Multivariate Behavioral Research

Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/hmbr20

On Latent Growth Models for Composites and Their Constituents

Gregory R. Hancock a , Xiulin Mao a & Hemant Kher b

a University of Maryland
b University of Delaware

To cite this article: Gregory R. Hancock, Xiulin Mao & Hemant Kher (2013) On Latent Growth Models for Composites and Their Constituents, Multivariate Behavioral Research, 48:5, 619-638

To link to this article: http://dx.doi.org/10.1080/00273171.2013.815579

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the “Content”) contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.
On Latent Growth Models for Composites and Their Constituents

Gregory R. Hancock and Xiulin Mao

University of Maryland

Hemant Kher

University of Delaware

Over the last decade and a half, latent growth modeling has become an extremely popular and versatile technique for evaluating longitudinal change and its determinants. Most common among the models applied are those for a single measured variable over time. This model has been extended in a variety of ways, most relevant for the current work being the multidomain and the second-order latent growth models. Whereas the former allows for growth function characteristics to be modeled for multiple outcomes simultaneously, with the degree of growth characteristics’ relations assessed within the model (e.g., cross-domain slope factor correlations), the latter models growth in latent outcomes, each of which has effect indicators repeated over time. But what if one has an outcome that is believed to be formative relative to its indicator variables rather than latent? In this case, where the outcome is a composite of multiple constituents, modeling change over time is less straightforward. This article provides analytical and applied details for simultaneously modeling growth in composites and their constituent elements, including a real data example using a general computer self-efficacy questionnaire.

Latent growth modeling (LGM), or latent curve analysis, is a structural equation modeling approach for understanding individuals’ longitudinal growth. Models
may address change in a measured outcome variable, change in a latent outcome variable, concurrent change in multiple domains, linear or nonlinear change, latent growth mixtures, accelerated longitudinal designs, and so forth. Specifically, LGM techniques can describe individuals’ behavior in terms of reference levels (e.g., initial amount) and their developmental trajectories to and from those levels (e.g., linear, quadratic). In addition, they can determine the variability across individuals in both reference levels and trajectories as well as provide a means for testing the contribution of other variables or constructs to explaining those reference levels and growth trajectories, thus utilizing more information available in the measured variables than do traditional methods. For didactic treatments, see Bollen and Curran (2006); Duncan, Duncan, and Strycker (2006); Hancock, Harring, and Lawrence (2013); and Preacher, Wichman, MacCallum, and Briggs (2008).

If one has multiple measures at each time point, a multidomain latent growth model (e.g., Willett & Sayer, 1996) may be of interest, whereby growth function characteristics (e.g., intercept and slope in a linear function) for the multiple outcomes are modeled simultaneously, with the degree of their growth characteristics’ relations assessed within the model (e.g., correlations of within-domain slopes). If, more specifically, the multiple measures are believed to be indicators of the same construct (e.g., multiple self-concept questionnaire items), then the researcher is likely more interested in growth in the latent construct underlying those measures rather than in the individual measures themselves. In such cases one may analyze a second-order latent growth model (e.g., Hancock, Kuo, & Lawrence, 2001). A second-order latent growth model, also known as a curve-of-factors model (McArdle, 1988) or a latent variable longitudinal curve model (Tisak & Meredith, 1990), models growth factors as second-order factors influencing the first-order constructs whose longitudinal change is of primary interest.

In the second-order latent growth model described earlier, the outcome factor is latent relative to its indicators. That is, the factor is believed to have a structural bearing on its manifest variables, which serve as its effect indicators. What if, on the other hand, the indicators are believed to be causes of the construct of interest? Socioeconomic status (SES), for example, is best interpreted such that changes in SES are the result of changes in its constituent elements (e.g., family income, mother’s education) rather than vice versa (see, e.g., Bollen & Lennox, 1991). Similarly, changes in General Health might make the most sense as a function of changes in constituent aspects of health (e.g., exercise, HDL and LDL levels, blood pressure, and BMI). Many other examples exist as well; for several from the business and economics literature, see Diamantopoulos, Riefler, and Roth (2005, 2008). For these scenarios, which are sometimes referred to as emergent (rather than traditional latent) systems that contain formative (rather than traditional reflective) factors, the key distinction is the underlying theoretical
belief regarding the direction of causality relating the construct and the indicators. That is, this is not typically a matter for empirical resolution through model comparison; rather, a researcher must specify a priori a measurement model that is consistent with his or her understanding of the measurement mechanisms at work. So, for example, the construct of Exposure to Discrimination could be influenced by race, sex, age, and disabilities but certainly not vice versa; similarly, Life Stress is more likely the result of, say, job loss, divorce, bodily injury, and death of a loved one rather than their cause (Bollen & Lennox, 1991).

When one has an outcome that is believed to be formative relative to its indicator variables, modeling change over time is less straightforward. One option is to form the composite of interest, either using theoretical or practical weights (e.g., unit weighting) or using weights derived through some external optimization process (e.g., principal components), and then model growth in that preformed composite. Doing so would provide an assessment of the growth in the resulting composite, although without concurrent insights into growth in its constituent elements or their relations to growth in the composite. For example, longitudinal change in a composite may be positive but due solely to growth in one of its constituent pieces. Alternatively, change over time in a composite may seem nonexistent but turns out to appear so because some constituent elements show positive growth, others show negative growth, and their combination yields a near zero net growth. Ideally, then, one would like to have a model that preserves the growth information in the constituent elements, as does a multidomain latent growth model, while simultaneously modeling growth in the composite as a whole and relating its growth characteristics to those of its constituent elements.

