioral Statistics, 34*, 293–318. doi:10.3102/1076998609332750

Harring, J. R., & Blozis, S. A. (2014). Fitting correlated residual error structures in nonlinear mixed-effects models using SAS PROC NLMIXED. *Behavior Research Methods*, *46*, 372–384. doi:10.3758/s13428-013-0397-z

Harring, J. R., Kohli, N., Silverman, R. D., & Speece, D. L. (2012). A second-order conditionally linear mixed effects model with observed and latent variable covariates. *Structural Equation Modeling*, *19*, 118–136. doi:10.1080/10705511.2012.634729

Harring, J. R., Weiss, B. A., & Hsu, J. C. (2012). A comparison of methods for estimating quadratic effects in nonlinear structural equation models. *Psychological Methods*, *17*, 193–214. doi:10.1037/a0027539

Heene, M., Franzens, K., Hilbert, S., Draxler, C., Ziegler, M., & Bühner, M. (2011). Masking misfit in confirmatory factor analysis by increasing unique variances: A cautionary note on the usefulness of cutoff values of fit indices. *Psychological Methods*, *16*, 319–336. doi:10.1037/a0024917

Hofstee, B. H. J. (1959). Non-inverted versus inverted plots in enzyme kinetics. *Nature*, *184*, 1296–1298. doi:10.1038/1841296b0

Horn, J. L., McArdle, J. J., & Mason, R. (1983). When is invariance not invariant: A practical scientist's look at the ethereal concept of factor invariance. *Southern Psychologist*, *1*, 179–188.

Hu, L. T., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, *3*, 424–453. doi:10.1037//1082-989x.3.4.424

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1–55. doi:10.1080/10705519909540118

Jones, L. V., & Thissen, D. (2007). A history and overview of psychometrics. In C. R. Rao and S. Sinharay (Eds.), *Handbook of Statistics* (Vol. 26, pp. 1–27). Amsterdam, The Netherlands: Elsevier.

Kang, Y., McNeish, D. M., & Hancock, G. R. (2016). The role of measurement quality on practical guidelines for assessing measurement and structural invariance. *Educational and Psychological Measurement*, *76*, 533–561. doi:10.1177/0013164415603764

Laird, A. K. (1964). Dynamics of tumour growth. *British Journal of Cancer*, *18*, 490. doi:10.1038/bjc.1964.55

Lawrence, F. R., & Hancock, G. R. (1998). Assessing change over time using latent growth modeling. *Measurement and Evaluation in Counseling and Development*, *30*, 211–224.

Lesaffre, E., & Spiessens, B. (2001). On the effect of the number of quadrature points in a logistic random effects model: An example. *Applied Statistics*, *50*, 325–335. doi:10.1111/1467-9876.00237

Lindstrom, M. J., & Bates, D. M. (1990). Nonlinear mixed effects models for repeated measures data. *Biometrics*, *46*, 673–687. doi:10.2307/2532087

Littell, R. C., Milliken, G. A., Stroup, W.W., Wolfinger, R.D., & Schabenberger, O. (2006). *SAS for Mixed Models*. Cary, NC: SAS Institute.

Little, T. D., Lindenberger, U., & Nesselroade, J. R. (1999). On selecting indicators for multivariate measurement and modeling with latent variables: When "good" indicators are bad and "bad" indicators are good. *Psychological Methods*, *4*, 192–211. doi:10.1037//1082-989x.4.2.192

Lopez, S., France, J., Gerrits, W. J., Dhanoa, M. S., Humphries, D. J., & Dijkstra, J. (2000). A generalized Michaelis-Menten equation for the analysis of growth. *Journal of Animal Science*, *78*, 1816–1828. doi:10.2527/2000.7871816x

Marsh, H. W., Hau, K. T., Balla, J. R., & Grayson, D. (1998). Is more ever too much? The number of indicators per factor in confirmatory factor analysis. *Multivariate Behavioral Research*, *33*, 181–220. doi:10.1207/s15327906mbr3302_1

McArdle, J. J. (1988). Dynamic but structural equation modeling of repeated measures data. In J.R. Nesselroade & R.B. Cattell (Eds.), *Handbook of multivariate experimental psychology* (pp. 561–614). New York, NY: Plenum Press.

