Published in final edited form as:

Struct Equ Modeling. 2019 ; 26(3): 418–429. doi:10.1080/10705511.2018.1547110.

Time-Varying Effect Sizes for Quadratic Growth Models in Multilevel and Latent Growth Modeling

Alan Feingold

Oregon Social Learning Center, Eugene, Oregon

Abstract

Multilevel and latent growth modeling analysis (GMA) is often used to compare independent groups in linear random slopes of outcomes over time, particularly in randomized controlled trials. The unstandardized coefficient for the effect of group on the slope from a linear GMA can be transformed into a model-estimated effect size for the group difference at the end of a study. Because effect sizes vary non-linearly in quadratic GMA, the effect size at the end of a study using quadratic GMA cannot be derived from a single coefficient, and cannot be used to estimate effect sizes at intermediate time points with backwards extrapolation. This article formulates equations and associated input commands in Mplus for time-varying effect sizes for quadratic GMA. Illustrative analyses that produced these time-varying effect sizes were presented, and Monte Carlo study found that bias in the effect sizes and their confidence intervals was ignorable.

Keywords

effect sizes; confidence intervals; multilevel analysis; latent growth models

The study of change is fundamental in longitudinal research in the behavioral, educational, social, and health sciences. The advent and proliferation of growth modeling analysis (GMA)--including multilevel modeling/hierarchical linear models (Goldstein, 2011; Hedeker & Gibbons, 2006; Hox, Moerbeek, & van de Schoot, 2010; Raudenbush & Bryk, 2002) and latent growth modeling/structural equation modeling (Bollen & Curran, 2006; Grimm, Ram, & Estabrook, 2016; McArdle & Nesselroade, 2014; Preacher, Wichman, MacCallum, & Briggs, 2008)--has had a profound impact on the study of change, particularly with repeated measures data collected from randomized controlled trials (RCTs). Applications of GMA have been increasing exponentially in clinical psychology, psychiatry, and prevention science (Feingold, 2009; Gueorguieva & Krystal, 2004). Indeed, Aiken, West, and Millsap (2008) deemed GMA one of the most important methodological innovations of the twenty-first century, and Kuljanin, Braun, and DeShon (2011) averred that these methods "are currently the dominant approaches to the analysis of longitudinal data in psychology" (p. 249).

However, little attention has been paid to methods for communicating the practical significance of effects for group differences in non-linear GMA (i.e., where outcome trajectories are not linear for, say, both the treatment and control groups). This article introduces a new conceptual model and associated equations for effect sizes for findings from quadratic GMA with categorical time-invariant covariates (e.g., group, with control = 0 and treatment = 1), which is frequently used to examine data collected from RCTs.

Point Estimates of Standardized Effect Sizes for Linear GMA

Until relatively recently, effect sizes for group differences (e.g., treatment effects, sex differences) in outcome trajectories examined with GMA by comparing the means of latent random slopes were often unreported. Moreover, equations for effects sizes that were used in different GMA studies were not mathematically equivalent, and the resulting effect sizes were not in the same metric as effect sizes from classical analysis (Feingold, 2009). Contemporary work on development of effect sizes for GMA has proposed the conversion of the coefficient for the difference between two groups in their linear slopes into a standardized mean difference (GMA d) at the end of a randomized study (Feingold, 2009, 2013, 2015, 2017),

$$\text{GMA } d = (b * \text{duration}) / SD. \quad (1)$$

The b in the numerator of Equation 1 is the coefficient for the effect of group on linear slope that is the difference in the rate of change in the outcome between the two groups per unit of time (e.g., per week when time is coded in weeks). Duration is the length of the study based on units associated with b (e.g., number of weeks from baseline if b is the group difference in rate of change per week). The numerator in Equation 1--the product of b and duration--is the model-estimated raw score mean difference between the two independent groups at the end of a randomized study, and is thus the equivalent of the observed raw score mean difference between two independent groups in a completely randomized design. SD is the pooled-within group SD of the outcome (y), which can be obtained from baseline data, observed data from other time points, or estimated from the GMA model (see discussions of SD estimation in Feingold, 2013, 2015, as Equation 1 assumes SD is estimated appropriately based on theoretical and statistical considerations). The GMA d obtained with Equation 1 is thus the model-estimated standardized mean difference (a Cohen's d equivalent) between groups at the end of a study when the expected value of the mean difference between the groups at baseline ($T1$) is zero.