Two seemingly useful, but in the end unviable, options are the following. One might try to impose formative factor models as discussed in, for example, Kline (2013) and then attempt to superimpose a growth structure atop those factors. Unfortunately, this strategy does not work because the model cannot sort the measured (cause indicator) inputs that go into the composites from the latent growth inputs that go into those same composites. An alternative might seem to be modeling the component variables and their cause indicators as part of the same model; however, as the composite is a weighted function of the constituents, having them in the same model suffers from fatal multicollinearity problems that render the model inestimable. This article offers one option that allows the simultaneous modeling of growth in composites and their constituent elements and the relations between the growth characteristics of both. This approach is developed later, starting with a simple illustrative case and then proceeding to cases that are sequentially more general in terms of number of constituents, number of time points, functional form of growth, and constituent weights. These developments are followed by a brief example using real data from a general computer self-efficacy questionnaire.
Simple Illustrative Case of a Linear Growth Model With Two Unit-Weighted Constituents Over Three Equally Spaced Time Points

To start, consider a scenario in which data are gathered at $T = 3$ ($t = 1, 2, 3$) equally spaced time points, where those data are on $C = 2$ ($c = 1, 2$) constituents ($y_{tc}$) that form a unit-weighted composite at each time point ($\Omega_t$). Thus, there are six measured variables ($y_{11}, y_{12}, y_{21}, y_{22}, y_{31}, y_{32}$) respectively forming three composites: $\Omega_1 = y_{11} + y_{12}$, $\Omega_2 = y_{21} + y_{22}$, and $\Omega_3 = y_{31} + y_{32}$. These composites may be modeled as phantom variables (i.e., latent variables with no observed indicators; e.g., Rindskopf, 1984), as displayed in Figure 1.

Further assume for this initial example that a linear functional form is of interest, with intercept ($\alpha$) and slope ($\beta$) factors for both the constituents and the composites; that is, for constituent 1 ($y_{t1}$) the growth factors are $\alpha_1$ and $\beta_1$, for constituent 2 ($y_{t2}$) they are $\alpha_2$ and $\beta_2$, and for the composite ($\Omega_t$) they are $\alpha_\Omega$ and $\beta_\Omega$. Given that the constituents and composites are from equally spaced time points in this example, the desired target loadings on their respective growth factors would follow the familiar pattern

$$
\begin{bmatrix}
1 & 0 \\
1 & 1 \\
1 & 2
\end{bmatrix}.
$$
The goal, then, is to create a comprehensive model where the constituents and composites have the desired relations with their respective growth factors.

Consider the covariance structure model presented in Figure 2 (displayed in three panels), in which the growth factors for the composite are a unit-weighted and error-free sum of the growth factors for its constituents: \( \alpha_{\Omega} = \alpha_1 + \alpha_2 \) and \( \beta_{\Omega} = \beta_1 + \beta_2 \). Writing out the structural equation for \( y_{11} \) from Panel 2a, followed by substitution for \( \alpha_c \) and \( \beta_c \), leads to the following:

\[
y_{11} = \frac{1}{2} \alpha_1 - \frac{1}{2} \alpha_2 + \frac{1}{2} \alpha_{\Omega} + 0 \beta_1 + 0 \beta_2 + 0 \beta_{\Omega} + \epsilon_{11} \\
= \frac{1}{2} \alpha_1 - \frac{1}{2} \alpha_2 + \frac{1}{2} (\alpha_1 + \alpha_2) + 0 \beta_1 + 0 \beta_2 + 0(\beta_1 + \beta_2) + \epsilon_{11} \\
= 1 \alpha_1 + 0 \beta_1 + \epsilon_{11}.
\]
Proceeding similarly for $y_{21}$ and $y_{31}$ from Panels 2b and 2c leads respectively to

$$y_{21} = (1/2)\alpha_1 - (1/2)\alpha_2 + (1/2)\alpha_\Omega + (1/2)\beta_1 - (1/2)\beta_2 + (1/2)\beta_\Omega + \varepsilon_{21}$$

$$= (1/2)\alpha_1 - (1/2)\alpha_2 + (1/2)(\alpha_1 + \alpha_2) + (1/2)\beta_1 - (1/2)\beta_2$$

$$+ (1/2)(\beta_1 + \beta_2) + \varepsilon_{21}$$

$$= 1\alpha_1 + 1\beta_1 + \varepsilon_{21}$$

$$y_{31} = (1/2)\alpha_1 - (1/2)\alpha_2 + (1/2)\alpha_\Omega + 1\beta_1 - 1\beta_2 + 1\beta_\Omega + \varepsilon_{31}$$

$$= (1/2)\alpha_1 - (1/2)\alpha_2 + (1/2)(\alpha_1 + \alpha_2) + 1\beta_1 - 1\beta_2 + 1(\beta_1 + \beta_2) + \varepsilon_{31}$$

$$= 1\alpha_1 + 2\beta_1 + \varepsilon_{31}.$$
As the reader may verify, equations for the second constituent simplify similarly to
\[
y_{12} = 1\alpha_2 + 0\beta_2 + \varepsilon_{12} \\
y_{22} = 1\alpha_2 + 1\beta_2 + \varepsilon_{22} \\
y_{32} = 1\alpha_2 + 2\beta_2 + \varepsilon_{32}.
\]
Finally, one may determine the total effects of the composite growth factors on the composites themselves either by path tracing from the former to the latter or by summing the constituents’ formulas. For example,
\[
\Omega_1 = y_{11} + y_{12} = (1\alpha_1 + 0\beta_1 + \varepsilon_{11}) + (1\alpha_2 + 0\beta_2 + \varepsilon_{12}) \\
= 1(\alpha_1 + \alpha_2) + 0(\beta_1 + \beta_2) + (\varepsilon_{11} + \varepsilon_{12}) \\
= 1\alpha_\Omega + 0\beta_\Omega + \zeta_1 \text{ where } \zeta_1 = \varepsilon_{11} + \varepsilon_{12}.
\]
It follows similarly that

\[ \Omega_2 = y_{21} + y_{22} = 1\alpha_\Omega + 1\beta_\Omega + \zeta_2 \text{ where } \zeta_2 = \varepsilon_{21} + \varepsilon_{22} \]

\[ \Omega_3 = y_{31} + y_{32} = 1\alpha_\Omega + 2\beta_\Omega + \zeta_3 \text{ where } \zeta_3 = \varepsilon_{31} + \varepsilon_{32}. \]