Meade, A. W., Johnson, E. C., & Braddy, P. W. (2008). Power and sensitivity of alternative fit indices in tests of measurement invariance. *Journal of Applied Psychology*, *93*, 568–592. doi:10.1037/0021-9010.93.3.568

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, *58*, 525–543. doi:10.1007/bf02294825

Meredith, W., & Horn, J. (2001). The role of factorial invariance in modeling growth and change. In A. Sayer & L. Collins (Eds.), *New methods for the analysis of change* (pp. 203–240). Washington, DC: American Psychological Association.

Meredith, W., & Tisak, J. (1990). Latent curve analysis. *Psychometrika*, *55*, 107–122. doi:10.1007/bf02294746

Michaelis, L., & Menten, M. L. (1913). Die Kinetik der Invertinwirkung. *Biochem. Z*, *49*, 333–369.

Miles, J., & Shevlin, M. (2007). A time and a place for incremental fit indices. *Personality and Individual Differences*, *42*, 869–874. doi:10.1016/j.paid.2006.09.022

Minick, N. (1987). Implications of Vygotsky's theories for dynamic assessment. In C. S. Lidz (Ed.), *Dynamic assessment: An interactional approach to evaluating learning potential* (pp. 116–140). New York, NY: Guilford Press.

Minick, N. (2005). The development of Vygotsky's thought: An introduction to thinking and speech. In H. Daniels (Ed.), *An introduction to Vygotsky* (2nd ed., pp. 33–57). New York, NY: Routledge.

Mislevy, D. R. J. (1984). Estimating latent distributions. *Psychometrika*, *49*, 359–381. doi:10.1007/bf02306026

Molenaar, P. C., & Campbell, C. G. (2009). The new person-specific paradigm in psychology. *Current Directions in Psychological Science*, *18*, 112–117. doi:10.1111/j.1467-8721.2009.01619.x

Monod, J. (1949). The growth of bacterial cultures. *Annual Review of Microbiology*, *3*, 371–394. doi:10.1146/annurev.mi.03.100149.002103

Morgan, P. H., Mercer, L. P., & Flodin, N. W. (1975). General model for nutritional responses of higher organisms. *Proceedings of the National Academy of Sciences*, *72*, 4327–4331. doi:10.1073/pnas.72.11.4327

Muthén, B. O., & Khoo, S. T. (1998). Longitudinal studies of achievement growth using latent variable modeling. *Learning and Individual Differences*, *10*, 73–101. doi:10.1016/s1041-6080(99)80135-6

Ortega, L., & Iberri-Shea, G. (2005). Longitudinal research in second language acquisition: Recent trends and future directions. *Annual Review of Applied Linguistics*, *25*, 26–45. doi:10.1017/S0267190505000024

Pan, Z., & Lin, D. Y. (2005). Goodness-of-fit methods for generalized linear mixed models. *Biometrics*, *61*, 1000–1009. doi:10.1111/j.1541-0420.2005.00365.x

Petersen, D. B., & Gillam, R. B. (2015). Predicting reading ability for bilingual Latino children using dynamic assessment. *Journal of Learning Disabilities*, *48*, 3–21. doi:10.1177/0022219413486930

Pinheiro, J.C., & Chao, E.C. (2006). Efficient Laplacian and adaptive Gaussian quadrature algorithms for multilevel generalized linear mixed models. *Journal of Computational and Graphical Statistics*, *15*, 58–81. doi:10.1198/106186006x96962

Plucker, J. A. (1999). Is the proof in the pudding? Reanalyses of Torrance's (1958 to present) longitudinal data. *Creativity Research Journal*, *12*, 103–114. doi:10.1207/s15326934crj1202_3

Poehner, M. E., & Lantolf, J. P. (2013). Bringing the ZPD into the equation: Capturing L2 development during Computerized Dynamic Assessment (C-DA). *Language Teaching Research*, *17* (3), 323–342. doi:10.1177/1362168813482935

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*, 84–101. doi: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*, 303–316. doi:10.1177/0165025407077751

Reise, S. P., & Waller, N. G. (2009). Item response theory and clinical measurement. *Annual Review of Clinical Psychology*, *5*, 27–48. doi:10.1146/annurev.clinpsy.032408.153553

Richards, F. J. (1959). A flexible growth function for empirical use. *Journal of Experimental Botany*, *10*, 290–301. doi:10.1093/jxb/10.2.290

Rose, L. T., Rouhani, P., & Fischer, K. W. (2013). The science of the individual. *Mind, Brain, and Education*, *7*, 152–158. doi:10.1111/mbe.12021

Rose, T. (2016). *The end of average: How to succeed in a world that values sameness*. London, UK: Penguin.