GMA d is typically calculated with a two-step approach. At step 1, a GMA is conducted to obtain the coefficient (b) for effect of group on slope (in primary research), or a previously reported b is extracted from a research document (in meta-analytic applications; see Feingold, 2017). At step 2, Equation 1 is used to transform b into GMA d . (A similar two-step approach is generally used to calculate d from descriptive statistics in classical analysis, as common statistical programs that compare observed means do not report classical d either.)

An alternative method for calculating GMA d uses software that has the capability to create new parameters from existing ones (e.g., with the MODEL CONSTRAINT command in Mplus; Muthén & Muthén, 2017). Equation 1 is then incorporated into the input used to conduct the GMA and the results in the output file include both b and the b -derived GMA d for the effect of group on slope (e.g., the treatment effect). These two approaches yield estimates of the same effect size parameter (GMA delta).

Although GMA d has now been reported in hundreds of studies, particularly evaluations of interventions (e.g., Chorpita et al., 2017; Felder et al., 2017; Goodnight et al., 2017; Parra-Cardona, et al., 2017; Stice, Rohde, Shaw, & Gau, 2017), its use has been restricted to linear models. By convention, linear models are fit when GMAs are conducted with data collected at three time points (e.g., baseline, end of RCT, and a follow-up). GMA d is typically reported in studies that collected data at three time points, as well as in studies with more than three time points but where the quadratic effects were either ignored or found to be non-significant and then deleted from the model. However, effect sizes for findings from quadratic GMAs are needed because studies often use more than three time points and trajectories are not always linear (e.g., Badour, Flanagan, Gros, Killeen, Pericot-Valverde, Korte, Allan, & Back, 2017; Marcoulides & Khojasteh, 2018; Feingold, Washburn, Tiberio, & Capaldi, 2015; Muthén, & Muthén, 2000).

Point Estimates of Standardized Effect Sizes for Quadratic GMA

In a linear GMA, Equation 1 can be used to estimate the GMA d at any given time point (a *time-specific* GMA d) because the effect sizes vary linearly over the course of the study (Feingold, 2013). In other words, the value for duration in Equation 1 could be the period from baseline to a given time point rather than the conventional definition of duration as the span from the baseline time point to the final time point (length of study), and the set of GMA d s would be the model-estimated standardized mean differences between groups at different time points that fall on a line when plotted over time.

If, for example, data were collected at 5 equidistant assessments and the conventional GMA d from a linear model was .80, the GMA d s would be .20, .40, .60, and .80 at T2, T3, T4, and T5, respectively. Thus, there would be no need to report *time-varying* GMA d s because the GMA d at the end of the study could be used to determine the GMA d s at intermediate time points by backwards linear extrapolation.

However, in a non-linear GMA, one group may change at a greater rate than the other group earlier than later in the study. For example, if there was a large immediate but dwindling effect of an intervention, the effect sizes might be 1.20, .90, .82, and .80 at T2, T3, T4, and T5, respectively. Alternatively, the intervention may not “kick in” immediately and the respective effect sizes might be .10, .20, .40, and .80. The GMA d at the end of the study is the same in these two examples but cannot be derived from a single coefficient (i.e., b in Equation 1), and does not communicate the non-linear changes in effects over the course of a study that differ among studies using quadratic GMA. Although the GMA d at the end of the study is informative for both linear and quadratic models, only research using the latter

should report time-varying GMA d s to convey the temporal changes in group differences that occur during the study.

Equations for time-varying GMA effect sizes.