Thus, the model in Figure 2 constitutes a system in which the total effect of growth factors on their respective outcomes precisely follows the prescribed linear form

\[
\begin{bmatrix}
y_{11} \\
y_{12} \\
y_{21} \\
y_{22} \\
y_{31} \\
y_{32}
\end{bmatrix} = \begin{bmatrix}
(+\frac{1}{2}) & (-\frac{1}{2}) & (+\frac{1}{2}) & 0 & 0 & 0 \\
(-\frac{1}{2}) & (+\frac{1}{2}) & (+\frac{1}{2}) & 0 & 0 & 0 \\
(+\frac{1}{2}) & (-\frac{1}{2}) & (+\frac{1}{2}) & (+\frac{1}{2}) & (-\frac{1}{2}) & (+\frac{1}{2}) \\
(-\frac{1}{2}) & (+\frac{1}{2}) & (+\frac{1}{2}) & (+\frac{1}{2}) & (+\frac{1}{2}) & (+\frac{1}{2}) \\
(+\frac{1}{2}) & (-\frac{1}{2}) & (+\frac{1}{2}) & (+\frac{1}{2}) & (+\frac{1}{2}) & (+\frac{1}{2}) \\
(-\frac{1}{2}) & (+\frac{1}{2}) & (+\frac{1}{2}) & (+\frac{1}{2}) & (+\frac{1}{2}) & (+\frac{1}{2})
\end{bmatrix} \begin{bmatrix}
\alpha_1 \\
\alpha_2 \\
\alpha_\Omega \\
\beta_1 \\
\beta_2 \\
\beta_\Omega
\end{bmatrix} + \begin{bmatrix}
\varepsilon_{11} \\
\varepsilon_{12} \\
\varepsilon_{21} \\
\varepsilon_{22} \\
\varepsilon_{31} \\
\varepsilon_{32}
\end{bmatrix}
\]

where

\[
\begin{bmatrix}
\alpha_\Omega \\
\beta_\Omega
\end{bmatrix} = \begin{bmatrix}
1 & 1 & 0 & 0 \\
0 & 0 & 1 & 1
\end{bmatrix} \begin{bmatrix}
\alpha_1 \\
\alpha_2 \\
\beta_1 \\
\beta_2
\end{bmatrix}.
\]

As such, this model as depicted facilitates the assessment of the growth factor characteristics, including the variation and covariation of all growth factors. These are directly assessed for the constituent growth factors and implied for the composite growth factors, where the latter are often provided in standard SEM software output. Further, note that the composites themselves \((\Omega_1, \Omega_2, \Omega_3)\) need not actually be modeled. Given that the growth portion of the model is designed around their implied existence, and given that modeling them as the latent outcomes depicted in Figure 2 would expend no degrees of freedom, their inclusion in the model is completely optional (and does not, in fact, appear in the real data example later in the article). Finally, although not depicted in Figure 2, the imposition of a mean structure would facilitate the estimation and testing of latent growth mean parameters, with those for the constituent growth factors again being direct parameters in the model and those for the composite growth factors being implied by other parameters in the model. A mean structure is included in the real data example presented later in this article.
Linear Growth in Composites and Any Number of Unit-Weighted Constituents Across Any Number of Equally Spaced Time Points

Following from the aforementioned simple example, it is possible to generalize the growth model with two unit-weighted constituents over three equally spaced time points to any number of unit-weighted constituents across any number of equally spaced time points. Specifically, for a linear growth model with \( C \) \((c = 1, \ldots, C)\) constituents and \( T \) \((t = 1, \ldots, T)\) equally spaced time points, the model is of the following form:

\[
\begin{bmatrix}
(\frac{-c}{T}) & (-\frac{1}{T}) & \cdots & (-\frac{1}{T}) & (\frac{1}{T}) \\
(-\frac{1}{T}) & (\frac{-c}{T}) & \cdots & (-\frac{1}{T}) & (\frac{1}{T}) \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
(-\frac{1}{T}) & (-\frac{1}{T}) & \cdots & (\frac{-c}{T}) & (\frac{1}{T}) \\
(-\frac{1}{T}) & (-\frac{1}{T}) & \cdots & (-\frac{1}{T}) & (\frac{-c}{T}) \\
(\frac{-c}{T}) & (-\frac{1}{T}) & \cdots & (-\frac{1}{T}) & (\frac{1}{T}) \\
(-\frac{1}{T}) & (-\frac{1}{T}) & \cdots & (-\frac{1}{T}) & (\frac{1}{T}) \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
(-\frac{1}{T}) & (-\frac{1}{T}) & \cdots & (-\frac{1}{T}) & (\frac{1}{T}) \\
\end{bmatrix}
\begin{bmatrix}
y_{11} \\
y_{12} \\
\vdots \\
y_{1C} \\
y_{21} \\
y_{22} \\
\vdots \\
y_{2C} \\
y_{T1} \\
y_{T2} \\
\vdots \\
y_{TC} \\
\end{bmatrix}
= 
\begin{bmatrix}
y_0 \\
y_1 \\
\vdots \\
y_{C-1} \\
y_C \\
\vdots \\
y_{2C-1} \\
y_{2C} \\
\vdots \\
y_{TC-1} \\
y_{TC} \\
\end{bmatrix}
+ 
\begin{bmatrix}
\alpha_1 \\
\alpha_2 \\
\vdots \\
\alpha_C \\
\beta_1 \\
\beta_2 \\
\vdots \\
\beta_C \\
\end{bmatrix}
\begin{bmatrix}
1 & 1 & \cdots & 1 & 0 & 0 & \cdots & 0 \\
0 & 0 & \cdots & 1 & 1 & 0 & \cdots & 0 \\
\alpha_C & \beta_1 & \beta_2 & \cdots & \beta_C \\
\end{bmatrix}
\]