Saris, W. E., Satorra, A., & Van der Veld, W. M. (2009). Testing structural equation models or detection of misspecifications? *Structural Equation Modeling*, *16*, 561–582. doi:10.1080/10705510903203433

Savalei, V. (2012). The relationship between root mean square error of approximation and model misspecification in confirmatory factor analysis models. *Educational and Psychological Measurement*, *72*, 910–932. doi:10.1177/0013164412452564

Schaie, K. W., (2005). *Developmental influences on adult intelligence: The Seattle longitudinal study*. New York, NY: Oxford University Press.

Schick, H., & Phillipson, S. N. (2009). Learning motivation and performance excellence in adolescents with high intellectual potential: What really matters? *High Ability Studies*, *20*, 15–37. doi:10.1080/13598130902879366

Schnute, J. (1981). A versatile growth model with statistically stable parameters. *Canadian Journal of Fisheries and Aquatic Sciences*, *38*, 1128–1140. doi:10.1139/f81-153

Seber, G. A. F, and C. J. Wild. (1989). *Nonlinear regression*. New York, NY: Wiley.

Sijtsma, K. (2012). Future of psychometrics: Ask what psychometrics can do for psychology. *Psychometrika*, *77*, 4–20. doi:10.1007/s11336-011-9242-4

Spearman, C. (1927). *The abilities of man: Their nature and measurement*. New York, NY: Macmillan.

Sterba, S. K. (2014). Fitting nonlinear latent growth curve models with individually varying time points. *Structural Equation Modeling*, *21*, 630–647. doi:10.1080/10705511.2014.919828

Sternberg, R. J., Grigorenko, E. L., Ngorosho, D., Tantufuye, E., Mbise, A., Nokes, C., … Bundy, D. A. (2002). Assessing intellectual potential in rural Tanzanian school children. *Intelligence*, *30*, 141–162. doi:10.1016/s0160-2896(01)00091-5

Tang, M., Slud, E. V., & Pfeiffer, R. M. (2014). Goodness of fit tests for linear mixed models. *Journal of Multivariate Analysis*, *130*, 176–193. doi:10.1016/j.jmva.2014.03.012

Thorndike, E. L. (1913). *Educational psychology*. New York, NY: Teachers College, Columbia University Press.

Tiekstra, M., Minnaert, A., & Hessels, M. G. P. (2016). A review scrutinising the consequential validity of dynamic assessment. *Educational Psychology*, *36*, 112–137. doi:10.1080/01443410.2014.915930

Torrance, E. P. (1962). *Guiding creative talent*. Englewood Cliffs, N.J.: Prentice-Hall.

Tourangeau, K., Nord, C., Lê, T., Sorongon, A. G., and Najarian, M. (2009). *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), combined user's manual for the ECLS-K Eighth-Grade and K–8 full sample data files and electronic codebooks (NCES 2009–004)*. Washington, DC: National Center for Education Statistics,

Institute of Education Sciences, U.S. Department of Education.
Tzuriel, D. (2001). *Dynamic assessment of young children*. New York, NY: Kluwer Academic.
Vygotsky, L. S. (1994). *The Vygotsky reader*. Cambridge, MA: Basil Blackwell.
Wolfinger, R. D. (1996). Heterogeneous variance: Covariance structures for repeated measures. *Journal of Agricultural, Biological, and Environmental Statistics*, *1*, 205–230. doi:10.2307/1400366
Wood, P. K., Steinley, D., & Jackson, K. M. (2015). Right-sizing statistical models for longitudinal data. *Psychological Methods*, *20*, 470–488. doi:10.1037/met0000037
Yilmaz, R., Behnke, S., Liepelt-Scarfone, I., Roeben, B., Pausch, C., Runkel, A., … Berg, D. (2016). Substantia nigra hyperechogenicity is related to decline in verbal memory in healthy elderly adults. *European Journal of Neurology*, *23*, 973–978. doi:10.1111/ene.12974