Time-varying GMA d s from quadratic models can be calculated by expanding Equation 1 to include a term for the coefficient for the effect of group on the quadratic slope and using varying values of duration in Equation 2,

$$\text{GMA } d = (b_s * \text{duration} + b_q * \text{duration}^2) / SD, \quad (2)$$

where b_s is the coefficient for the effect of group on the linear slope, b_q is the coefficient for the effect on the quadratic slope, and SD is the same estimate of within-group outcome dispersion used in linear GMA. In Equation 1, duration generally refers to the span from baseline to the end of the study (final time point). In Equation 2, duration is the interval from baseline to a particular time point, and this equation is used with different values of duration to calculate the set of time-varying GMA d s for the study. The number of time-varying GMA d s in a study equals the number of time points after baseline (based on frequency of assessments or time points of interest in the study).

For example, in a quadratic GMA with 4 equidistant time points (and time coded 0, 1, 2, and 3), the time-varying GMA d s for the study would be calculated as,

$$d_{T4} = (b_s * \text{duration-at-T4} + b_q * \text{duration-at-T4}^2) / SD = (b_s * 3 + b_q * 9) / SD \quad (3)$$

and

$$d_{T3} = (b_s * \text{duration-at-T3} + b_q * \text{duration-at-T3}^2) / SD = (b_s * 2 + b_q * 4) / SD$$

and

$$d_{T2} = (b_s * \text{duration-at-T2} + b_q * \text{duration-at-T2}^2) / SD = (b_s * 1 + b_q * 1) / SD.$$

where b_s is the unstandardized coefficient for the effect on linear slope, b_q is the unstandardized coefficient for the effect on the quadratic slope, d_{T4} is the time-specific GMA d at time point 4 (T4, end of study), d_{T3} is GMA d at T3, and d_{T2} is GMA d at T2. (Given randomization, the expected value of d_{T1} is zero).

When Equation 2 is applied to data where outcomes from both groups have linear trajectories, the expected value of b_q is zero and Equation 2 reduces to Equation 1. When the coefficient for the quadratic effect is zero, the time-varying GMA d s are the same as the d s obtained with Equation 1 (and varying values of duration)--and these time-varying GMA d s would then also fall on a line if plotted over time.

Time-varying GMA d s for a quadratic GMA can be obtained with a 2-step approach that uses the coefficients (b_s and b_q) for the effects of group on slopes and different values of duration. At step 1, a quadratic GMA is conducted to determine b_s and b_q . These coefficients are then included in Equation 2 at step 2 to produce the set of time-specific GMA d s for the study.

As with effect sizes from linear models, software can be used to directly obtain time-varying GMA d s. A key purpose of this article is to illustrate the application of the Mplus program to derive time-varying GMA d s and their confidence intervals (CIs) from the unstandardized coefficients for effects of group on linear and quadratic slopes. However, this approach is applicable to any software for GMA that can incorporate user-prescribed parameter functions, including lavaan in R (Rosseel, 2012), LISREL (Jöreskog & Sörbom, 2006), and PROC CALIS in SAS (SAS Institute Inc., 2011).

Standard Errors and Confidence Intervals for Standardized Effect Sizes for Quadratic GMA

When an example of the post hoc equation approach--which uses Equation 1 to compute GMA d following a linear GMA to obtain the b included in the equation--was first explicated, the estimation of the standard error (SE) and the CI for GMA d was not addressed (Feingold, 2009, 2013). However, post hoc equations for the SE and CI for GMA d were later formulated and validated (Feingold, 2015).

In the alternative method, where GMA d is directly calculated by software and included in output files, equations for the SE and CI for GMA d need not be specified in the input. In Mplus, for example, the SE for new parameters (e.g., GMA d) is automatically produced by the same delta method (Benichou & Gail, 1989; Kendall & Stuart, 1977) used by the program to obtain respective SE s for standard parameters (e.g., b). Thus, the input used to obtain GMA d also produces its SE by default. Optional Mplus input requesting unspecified CIs would produce CIs for both b and GMA d with the delta method. Moreover, CIs for both b and GMA d can also be obtained in Mplus with either the standard (percentile) bootstrap (Efron & Tibshirani, 1993) or the Bollen-Stine (1992) residual bootstrap in lieu of CIs calculated with the delta method. This article illustrates the application of each of three different methods that can be used by Mplus to estimate SE s and CIs for time-varying effect sizes from quadratic GMA.