where

\[
\begin{bmatrix}
\alpha_O \\
\beta_O \\
\end{bmatrix}
= 
\begin{bmatrix}
1 & 1 & \cdots & 1 & 0 & 0 & \cdots & 0 \\
0 & 0 & \cdots & 1 & 1 & 0 & \cdots & 0 \\
\alpha_C & \beta_1 & \beta_2 & \cdots & \beta_C \\
\end{bmatrix}
\]

As seen in the loading matrix, regarding the intercept factors in this model the loading path from a constituent intercept factor \( \alpha_c \) to a corresponding constituent \( y_{tc} \) is \( \frac{c-1}{T} \), whereas the path to a constituent not corresponding to the intercept factor is \( -\frac{1}{T} \). Meanwhile, the loading path from the composite intercept factor \( \alpha_O \) to all constituents is \( \frac{1}{T} \). As for the slope factors in this model, the loading path from a constituent slope factor \( \beta_c \) to a corresponding constituent \( y_{tc} \) is
whereas the path to a constituent not corresponding to the slope factor is \(-\frac{T-1}{C}\). As for the composite slope factor \(\beta_\Omega\), its loading paths to all constituents is \(\frac{T-1}{C}\). Thus, as with typical linear growth modeling, the loadings associated with the slope factors are time dependent whereas those associated with the intercept factors are not.

For illustration, consider the case of data gathered at \(T = 3\) \((t = 1, 2, 3)\) equally spaced time points on \(C = 3\) \((c = 1, 2, 3)\) constituents \((y_{tc})\) that form a unit-weighted composite at each time point \((\Omega_t)\). Thus, there are nine measured variables \((y_{11}, y_{12}, y_{13}, y_{21}, y_{22}, y_{23}, y_{31}, y_{32}, y_{33})\) respectively forming three composites: \(\Omega_1 = y_{11} + y_{12} + y_{13},\ \Omega_2 = y_{21} + y_{22} + y_{23},\ \text{and}\ \Omega_3 = y_{31} + y_{32} + y_{33}\). Further assume for this example that a linear functional form is again of interest, with intercept \((\alpha)\) and slope \((\beta)\) factors for both the constituents and the composites; that is, for Constituent 1 the growth factors are \(\alpha_1\) and \(\beta_1\), for Constituent 2 they are \(\alpha_2\) and \(\beta_2\), for Constituent 3 they are \(\alpha_3\) and \(\beta_3\), and for the composite they are \(\alpha_\Omega\) and \(\beta_\Omega\).

Based on the aforementioned generic expressions, the loading path from a constituent intercept factor \(\alpha_c\) to a corresponding constituent \((y_{tc})\) is \(\frac{2}{3}\) and to a noncorresponding constituent is \(-\frac{1}{3}\); the loading from composite intercept factor \(\alpha_\Omega\) to all the constituents is \(\frac{1}{4}\). For the slopes, the loading path from a constituent slope factor \(\beta_c\) to a corresponding constituent \((y_{tc})\) is 0 at the first time point \((t = 1)\), \(\frac{2}{3}\) at the second time point \((t = 2)\), and \(\frac{4}{3}\) at the third time point \((t = 3)\). The loading path from a constituent slope factor \(\beta_c\) to a noncorresponding constituent is 0 at the first time point \((t = 1)\), \(-\frac{1}{3}\) at the second time point \((t = 2)\), and \(-\frac{2}{3}\) at the third time point \((t = 3)\). The loading paths from composite slope factor \(\beta_\Omega\) to all constituents are 0 at the first time point \((t = 1)\), \(\frac{1}{3}\) at the second time point \((t = 2)\), and \(\frac{2}{3}\) at the third time point \((t = 3)\). Thus, the growth model for this case may be represented as