Appendix A: Structured latent curve model specification

The first-order model is unaffected by the decision to model the second-order curve as a structured latent curve model (SLCM) or a nonlinear mixed-effect model (NLME). As such, the first-order measurement model can be written as

$$\begin{bmatrix} MathFallK \\ ReadFallk \\ \vdots \\ MathSpringG8 \\ ReadSpringG8 \end{bmatrix} = \begin{bmatrix} 0 \\ \tau_2 \\ \vdots \\ 0 \\ \tau_{14} \end{bmatrix} + \begin{bmatrix} 1 & & \\ \lambda_{2,1} & & \\ & \ddots & \\ & & 1 \\ & & \lambda_{14,7} \end{bmatrix} \begin{bmatrix} Academic\ Ability_1 \\ Academic\ Ability_2 \\ \vdots \\ Academic\ Ability_6 \\ Academic\ Ability_7 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_{13} \\ \varepsilon_{14} \end{bmatrix}$$

$$\textbf{Assessment Scores}_{it} = \boldsymbol{\tau}_{\mathbf{t}} + \boldsymbol{\Lambda}_t \ (\textbf{Academic Ability})_{it} + \boldsymbol{\varepsilon}_{it} \tag{B1}$$

where

$$\textbf{Academic Ability} \sim MVN\left(\begin{bmatrix} \nu_1 \\ \vdots \\ \nu_7 \end{bmatrix}, \begin{bmatrix} \psi_{11} & \cdots & \psi_{17} \\ \vdots & \ddots & \vdots \\ \psi_{71} & \cdots & \psi_{77} \end{bmatrix}\right) \tag{B2}$$

$$\boldsymbol{\varepsilon} \sim MVN\left(\begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} \theta_{1,1} & 0 & \cdots & 0 \\ 0 & \theta_{2,2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \theta_{14,14} \end{bmatrix}\right). \tag{B3}$$

However, the $\frac{(\beta_{Ui}-\beta_{0i})t}{\beta_{Ri}+t}$ term in the Michaelis-Menten model has parameters that enter the model in a nonlinear fashion (namely the β_{Ri} parameter), meaning that the model cannot be directly fit in structural equation modeling software (Blozis, Harring, & Mels, 2008; Sterba, 2014). The SLCM linearizes the model by setting constraints on the loadings from the second-order growth factors to the first-order factors. The constraints are based upon the first partial derivatives of the target function with respect to each parameter in the target function (Blozis & Harring, 2016a; Browne & du Toit, 1991; Browne, 1993), which is the Michaelis-Menten mean function in this case. The mean response of the Michaelis-Menten target function is $f(\alpha_0, \alpha_U, \alpha_R, t) = \alpha_0 + \frac{(\alpha_U-\alpha_0)t}{\alpha_R+t}$. The partial derivatives are as follows:

$$\frac{\partial}{\partial\alpha_0} f(\alpha_0, \alpha_U, \alpha_R, t) = \frac{d}{d\alpha_0}\alpha_0 + \frac{t\left(\frac{d}{d\alpha_0}\alpha_U - \frac{d}{d\alpha_0}\alpha_0\right)}{\alpha_R + t} = 1 - \frac{t}{t+\alpha_R} \tag{B4}$$

$$\frac{\partial}{\partial\alpha_U} f(\alpha_0, \alpha_U, \alpha_R, t) = \frac{d}{d\alpha_U}\alpha_0 + \frac{t\left(\frac{d}{d\alpha_U}\alpha_U - \frac{d}{d\alpha_U}\alpha_0\right)}{\alpha_R + t} = \frac{t}{t+\alpha_R} \tag{B5}$$

$$\begin{aligned} \frac{\partial}{\partial\alpha_R} f(\alpha_0, \alpha_U, \alpha_R, t) &= (\alpha_U - \alpha_0)t \cdot \frac{d}{d\alpha_R}\left(\frac{1}{\alpha_R+t}\right) + \frac{d}{d\alpha_R}\alpha_0 \\ &= -\frac{\frac{d}{d\alpha_R}(\alpha_R+t)}{(\alpha_R+t)^2}(\alpha_U-\alpha_0)t \\ &= -\frac{\left(\frac{d}{d\alpha_U}\alpha_U + \frac{d}{d\alpha_U}t\right)(\alpha_U-\alpha_0)t}{(\alpha_R+t)^2} \\ &= -\frac{(\alpha_A-\alpha_0)t}{(\alpha_R+t)^2} \end{aligned} \tag{B6}$$