Illustrative Analysis: Time-Varying GMA d s from Quadratic Models

Example 6.10 in the Mplus user's guide (Muthén & Muthén, 2017) is an input statement for a linear GMA--with 4 equidistant time points (coded 0, 1, 2, and 3 for $y11$, $y12$, $y13$, and $y14$, at T1, T2, T3, and T4, both respectively), 2 continuous time-invariant covariates ($x1$ and $x2$), and a single time-varying covariate ($a31$ - $a34$)--used for illustrative analysis of an accompanying dataset included with the Mplus software (ex6.10.dat, $N = 500$). The current illustrative analysis used this example as a foundation for demonstrating the calculation of time-varying GMA d s--and their SE s and CIs--from a quadratic GMA conducted with Mplus.

However, because GMA d is not applicable with continuous independent variables, a DEFINE command was added to Example 6.10 to dichotomize $x1$ (based on a mean split), and the requisite quadratic term was added to adapt the model for use as a quadratic GMA example. The MODEL CONSTRAINT command was added to generate the three time-specific GMA d s for T2, T3, and T4. Given 4 equidistant time points with codes differing by 1 unit between assessments, the values for duration at the post-baseline time points were 1, 2, and 3 at T2, T3, and T4, respectively (see Equation 3).

Multiple covariates GMA model.

The two-step method needed to obtain time-varying GMA d s from a model with multiple covariates (e.g., Mplus Example 6.10 and its initial modification) requires specifications of (a) duration from baseline to each time point, and (b) a predetermined fixed SD used to calculate each d . The SD of 1.478 in this example was the within-group SD at baseline ($y11_{\text{POOLED}}$), which is commonly used for SD estimation in RCTs because it ensures the dispersion estimate is unbiased by effect of treatment or attrition--and is also used to examine the effectiveness of the randomization when comparing groups at baseline (see Feingold, 2009, 2013). The SD was obtained by regressing $y11$ on $x1$ and taking the square root of the residual variance from a regression analysis. (This two-step approach is similar to the method used in Mplus to obtain effect sizes from mediation analysis with a continuous predictor and a binary outcome that also requires a predetermined SD to be specified in the input; Muthén, Muthén, & Asparouhov, 2016.) Values for the coefficients for the group differences in linear and quadratic slopes were not specified in the input because they were parameters in the model and estimated from the GMA.

Three input statements were used to conduct the analyses that produced the time-varying GMA d s, SE s, and CI s from the quadratic GMA as a function of SE and CI estimation method: (a) the default delta method (see Appendix A), (b) the standard (percentile) bootstrap, and (c) the Bollen-Stine (residual) bootstrap. Specification for the use of a bootstrap instead of the default delta method for estimating SE s and CI s was the same in this example as in all Mplus input statements. (The bootstrap used 500 draws in the illustrative analysis, the standard number of draws when the bootstrap was requested in input statements in Mplus user's guide examples.)

The observed time-varying GMA d s of .335, .624, and .865 for T2, T3, and T4, respectively, were the same in all three analyses because point estimates are not affected by the method used to estimate SE s and CI s. The SE s and CI s from the three analyses are reported on the right side of Table 1, and were nearly identical across the three estimation methods.

Single covariate GMA model.

The analyses were first rerun with only $x1$ as a covariate to illustrate calculation of time-varying GMA d s from a single covariate model with a specified SD (i.e., 1.478, as used in the preceding analysis of a model with multiple covariates). However, in a GMA with a single covariate, SD can be estimated from the model using the variance of the intercept growth factor and residual variances of y with either of two equations. The first equation for estimating SD from the model takes the square root of the sum of the intercept growth factor

variance and the mean of all the residual variances of y (Feingold, 2015)--and is most appropriate when the residual y variances are assumed homogeneous (or are specified to be equal in a Monte Carlo study). The second equation takes the square root of the sum of the intercept growth factor variance and the residual y variance associated with a time code of 0 ($y11$ in this example). (See Appendix B for the input statement for a single covariate model using a specified SD , and the appendix note for the two equations that can be used to estimate SD from the model.)