\[
\begin{bmatrix}
    y_{11} \\
    y_{12} \\
    y_{13} \\
    y_{21} \\
    y_{22} \\
    y_{23} \\
    y_{31} \\
    y_{32} \\
    y_{33}
\end{bmatrix} =
\begin{bmatrix}
    \left(\frac{2}{3}\right) & \left(-\frac{1}{3}\right) & \left(-\frac{1}{3}\right) & \left(\frac{1}{3}\right) & 0 & 0 & 0 & 0 & 0 \\
    \left(-\frac{1}{3}\right) & \left(\frac{2}{3}\right) & \left(-\frac{1}{3}\right) & \left(\frac{1}{3}\right) & 0 & 0 & 0 & 0 & 0 \\
    \left(-\frac{1}{3}\right) & \left(-\frac{1}{3}\right) & \left(\frac{2}{3}\right) & \left(\frac{1}{3}\right) & 0 & 0 & 0 & 0 & 0 \\
    \left(\frac{2}{3}\right) & \left(-\frac{1}{3}\right) & \left(-\frac{1}{3}\right) & \left(\frac{2}{3}\right) & \left(-\frac{1}{3}\right) & \left(-\frac{1}{3}\right) & \left(\frac{2}{3}\right) & \left(-\frac{1}{3}\right) & \left(\frac{1}{3}\right) \\
    \left(-\frac{1}{3}\right) & \left(\frac{2}{3}\right) & \left(-\frac{1}{3}\right) & \left(\frac{2}{3}\right) & \left(-\frac{1}{3}\right) & \left(-\frac{1}{3}\right) & \left(\frac{2}{3}\right) & \left(-\frac{1}{3}\right) & \left(\frac{1}{3}\right) \\
    \left(-\frac{1}{3}\right) & \left(-\frac{1}{3}\right) & \left(\frac{2}{3}\right) & \left(\frac{1}{3}\right) & \left(-\frac{1}{3}\right) & \left(-\frac{1}{3}\right) & \left(\frac{2}{3}\right) & \left(-\frac{1}{3}\right) & \left(\frac{1}{3}\right) \\
    \left(\frac{2}{3}\right) & \left(-\frac{1}{3}\right) & \left(-\frac{1}{3}\right) & \left(\frac{2}{3}\right) & \left(-\frac{1}{3}\right) & \left(-\frac{1}{3}\right) & \left(\frac{2}{3}\right) & \left(-\frac{1}{3}\right) & \left(\frac{1}{3}\right) \\
    \left(-\frac{1}{3}\right) & \left(\frac{2}{3}\right) & \left(-\frac{1}{3}\right) & \left(\frac{2}{3}\right) & \left(-\frac{1}{3}\right) & \left(-\frac{1}{3}\right) & \left(\frac{2}{3}\right) & \left(-\frac{1}{3}\right) & \left(\frac{1}{3}\right)
\end{bmatrix}
\begin{bmatrix}
    \alpha_1 \\
    \alpha_2 \\
    \alpha_3 \\
    \alpha_\Omega \\
    \beta_1 \\
    \beta_2 \\
    \beta_3 \\
    \beta_\Omega
\end{bmatrix} +
\begin{bmatrix}
    \varepsilon_{11} \\
    \varepsilon_{12} \\
    \varepsilon_{13} \\
    \varepsilon_{21} \\
    \varepsilon_{22} \\
    \varepsilon_{23} \\
    \varepsilon_{31} \\
    \varepsilon_{32} \\
    \varepsilon_{33}
\end{bmatrix}
\]
where

\[
\begin{bmatrix}
\alpha_\Omega \\
\beta_\Omega
\end{bmatrix}
= \begin{bmatrix}
1 & 1 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 1 & 1
\end{bmatrix}
\begin{bmatrix}
\alpha_1 \\
\alpha_2 \\
\alpha_3 \\
\beta_1 \\
\beta_2 \\
\beta_3
\end{bmatrix}.
\]

Finally, as the reader may verify, this loading pattern results in the following implied relations for the composites:

\[
\begin{align*}
\Omega_1 &= y_{11} + y_{12} + y_{13} = 1\alpha_\Omega + 0\beta_\Omega + \zeta_1, \text{ where } \zeta_1 = \varepsilon_{11} + \varepsilon_{12} + \varepsilon_{13} \\
\Omega_2 &= y_{21} + y_{22} + y_{23} = 1\alpha_\Omega + 1\beta_\Omega + \zeta_2, \text{ where } \zeta_2 = \varepsilon_{21} + \varepsilon_{22} + \varepsilon_{23} \\
\Omega_3 &= y_{31} + y_{32} + y_{33} = 1\alpha_\Omega + 2\beta_\Omega + \zeta_3, \text{ where } \zeta_3 = \varepsilon_{31} + \varepsilon_{32} + \varepsilon_{33}.
\end{align*}
\]

MODELING GENERAL GROWTH IN COMPOSITES
AND CONSTITUENTS

In the previous section of the article we started with a simple but highly restrictive illustrative example and expanded it to accommodate linear growth in any number of unit-weighted constituents across any number of equally spaced time points. This section of the article continues the generalization of this approach, first to incorporate growth of any functional form (i.e., not just linear) for unit-weighted composites and their constituents, and then ultimately to allow for any weighting scheme (i.e., not just unit weighted).

General Growth in Composites and Unit-Weighted Constituents

The linear growth model with any number of unit-weighted constituents across equally spaced time points can be expanded to a more general growth model with any possible loading pattern by embedding the desired functional form’s target loadings directly. Consider a growth model across \(T\) time points with \(C\) unit-weighted constituents \((y_{tc})\), where the growth form is operationalized using \(K\) factors for each \(c\)th constituent \((\eta_{c1}, \ldots, \eta_{cK})\) as well as for the composite \((\eta_{\Omega1}, \ldots, \eta_{\Omega K})\). The ultimate aim then, as before, is to create a comprehensive model where the actual loadings are strategically chosen such that the growth factors’ total effects on the constituents and composites constitute the target...
loadings of the desired functional form, where those target loadings across the $T$ (not necessarily equally spaced) time points may be represented as

$$\begin{bmatrix}
1 & \lambda_{12} & \cdots & \lambda_{1K} \\
\vdots & \vdots & \ddots & \vdots \\
1 & \lambda_{T2} & \cdots & \lambda_{TK}
\end{bmatrix}.$$ 

In order to achieve this, the loading path from a constituent growth factor to a corresponding constituent ($y_{tc}$) should be set as $\frac{(C-1)\lambda_{tk}}{C}$ and to a noncorresponding constituent as $-\frac{\lambda_{tk}}{C}$; the loading path from the composite growth factor to all constituents is $\frac{\lambda_{tk}}{C}$.