Thus, the second-order growth model as an SLCM for the ECLS-K data would be represented as

$$\begin{bmatrix} \textit{Academic Ability, Fall K} \\ \textit{Academic Ability, Spring K} \\ \textit{Academic Ability, Fall G1} \\ \textit{Academic Ability, Spring G1} \\ \textit{Academic Ability, Spring G3} \\ \textit{Academic Ability, Spring G5} \\ \textit{Academic Ability, Spring G8} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1-\frac{0.5}{0.5+\alpha_R} & \frac{0.5}{0.5+\alpha_R} & -\frac{(\alpha_U-\alpha_0)\,0.5}{(\alpha_R+0.5)^2} \\ 1-\frac{1}{1+\alpha_R} & \frac{1}{1+\alpha_R} & \frac{(\alpha_U-\alpha_0)}{(\alpha_R+1)^2} \\ 1-\frac{1.5}{1.5+\alpha_R} & \frac{1.5}{1.5+\alpha_R} & \frac{(\alpha_U-\alpha_0)\,1.5}{(\alpha_R+1.5)^2} \\ 1-\frac{3.5}{3.5+\alpha_R} & \frac{3.5}{3.5+\alpha_R} & \frac{(\alpha_U-\alpha_0)\,3.5}{(\alpha_R+3.5)^2} \\ 1-\frac{5.5}{5.5+\alpha_R} & \frac{5.5}{5.5+\alpha_R} & \frac{(\alpha_U-\alpha_0)\,5.5}{(\alpha_R+5.5)^2} \\ 1-\frac{8.5}{8.5+\alpha_R} & \frac{8.5}{8.5+\alpha_R} & \frac{(\alpha_U-\alpha_0)\,8.5}{(\alpha_R+8.5)^2} \end{bmatrix} \begin{bmatrix} \alpha_0+\zeta_0 \\ \alpha_U+\zeta_U \\ \zeta_R \end{bmatrix} + \begin{bmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \\ d_5 \\ d_6 \\ d_7 \end{bmatrix}$$

$$\textbf{Academic Ability}_i = \quad \boldsymbol{\Lambda} \quad \textbf{(Growth Parameters)}_i + \mathbf{d}_i \tag{B7}$$

Note that α_R enters the model nonlinearly and is thus estimated as part of $\boldsymbol{\Lambda}$ and does not appear in the vector of individual growth parameters (Sterba, 2014).

Appendix B: Micahelis-Menten model implied mean and covariance structures

As shown in Section 3 of Harring (2009), the conditional model-implied mean and covariance structures for nonlinear growth models for latent variables take the same general form as the model-implied structures were the model linear. In the linear case, for a growth model of the form

$$\begin{aligned} \mathbf{y}_{it} &= \boldsymbol{\tau}_t + \boldsymbol{\Lambda}_t\boldsymbol{\eta}_i + \boldsymbol{\varepsilon}_{it} \\ \boldsymbol{\eta}_i &= \boldsymbol{\alpha} + \boldsymbol{\Gamma}\mathbf{x} + \boldsymbol{\zeta}_i, \end{aligned} \tag{C1}$$

where $\mathbf{y}$ is a vector of manifest outcomes, $\boldsymbol{\eta}$ is a vector of growth factors, $\boldsymbol{\Lambda}$ is matrix of loadings relating the growth factors to the manifest outcomes, ε is a vector of residuals, $\boldsymbol{\alpha}$ is a vector of growth factor means, $\boldsymbol{\Gamma}$ is a matrix of coefficients for the predicted effect of time-invariant covariates on the growth factors, $\mathbf{x}$ is a vector of covariate values, and $\boldsymbol{\zeta}$ is a vector of random effects, the model-implied mean structure and model-implied covariance structure, respectively, are