The two model-estimated approaches for SD estimation produce identical SD s (and thus the same time-varying GMA d s that are calculated using that SD as the denominator) when the residual y variance associated with a time code of 0 equals the mean of all the other residual y variances. In addition, with the equation using a single residual y variance, the GMA d s (but not their CIs) are the same as the GMA d s obtained with the predetermined SD using the previously described two-step approach. Because there is no specified value for SD when SD is estimated from the model, the coefficients for the group differences in slopes and the time-varying GMA d s are obtained simultaneously in a single step.

In the illustrative GMA that included only $x1$ as a covariate and either specified the SD of 1.478 or had SD estimated from the GMA using the $y11$ residual variance, the GMA d s were .396, .739, and 1.028 at T2, T3, and T4, respectively; When SD was instead estimated by Mplus with an equation using the mean of the 4 y residual variances, the corresponding GMA d s were .409, .762, and 1.061. The left side of Table 1 reports the SE s and CIs obtained with the different methods of SD , SE , and CI estimation.. The observed differences among the SE s and CIs obtained with the different estimation methods were not meaningful.

Current Monte Carlo Study

Errors in parameter estimates (*bias*)--differences between expected values of the estimates and the population values--can be assessed with Monte Carlo simulation studies (e.g., Cheung, 2009; Hedges, Pustejovsky, & Shadish, 2012; Lau & Cheung, 2012; MacKinnon, Lockwood, & Williams, 2004). A previous Monte Carlo study specifying a linear GMA had examined bias in both b for the effect of group on slope and the GMA d derived from b --and in the SE s and the 95% CIs for both b and d (Feingold, 2015). The study found that b (as calculated directly by Mplus with the delta method) and GMA d (obtained post hoc with Equation 1 using delta method statistics from a GMA) were both good estimators of their parameters, irrespective of effect size and sample size. (Bootstrap CIs for b could not be evaluated for bias in the earlier study because the version of Mplus then available did not have its current capability to include bootstrap statistics in Monte Carlo analysis.)

This Monte Carlo study of time-varying GMA d s from a quadratic GMA was modeled after the previous study of bias in effect sizes from linear models (Feingold, 2015). However, because these GMA d s were now calculated directly by Mplus instead of with a post hoc equation, the bias in the d s was evaluated with the method Mplus used to examine bias in b in the previous study. Moreover, whereas the earlier study of bias in the SE s and CIs of b only examined SE s and CIs obtained with the delta method, the current study also evaluated bias in bootstrap CIs and SE s for the time-varying GMA d s.

The study first examined bias in the point estimates of the time-varying GMA deltas (the effect size parameters) calculated by Mplus from a quadratic GMA, as a function of sample size, time point, and magnitude of slope differences between the groups. Next, the bias in the *SEs* and *CI*s for these GMA *d*s was examined separately for each estimation method. Because past work indicated that bias in effect sizes was largely related to sampling errors, a key objective of this study was to determine the sample size needed to obtain time-varying GMA *d*s--and their *SEs* and *CI*--that had ignorable amounts of bias.

Method

Eight Monte Carlo simulations--each using 10,000 replications and specifying two linear slope differences (.2000 and .2999) and a single quadratic slope difference (-.0333), crossed with four sample sizes (100, 150, 250, and 500)--were conducted to examine biases in the time-varying GMA *d*s from quadratic models. The values specified for the coefficients for the effects of *x1* on the linear and quadratic slopes were chosen so that the time-specific GMA *d* at T4 would have the same expected value as the GMA *d* from the previously examined linear model (Feingold, 2015), thus allowing for meaningful comparisons of bias in point estimates, *SEs*, and *CI*s of effect sizes from linear and quadratic models.

As in the earlier study, each replication manufactured and analyzed replications from a balanced design with a dichotomous time-invariant covariate (i.e., two groups of equal size), and 4 equidistant time points differing by 1 point (duration = 1, 2, and 3 at T2, T3, and T4, respectively). (See Appendix C for an Mplus input statement for one of the simulations that used the default delta method for *SE* and *CI* estimation.). Bias in *SEs* and *CI*s for time-varying GMA *d*s obtained from the quadratic GMA with the delta method, the standard bootstrap, and the residual bootstrap--with 500 draws used for the bootstraps--was assessed following conventional practices (Feingold, 2015; Muthén & Muthén, 2002, 2017).