For illustration, consider the case of $C = 3$ ($c = 1, 2, 3$) constituents measured at $T = 4$ ($t = 1, 2, 3, 4$) equally spaced time points, where both linear and quadratic trends are of interest as expressed by the target loadings of

$$\begin{bmatrix}
1 & 0 & 0 \\
1 & 1 & 1 \\
1 & 2 & 4 \\
1 & 3 & 9
\end{bmatrix}.$$ 

In its entirety the model has 12 growth factors, 9 for constituents and 3 for the composites. Based on the aforementioned relations, the growth model for this case may be represented as follows:
where

\[
\begin{bmatrix}
\eta_{\Omega 1} \\
\eta_{\Omega 2} \\
\eta_{\Omega 3}
\end{bmatrix} = \begin{bmatrix}
1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1
\end{bmatrix} \begin{bmatrix}
\eta_{11} \\
\eta_{12} \\
\eta_{13} \\
\eta_{21} \\
\eta_{22} \\
\eta_{23} \\
\eta_{31} \\
\eta_{32} \\
\eta_{33}
\end{bmatrix}.
\]

General Growth in Composites and Weighted Constituents

Finally, the generic form established earlier for unit-weighted composites can be generalized to constituents with any system of weights. Consider a growth model across \( T \) time points with \( C \) constituents \( y_{tc} \) with corresponding weights \( w_1, \ldots, w_C \) used to form each \( t \)th composite; that is, \( \Omega_t = \sum_{c=1}^{C} w_c y_{tc} \). The growth form is again operationalized using \( K \) factors for each \( c \)th constituent \( (\eta_{c 1}, \ldots, \eta_{c K}) \) as well as for the composite \( (\eta_{\Omega 1}, \ldots, \eta_{\Omega K}) \). In order to create a comprehensive model where the growth factors’ total effects on the constituents and composites constitute the target loadings of the desired functional form, loadings must be assigned as follows. Specifically, defining \( W = \sum_{c=1}^{C} w_c \), the loading path from a constituent growth factor to a corresponding constituent \( y_{tc} \) should be set as \( \frac{W-w_c}{W} \lambda_{ck} \) and to a noncorresponding constituent as \( -\frac{w_c}{W} \lambda_{ck} \); the loading path from the composite growth factor to all constituents is \( \lambda_{\Omega k} \).

For illustration, consider the case of data gathered at \( T = 3 \) \((t = 1, 2, 3)\) equally spaced time points on \( C = 3 \) \((c = 1, 2, 3)\) constituents \( y_{tc} \) that form a weighted composite with weights \( w_c \) for each \( c \)th constituent of \( w_1 = 1, w_2 = 2, \) and \( w_3 = 3 \) at each time point (i.e., \( \Omega_t = w_1 y_{t1} + w_2 y_{t2} + w_3 y_{t3} \)). Thus, there are nine measured variables \((y_{11}, y_{12}, y_{13}, y_{21}, y_{22}, y_{23}, y_{31}, y_{32}, y_{33})\) respectively forming three composites: \( \Omega_1 = 1y_{11} + 2y_{12} + 3y_{13}, \Omega_2 = 1y_{21} + 2y_{22} + 3y_{23}, \) and \( \Omega_3 = 1y_{31} + 2y_{32} + 3y_{33} \). Further assume for this example that a linear functional form is again of interest, with intercept \((\alpha)\) and slope \((\beta)\) factors for both the constituents and the composites; that is, for Constituent 1 the growth factors are \( \alpha_1 \) and \( \beta_1 \), for Constituent 2 they are \( \alpha_2 \) and \( \beta_2 \), for Constituent 3 they are \( \alpha_3 \) and \( \beta_3 \), and for the composite they are \( \alpha_\Omega \) and \( \beta_\Omega \). Because of the different weights imposed, the relations between factors for the constituents and the composites are \( \alpha_\Omega = 1\alpha_1 + 2\alpha_2 + 3\alpha_3 \) and \( \beta_\Omega = 1\beta_1 + 2\beta_2 + 3\beta_3 \). Thus, the growth model for this case may be represented as follows:
Data for this example come from a general computer self-efficacy (GCSE) scale administered to undergraduate students in an introductory management information systems course within the business school of a university in the northeastern United States. In this course students learn about computers in general and learn to use office productivity software programs such Microsoft Excel (spreadsheets) and Microsoft Access (database). Repeated measurements on the GCSE instrument for this example were taken in the 5th, 9th, and 13th weeks of a 13-week long semester, making the time between successive measurements 4 weeks. The items used on the GCSE scale were taken from Marakas, Johnson, and Clay (2007), with each addressing whether or not the student possessed a specific skill (yes/no). If the student possessed the skill, he or she was then asked to rate the level of confidence in utilizing that skill on a percentage scale from 10% to 100% (in 10% increments). The student’s score for a given item in a given time period was recorded as a percentage from 0% (did not possess the skill) to 100% (completely confident in the skill).