$$E[\mathbf{y}_i] = \boldsymbol{\tau} + \boldsymbol{\Lambda}(\boldsymbol{\alpha} + \boldsymbol{\Gamma}\boldsymbol{\kappa}) \tag{C2}$$

$$Var[\mathbf{y}_i] = \boldsymbol{\Lambda}(\boldsymbol{\Gamma}\boldsymbol{\Phi}\boldsymbol{\Gamma}^{\mathrm{T}})\boldsymbol{\Lambda}^{\mathrm{T}} + \boldsymbol{\Theta}, \tag{C3}$$

where $\boldsymbol{\mu}$ is a vector of model-implied means of the outcome variables, $\boldsymbol{\Sigma}$ is the model-implied covariance matrix of the outcome variables, Ψ is the covariance matrix of the growth factors, $\boldsymbol{\Theta}$ is a matrix of residual variances and covariances among the repeated measures, $\boldsymbol{\kappa}$ is a vector of covariate means, and $\boldsymbol{\Phi}$ is a covariance matrix of the covariates.

While this general structure is intact, two aspects must be considered: (a) how to adapt the matrices to account for the nonlinear growth trajectory and (b) how to incorporate the second-order portion of the model.

To account for nonlinearity in the growth trajectory, Blozis and Harring (2016a) explain that one must find the basis function for the growth trajectory, which is similarly needed as a first step in the SLCM specification outlined in Appendix A. The general basis function for nonlinear models is $f(t,\ \theta) = \boldsymbol{\Lambda}\boldsymbol{\zeta}$ with $\boldsymbol{\Lambda} = \frac{\partial}{\partial\theta'}f(t,\ \theta)$ for θ a vector of model parameters. In the case of the Michaelis-Menten model, $\theta = (\alpha_0,\ \alpha_U,\ \alpha_R)'$. The basis function is invariant to scaling constants, so the basis function can be rewritten as $f(\theta^*) = \upsilon f(\theta)$where υ is a scaling constant. After scaling, the Michaelis-Menten function can be written as

$$\upsilon\alpha_0 + \frac{(\upsilon\alpha_U - \upsilon\alpha_0)t}{\alpha_R + t} \text{ where } \theta^* = (\upsilon\alpha_0,\ \upsilon\alpha_U,\ \alpha_R)'. \tag{C4}$$

To find elements of $\boldsymbol{\zeta}$ so that $f(t,\ \theta) = \boldsymbol{\Lambda}\boldsymbol{\zeta}$,

$$\boldsymbol{\zeta} = \boldsymbol{\zeta}(\theta) = \left.\frac{\partial\theta^*(\upsilon,\theta^*)}{\partial\upsilon}\right|_{\upsilon=1} = \begin{pmatrix} \upsilon\alpha_0 \\ \upsilon\alpha_U \\ \alpha_R \end{pmatrix} = \begin{pmatrix} \alpha_0 \\ \alpha_U \\ 0 \end{pmatrix}. \tag{C5}$$

The target function can be written as a growth model $f(t,\ \theta) = \boldsymbol{\Lambda}\boldsymbol{\zeta}$ with a general row of $\boldsymbol{\Lambda}$,

$$f(t,\ \theta) = \boldsymbol{\eta}_i = \boldsymbol{\Lambda}\boldsymbol{\zeta} \tag{C6}$$

$$= \left(1 - \frac{t}{t+\alpha_R} \quad \frac{t}{t+\alpha_R} \quad -\frac{(\alpha_U-\alpha_0)t}{(\alpha_R+t)^2}\right)\begin{pmatrix} \alpha_0 \\ \alpha_U \\ 0 \end{pmatrix} \tag{C7}$$

$$= \alpha_0 - \frac{\alpha_0 t}{t+\alpha_R} + \frac{\alpha_U t}{t+\alpha_R} + 0 \tag{C8}$$

$$= \alpha_0 + \frac{(\alpha_U - \alpha_0)t}{\alpha_R + t}, \tag{C9}$$

which shows that $\mathbf{\Lambda} = \frac{\partial}{\partial \boldsymbol{\theta}'} f(t, \boldsymbol{\theta})$reproduces the original Michaelis-Menten structure but in the form of a linear model, meaning that Equations (C2) and (C3) hold despite the fact that the model takes a nonlinear form.