The commands and options included in the inputs for the simulations were Monte Carlo counterparts to the inputs for the illustrative analysis that used the quadratic modification of Example 6.10 in the Mplus user's guide for a single covariate model. *SD* for each replication was estimated using the mean of all *y* residual variances (as in Feingold, 2015, and *SD3* in Table 1) because the *y* residual variances were specified to be the same. However, the CINTERVAL option that was specified in the input for the illustrative analysis was not included in the input statements for the simulations because bias in *CI*s is addressed in Monte Carlo studies with *coverage*: the proportion of the replications in which the *CI* contains the parameter (Muthén & Muthén, 2002, 2017). Thus, perfect coverage for the 95% *CI* is .950, and the smaller the bias in the *SE*, the closer the coverage is to .950. In addition, Feingold (2015) proposed and illustrated the use of a *CI* bias statistic obtained by subtracting .950 from the coverage value, which is also reported in the current study.

Results

Point Estimates of Time-Varying Effect Sizes for Quadratic GMA

Table 2 consists of 3 panels of results from the 8 Monte Carlo simulations addressing bias in the point estimate for each effect size parameter (GMA delta). The top panel reports the

results for different sample sizes and effect magnitudes at end of the study (T4); the middle panel reports the corresponding results for effect sizes at T3; and the last panel reports the results for T2. (The GMA d s for T1, which have expected values of zero, are not included in the table.)

The sample size and delta used in each Monte Carlo simulation are reported in the first two columns in Table 2, with the 8 rows in each panel presenting findings obtained by crossing 4 sample sizes with 2 effect size parameters for each of 3 time points. Column 3 (Estimates Avg) reports the mean of the 10,000 GMA d s for each analysis. The differences between the means in this column and the respective GMA deltas of .3464 or .6928 at T4 (reported in column 2) indicate the absolute bias in the time-varying GMA d s (see column 4).

As described in Muthén and Muthén (2002), percent bias was calculated by dividing the absolute bias by the respective parameter (column 2) and multiplying by 100 to convert the proportion to a percent. The percent biases in d s reported in column 5 did not vary notably with effect magnitude or time point but decreased monotonically as sample size increased. According to Muthén and Muthén's (2002) 10% percent rule for ignorable bias in point estimates, the percent bias in GMA d was marginal (clustering at 10% across simulation conditions) at $N = 100$, better (clustering at 5.0%) at $N = 150$, and best (2.9% or less) at $N > 250$.

Standard Errors and Coverage of Effect Sizes for Quadratic Models

Table 3 is comprised of 3 panels that report the findings from the 8 Monte Carlo analyses addressing biases in the SE s and CIs of the time-varying effect sizes--as a function of estimation method, sample size, effect magnitude, and time point. The top panel reports the results of the simulations for different sample sizes and effect magnitude at end of the study (T4); the middle panel reports the corresponding results for T3; and the last panel reports results for T2.

As in Table 2, sample size and delta at T4 for each Monte Carlo simulation are reported in the first two columns. Column 3 reports the Estimates SD , which is the SD of the 10,000 GMA d s over replications. Columns 4, 5, and 6 (SE Avgs) report the means of the 10,000 SE s of d for the delta method, standard bootstrap, and residual bootstrap, respectively. Absolute bias in SE obtained with the delta method was determined by subtracting the statistics in column 4 from the respective Estimates SD s in column 3; the absolute bias in SE for the standard bootstrap was obtained by subtracting the statistics in column 5 from the respective Estimates SD s in column 3; and the absolute bias in the SE for the residual bootstrap was obtained by subtracting the statistics in column 6 from the respective Estimates SD s in column 3.

Each absolute bias value was divided by the respective Estimates SD and multiplied by 100 to obtain the percent bias--as a function of estimation method, effect size parameter at T4, sample size, and time point (see columns 10–12). As with the point estimates, percent bias in the SE s obtained with the delta method did not vary notably with effect magnitude or time point but decreased with sample size. According to Muthén and Muthén's (2002) 5% or less ignorable bias rule for SE s, bias in the SE of d obtained with the delta method was non-

ignorable (consistently greater than 17%) at $N = 100$. At sample sizes of 150 or greater, the bias in the SE obtained with the delta method was acceptable (2% or less). By contrast, the bias in the SE obtained with either the standard or the residual bootstrap was ignorable only at the largest sample size ($N = 500$).