For this illustration, three items were selected from the GCSE scale: Item 2 (I believe I have the ability to install new software applications), Item 5 (I believe I have the ability to remove information that I no longer need from a computer), and Item 6 (I believe I have the ability to use a computer to display or present information in a desired manner). A composite of these three items was also of interest for this example, with a weight of \( w_1 = 3 \) for Item 2, \( w_2 = 2 \) for Item 5, and \( w_3 = 1 \) for Item 6. In sum, complete data on \( n = 93 \) cases were modeled, where the data were from \( T = 3 \) equally spaced time points on \( C = 3 \) constituents that form a weighted composite at each time point.
The resulting growth model, with respective constituent weights of \( w_1 = 3, w_2 = 2, \) and \( w_3 = 1 \) as mentioned earlier, can be determined to be as follows:

\[
\begin{bmatrix}
  y_{21} \\
  y_{22} \\
  y_{23} \\
  y_{31} \\
  y_{32} \\
  y_{33}
\end{bmatrix} = \begin{bmatrix}
  \left(\frac{3}{6}\right) & -\left(\frac{2}{6}\right) & -\left(\frac{1}{6}\right) \\
  \left(\frac{3}{6}\right) & -\left(\frac{2}{6}\right) & -\left(\frac{1}{6}\right) \\
  \left(\frac{3}{6}\right) & -\left(\frac{2}{6}\right) & -\left(\frac{1}{6}\right) \\
  \left(\frac{3}{6}\right) & -\left(\frac{2}{6}\right) & -\left(\frac{1}{6}\right) \\
  \left(\frac{3}{6}\right) & -\left(\frac{2}{6}\right) & -\left(\frac{1}{6}\right) \\
  \left(\frac{3}{6}\right) & -\left(\frac{2}{6}\right) & -\left(\frac{1}{6}\right)
\end{bmatrix} \begin{bmatrix}
  \alpha_1 \\
  \alpha_2 \\
  \alpha_3 \\
  \alpha_\Omega \\
  \beta_1 \\
  \beta_2 \\
  \beta_3 \\
  \beta_\Omega
\end{bmatrix} + \begin{bmatrix}
  \varepsilon_{11} \\
  \varepsilon_{12} \\
  \varepsilon_{13} \\
  \varepsilon_{21} \\
  \varepsilon_{22} \\
  \varepsilon_{23} \\
  \varepsilon_{31} \\
  \varepsilon_{32} \\
  \varepsilon_{33}
\end{bmatrix}.
\]

(i.e., \( \Omega_t = 3y_{t1} + 2y_{t2} + y_{t3} \)). Correlations, means, and standard deviations for the nine observed variables appear in Table 1.

For a linear model with the last time point (the end of the course) as the reference level, the target loadings would be

\[
\begin{bmatrix}
  1 & -2 \\
  1 & -1 \\
  1 & 0
\end{bmatrix}.
\]
with additional structural relations of

\[
\begin{bmatrix}
\alpha_\Omega \\
\beta_\Omega
\end{bmatrix} =
\begin{bmatrix}
3 & 2 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 3 & 2 & 1
\end{bmatrix}
\begin{bmatrix}
\alpha_1 \\
\alpha_2 \\
\alpha_3 \\
\beta_1 \\
\beta_2 \\
\beta_3
\end{bmatrix}.
\]

For the covariance structure the parameters to be estimated are thus the variances and covariances of the growth factors \(\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3\); the variances of the errors \(\varepsilon_{11}, \varepsilon_{12}, \varepsilon_{13}, \varepsilon_{21}, \varepsilon_{22}, \varepsilon_{23}, \varepsilon_{31}, \varepsilon_{32}, \varepsilon_{33}\); and the covariances among the errors of the constituents at each time point, \((\varepsilon_{11}, \varepsilon_{12}, \varepsilon_{13})\), \((\varepsilon_{21}, \varepsilon_{22}, \varepsilon_{23})\), and \((\varepsilon_{31}, \varepsilon_{32}, \varepsilon_{33})\). As for the mean structure, only the means of the growth factors are estimated: \(\alpha_1, \alpha_2, \alpha_3\). The model’s nine variables have 45 (co)variances and nine means with which to estimate the 39 parameters of the covariance structure and the 6 parameters of the mean structure, leaving nine degrees of freedom overall.

Data analysis was carried out using maximum likelihood estimation with EQS 6.2 (Bentler & Wu, 2012), with syntax appearing in the Appendix. The linear growth model for three time points and three weighted constituents fit very well: \(\chi^2(9) = 4.794\), SRMR = 0.013, RMSEA = 0.000, CFI = 1.000. Most interesting for this example are the correlations, means, and standard deviations for the composite growth factors, which are presented in Table 2.

To start, note that whereas the constituents were on a 0–100 percentage scale, the composites were on a 0–600 scale due to the weighting; thus, if one prefers

<table>
<thead>
<tr>
<th>TABLE 2</th>
<th>Descriptive Statistics for Constituent and Composite Growth Factors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(\alpha_1)</td>
</tr>
<tr>
<td>(\alpha_1)</td>
<td>1.000</td>
</tr>
<tr>
<td>(\alpha_2)</td>
<td>.752*</td>
</tr>
<tr>
<td>(\alpha_3)</td>
<td>.842*</td>
</tr>
<tr>
<td>(\alpha_c)</td>
<td>.944(^b)</td>
</tr>
<tr>
<td>(\beta_1)</td>
<td>.255</td>
</tr>
<tr>
<td>(\beta_2)</td>
<td>.004</td>
</tr>
<tr>
<td>(\beta_3)</td>
<td>.039</td>
</tr>
<tr>
<td>(\beta_c)</td>
<td>.132(^b)</td>
</tr>
<tr>
<td>(M)</td>
<td>89.820*</td>
</tr>
<tr>
<td>(SD)</td>
<td>12.906*</td>
</tr>
</tbody>
</table>

\(^a\)Tested using the total effect test results. \(^b\)No test available in EQS.