Extending this principle to a second-order model, an additional model equation is needed for the additional order of the model such that

$$\begin{aligned} \mathbf{y}_{it} &= \mathbf{\Lambda}_t \boldsymbol{\eta}_{it} + \boldsymbol{\varepsilon}_{it} \\ \boldsymbol{\eta}_{it} &= \mathbf{\Lambda}_\eta \boldsymbol{\beta}_i + \mathbf{d}_{it} \\ \boldsymbol{\beta}_i &= \boldsymbol{\alpha} + \mathbf{\Gamma}\mathbf{x} + \boldsymbol{\zeta}_i, \end{aligned} \tag{C10}$$

where now $\mathbf{y}$ is a vector of manifest outcomes, $\boldsymbol{\eta}$ is a vector of first-order latent variables, and $\boldsymbol{\beta}$ is a vector of subject-specific growth factors. There are now two loadings loading matrices: $\mathbf{\Lambda}_t$, which includes loadings from the first-order factors to the manifest variables (the measurement model), and $\mathbf{\Lambda}_\eta$, which includes loadings from the growth factors to the first-order factors (the growth model). Substituting this more complex model specification from (C10) into (C2) and (C3) yields the model-implied mean and covariance structures for the model specified in text in Equations (8) through (13). Namely,

$$E[\mathbf{y}_i] = \boldsymbol{\tau} + \mathbf{\Lambda}_t[\mathbf{\Lambda}_\eta(\boldsymbol{\alpha} + \mathbf{\Gamma}\mathbf{x})] \tag{C11}$$

$$Var[\mathbf{y}_i] = \mathbf{\Lambda}_t[\mathbf{\Lambda}_\eta(\mathbf{\Gamma}\Phi\Gamma^{\mathrm{T}} + \mathbf{G})\mathbf{\Lambda}_\eta^{\mathrm{T}} + \mathbf{R}]\mathbf{\Lambda}_t^{\mathrm{T}} + \boldsymbol{\Theta}, \tag{C12}$$

where $\mathbf{G}$ is the covariance matrix of the second-order random effects $\boldsymbol{\zeta}$ and $\mathbf{R}$ is the covariance matrix of the first-order factor residuals $\mathbf{d}$.

Appendix C: Morgan-Mercer-Flodin model for ECLS-K data

In the body of the main text, we reported the results from the Michaelis-Menten growth trajectory in the ECLS-K empirical example. Although we defend the choice of this model due to its interpretable parameters that map well onto the components of potential for which we argue, there are, of course, other types of models that could be parameterized in such a way. As noted by an anonymous reviewer, the Morgan-Mercer-Flodin (MMF; Morgan, Mercer, & Flodin, 1975) model is a more general type of nonlinear growth model that includes the Michaelis-Menten model as a special case. The MMF model is written as

$$Academic\,Ability_{it} = \beta_{Ui} + \frac{(\beta_{Ui} - \beta_{0i})t}{1 + (\beta_{Ri}t)^{\delta}} + d_{it}, \tag{D1}$$

where

$$\begin{aligned} \beta_{0i} &= \alpha_0 + \zeta_{0i} \\ \beta_{Ui} &= \alpha_U + \zeta_{Ui} \\ \beta_{Ri} &= \alpha_R + \zeta_{Ri} \\ \delta &= \alpha_D \end{aligned} \tag{D2}$$

and

$$\zeta_i \sim MVN\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} g_{00} & g_{0U} & g_{0R} \\ g_{U0} & g_{UU} & g_{UR} \\ g_{R0} & g_{RU} & g_{RR} \end{bmatrix}\right) \tag{D3}$$

$$\mathbf{d}_i \sim MVN\left(\mathbf{0}_7, \, diag\left(\sigma_1^2\ \sigma_2^2\ \sigma_3^2\ \sigma_4^2\ \sigma_5^2\ \sigma_6^2\ \sigma_7^2\right)\right). \tag{D4}$$