Because the bias in the bootstrap SE in the smallest sample size examined ($N = 100$) was so much larger than the respective bias in the SE obtained with the delta method, a standard bootstrap analysis with the larger effect size parameter specification and $N = 100$ was rerun with the number of draws increased to 2000 to determine whether an inadequate number of draws resulted in overestimation of bias in bootstrap SE s for the time-varying GMA d s. However, quadrupling the number of draws increased rather than reduced the bias in the bootstrap SE s.

The coverage for GMA d obtained with the delta and bootstrap methods are reported in columns 7–9. (CI bias was calculated by subtracting the coverage values from .95, and reported in the last 3 columns of the table.). Muthén and Muthén (2002) suggested coverage should be in the .91–.98 range. Thus, when calculated with the delta method, the coverage for the time-varying GMA d s clustering at .95–.96 across conditions indicated excellent CI estimation. Coverage for the bootstrap CIs was consistently poorer than coverage for the delta method CIs but was only non-ignorable at the smallest sample size (where the coverage was virtually identical whether the standard bootstrap used 500 or 2000 draws).

Discussion

This article introduced and illustrated the application of equations for time-varying effect sizes--model-estimated standardized mean differences at different time points during the study--that are derived from the coefficients for the effects of group on linear and quadratic slopes from a quadratic GMA. For example, in an RCT with 4 time points, T2 may be the end of treatment, T3 a follow-up assessment, and T4 a second follow-up. A time-specific GMA d would convey the treatment effect at each of these post-baseline measurement occasions, and a curve of these time-varying effect sizes could be obtained by plotting them over time. (In a linear model, the plots of time-varying GMA d s would fall on a line instead of a curve.)

When time is measured in weeks rather than coded by measurement occasions, the time-varying GMA d s could be determined for the weeks corresponding to meaningful time points during the study. If, for example, the end of the treatment was at week 6 and the follow-up was at week 10, GMA d s could be obtained for time codes of 6 and 10. (If week is mean centered for use in the GMA, then the time points would be based on the transformed time codes for weeks 6 and 10.)

This article first presented equations for time-varying GMA d s that can be used post hoc with the unstandardized coefficients for the effects of group on linear and quadratic slopes from an analysis conducted with any GMA software. Next, Mplus input statements were presented for directly calculating both time-varying GMA d s and their respective SE s and CIs in Mplus by expanding a GMA example from the Mplus user's guide. Results from an

illustrative analysis using artificial data available with Mplus and quadratic GMA models based on this example were reported to demonstrate the calculation of time-varying GMA d s.

Finally, a Monte Carlo study of time-varying GMA d s was conducted with Mplus that manipulated SE and CI estimation method, sample size, effect size (GMA delta), and time point. Bias in the time-varying GMA d s and their CIs obtained with the delta method was ignorable across all conditions, although the bias in the delta method SE s for these d s was non-ignorable at the smallest sample size ($N = 100$). The Monte Carlo study also found that bootstrap CIs exhibited greater (and generally non-ignorable) bias than did respective delta method CIs. Thus, the default delta method should be used to obtain SE s and CIs for time-varying GMA d s from quadratic GMA.

Randomized vs. Non-Equivalent Group Designs

Time-varying GMA d s are model-estimated standardized mean differences (Cohen's d equivalents) at different time points only when the design uses randomization (e.g., in an RCT) or matching to ensure that the expected mean difference between the two independent groups at baseline (T_1) is zero. Because Equations 1 and 2 do not include a term for the effect of group on the random intercepts from the GMA, GMA d s are derived exclusively from differences between the groups in rate of growth from their respective--and potentially different--baselines.