*\(p < .05\).
the original scale, adjustments may be made using the sum of the weights, \( W = 6 \). So, the mean for \( \alpha_c \) was estimated at 535.562 and for \( \beta_c \) at 18.828, which translates to 89.260 and 3.138 on a standard percentage scale, respectively. This means that students finished the semester with an average composite score equivalent to an 89.260%, with average linear growth every 4 weeks estimated at 3.138% (which is statistically significantly greater than 0). Further, the standard deviation for \( \alpha_c \) was estimated at 74.698 and for \( \beta_c \) at 20.624, which respectively become 12.450 and 3.437 on a standard percentage scale. As for the correlation between students’ Week 13 composite performance and rate of linear growth in the composite, that was a fairly weak .226.

Considering the growth factor correlations more broadly, the relations among the constituent \( \alpha \) factors were all high, positive, and statistically significant, whereas none of the relations among the constituent \( \beta \) factors were statistically significant. Most interestingly, although no significance tests existed, the relations of the constituent \( \alpha \) factors to the composite factor \( \alpha_c \) were all in the .9s, whereas the relations of the constituent \( \beta \) factors to the composite factor \( \beta_c \) showed only the slope factor for GCSE Item 5 (\( \beta_2 \)) to be highly related to the composite slope factor at .913. The slope factors for GCSE Items 2 and 6 showed much weaker relations with the composite slope factor, .529 and .566, respectively. Thus, growth in the composite of interest appears to have been associated primarily with growth Item 5, the question regarding the ability to remove information no longer needed from a computer.

**SUMMARY AND CONCLUSIONS**

This article has offered a strategy for simultaneously modeling growth in composites as well as their constituents without actually requiring that the composites themselves be included in the model. The strategy presented is sufficiently general as to allow for any number of constituents, any number of time points, growth of any linearizable functional form, and any weighting scheme for forming the composites from the constituents. The model was illustrated with an example that included a mean structure, which revealed that growth in the composite was primarily a reflection of growth in one particular constituent. One could imagine many extensions of the models presented herein, including unspecified latent trajectory models with only two loadings fixed (and then the precipitating constraints relating constituents’ and composites’ free loadings), or higher order growth models evaluating change in latent outcomes, or the inclusion of measured or latent covariates. The latter might be especially interesting as it would allow a researcher to understand why a covariate might or might not be related to a composite vis-à-vis its relations to the individual constituents.
REFERENCES


APPENDIX

EQS Syntax for Real Data Example

/TITLE
Real data example
/SPECIFICATIONS
DATAFILE='GCSE.ess'; MA=raw; AN=moment;
/LABELS
V1=T1ITEM2; V2=T1ITEM5; V3=T1ITEM6;
V4=T2ITEM2; V5=T2ITEM5; V6=T2ITEM6;
V7=T3ITEM2; V8=T3ITEM5; V9=T3ITEM6;
F1=ALPHA1; F2=ALPHA2; F3=ALPHA3; F4=ALPHAc;
F5=BETA1; F6=BETA2; F7=BETA3; F8=BETAc;
/EQUATIONS
V1 = 0.5F1 - 0.3333F2 - 0.1667F3 + 0.1667F4
    - 1F5 + 0.6667F6 + 0.3333F7 - 0.3333F8 + E1;
V2 = -0.5F1 + 0.6667F2 - 0.1667F3 + 0.1667F4
    + 1F5 - 1.3333F6 + 0.3333F7 - 0.3333F8 + E2;
V3 = -0.5F1 - 0.3333F2 + 0.8333F3 + 0.1667F4
    + 1F5 + 0.6667F6 - 1.6667F7 - 0.3333F8 + E3;
V4 = 0.5F1 - 0.3333F2 - 0.1667F3 + 0.1667F4
    - 0.5F5 + 0.3333F6 + 0.1667F7 - 0.1667F8 + E4;
V5 = -0.5F1 + 0.6667F2 - 0.1667F3 + 0.1667F4
    + 0.5F5 - 0.6667F6 + 0.1667F7 - 0.1667F8 + E5;
V6 = -0.5F1 - 0.3333F2 + 0.8333F3 + 0.1667F4
    + 0.5F5 + 0.3333F6 - 0.8333F7 - 0.1667F8 + E6;
V7 = 0.5F1 - 0.3333F2 - 0.1667F3 + 0.1667F4 + E7;
V8 = -0.5F1 + 0.6667F2 - 0.1667F3 + 0.1667F4 + E8;
V9 = -0.5F1 - 0.3333F2 + 0.8333F3 + 0.1667F4 + E9;
F1 = *V999 + D1;
F2 = *V999 + D2;
F3 = *V999 + D3;
F4 = 3F1 + 2F2 + 1F3 + D4;
F5 = *V999 + D5;
F6 = *V999 + D6;
F7 = *V999 + D7;
F8 = 3F5 + 2F6 + 1F7 + D8;
/VARIANCES
E1 to E9 = *;
D1 to D3 = *;
D4 = 0;
D5 to D7 = *;
D8 = 0;
/COVARIANCES
E1 to E3 = *;
E4 to E6 = *;
E7 to E9 = *;
D1 to D3 = *;
D5 to D7 = *;
D1,D5 = *;
D1,D6 = *;
D1,D7 = *;
D2,D5 = *;
D2,D6 = *;
D2,D7 = *;
D3,D5 = *;
D3,D6 = *;
D3,D7 = *;
/TECHNICAL
  ITR=1000;
/PRINT
  FIT=ALL; COVARIANCE=YES; CORRELATIONS=YES; EFFECTS=YES;
/END