Much like the Michaelis-Menten model shown in the main text in Equations (8) through (11), β_{Ui} is the upper asymptote for the ith person, and β_{0i} is the initial value for the ith person. The β_{Ri} parameter is similarly a rate parameter although it does not have an identical interpretation to the rate parameter in the Michaelis-Menten model. Namely, the MMF and Michaelis-Menten rate parameters are inverses. So, for example, an MMF rate parameter estimate of 0.125 would indicate that the half of the growth from the initial value to the asymptote occurs at 8 years after the initial timepoint. When reporting model results in the following, we will transform the MMF parameter onto the Michaelis-Menten midpoint scale because, as mentioned previously, a meaningful midpoint has a useful function in our definition of potential. The other notable addition in the MMF model is the δ parameter in the denominator. This parameter controls the inflection point of the curve sublogistically (i.e., the inflection point of a logistic curve is the midpoint, so a sublogistic inflection point is bounded above by the midpoint but can occur at any point between the initial value and the midpoint) where the timepoint of the inflection point (as a proportion of the midpoint) is located at $\left[\frac{\delta-1}{\delta+1}\right]^{1/\delta}$for $\delta > 1$ and the outcome variable at the inflection point is at $\frac{\delta-1}{2\delta}$ (as a proportion of the upper asymptote) for $\delta > 1$. When $\delta < 1$, no inflection point is present and δ is defined to be bounded below by 0. Thus, if $\delta = 1$, then $t_{Inflection} = 0$ and the MMF model reduces to the Micahelis-Menten model. Conversely, as $\delta \to \infty$, the MMF approaches the logistic curve (Seber & Wild, 1989; demonstrating why the general form of the MMF curve is desirable because other models can be shown to be special cases).

The added value of fitting the MMF model is to determine whether the parameter estimates of a more general model are estimated to be at or near values that would reduce the more general model to the Michaelis-Menten model. If this were the case, we have added empirical support for our theoretical decision to employ the

Michaelis-Menten model. Similar to the empirical example in the main text, we fit the model in SAS Proc NLMIXED with four quadrature points using the same starting values (starting values were transformed for the rate parameter so that the scale was appropriate). We placed subject-specific random effects on the initial value, the upper asymptote, and the rate parameter but not on the inflection point parameter. The residual error structure was also set to be a heterogeneous diagonal because the variance appears to increase over time. Table D1 shows the Michaelis-Menten estimates from the main text as compared to the MMF estimates.

In Table D1, it can be seen that the estimates of α_0, α_A, and all random-effect variances are extremely close between the Michaelis-Menten and MMF models. The rate parameter is somewhat lower in the MMF model than in the Michaelis-Menten, with the average student being estimated to reach half growth at about the fall of seventh grade instead of the beginning of ninth grade. The inflection parameter is not exactly equal to 1, although it is quite close (the 95% confidence interval is [0.84, 1.40], so the estimate is not significantly different from 1.00). The BIC of the Michaelis-Menten model is also lower than that of the MMF model, suggesting that the Michaelis-Menten model is more parsimonious and that it is redundant to estimate α_D rather than constraining it to 1.0.

Table D1. Comparison of Michaelis-Menten and Morgan-Mercer-Flodin parameter estimates.

Second-order growth model			
Parameter name	Parameter	MM	MMF
Initial Value	α_0	22.82	25.08
Capacity	α_A	249.47	251.38
Rate	α_R	9.07	7.11
Inflection	α_D	—	1.12
Var (Initial Value)	g_{00}	45.12	49.068
Var (Capacity)	g_{UU}	2,498.02	2,497.59
Var (Rate)	g_{RR}	9.26	8.94
Corr (Initial Value, Capacity)	$r_{(g_{00}, g_{UU})}$	0.24	0.23
Corr (Initial Value, Rate)	$r_{(g_{00}, g_{RR})}$	-0.15	-0.19
Corr (Capacity, Rate)	$r_{(g_{UU}, g_{RR})}$	0.17	0.09
Model fit			
-2Log-Likelihood		183,326	183,323
No. of Parameters		31	32
BIC		183,563	183,568

Note. MM = Michaelis-Menten; MMF = Margan-Mercer-Flodin.