Assume, for example, that the dummy-coded grouping variable is gender and men have a higher mean than women on the modeled variable at onset but the two sexes exhibit the same non-linear rate of growth over the course of the study, resulting in parallel curves from model-estimated scores of the two groups that intersect the y axis at different points when time = 0 (T_1). The resulting time-varying GMA d s of zero would not indicate gender equality but rather that the sex difference at each time point was the same as the baseline sex difference--and that the sex difference was constant across the study. Thus, Equations 1 and 2 produce GMA d s that are model-estimated standardized mean differences that are adjusted for baseline differences by ignoring effects of group on random intercepts.

When the expected value of the mean difference at baseline is zero, the expected values of the time-varying GMA d s are the same as the expected value of Cohen's d at different time points. Otherwise, the expected value of the unadjusted model-estimated standardized mean difference at each time point would be determined by group differences in the random intercepts and the random slopes. Such unadjusted model-estimated effect sizes could be computed by (a) using the full quadratic GMA model (rather than Equation 2) to estimate the mean for each group at each time point, (b) determining the group difference in the estimated means at each time point, and (c) dividing each estimated mean difference by SD . In a randomized design, the expected values of the time-varying GMA d s obtained with the full quadratic GMA equation would be the same as the expected values of the GMA d s obtained with Equation 2.

However, when the time-varying GMA d s are estimated from the full GMA model and the two groups differ at onset, the effect sizes at baseline are confounded with group differences

in rate of growth. A time-specific GMA d is essentially Cohen's d at each time point minus the Cohen's d at baseline. Given randomization, the expected value of the latter is zero and the expected value of each GMA d equals the expected value of Cohen's d . Similarly, the adjusted and unadjusted mean differences at posttest from an analysis of covariance (ANCOVA) have the same expected value only when the groups are equivalent at baseline. The GMA d is thus less like a Cohen's d from a completely randomized design and more analogous to an effect size calculated with the adjusted mean difference from a pretest-posttest control group design obtained from an ANCOVA.

Moreover, in a randomized design, the expected values of the correlations between the treatment covariate and other covariates in the model are zero. Therefore, the expected values of the time-varying GMA d s do not change when additional covariates are added to the model. However, in the absence of randomization, there may be differences between groups in initial status (baseline) on the outcome variable that are correlated with other covariates, and the expected values of the time-varying GMA d s would then change when other covariates are added to the model.

For example, the illustrative analysis that was based on a quadratic modification of Example 6.10 in the Mplus user's guide had not been conducted with data generated under random assignment specifications. The groups did not have approximately the same baseline means on y , and the baseline scores were correlated with the other covariates in the multiple covariates model. Thus, the time-varying GMA d s were smaller when covariates other than treatment were included in the GMA, and none of these GMA d s were equivalent to Cohen's d (regardless of whether or not additional covariates were included in the model).

Limitations and Directions for Future Research

Issues of bias.—The Monte Carlo study found time-varying GMA d s and their SE s both exhibited greater bias in small samples ($N < 150$) than had been found for GMA d s from a linear model (Feingold, 2015). However, the small sample bias was only non-ignorable for the SE s at the smallest sample size examined ($N = 100$). This would be problematic when using Monte Carlo studies to calculate power for the time-varying GMA d s (Muthén & Muthén's, 2002). However, in a quadratic GMA, power is most likely to be examined for the linear and quadratic coefficients rather than for GMA d s. The SE bias in time-specific GMA d s from small N s would also be problematical for meta-analysis, which requires the use of the squares of SE s (effect size variances), and where the meta-analyst has no control over the sample size used in retrieved studies. Development of bias correction equations for these SE s in small samples may be feasible in future research, and would be particularly useful when findings from such GMA studies are later included in meta-analysis.

Post hoc equations for SE and CI estimation.—Equation 1 was initially presented for post hoc applications with findings for effects of group on slopes from linear GMAs (Feingold, 2009). Subsequent work formulated post hoc equations that could be used in conjunction with Equation 1 to estimate the SE and CI for GMA d (Feingold, 2015). This article introduced Equation 2 for calculating time-varying GMA d s from quadratic GMAs, which is a counterpart to Equation 1 for linear GMA, but did not provide a corresponding

post hoc equation for obtaining *SEs* or CIs for time-varying GMA *d*s. Yet, post hoc equations would be useful to obtain CIs for time-varying GMA *d*s with findings from GMA programs that do not have the capability to generate CIs for user created parameters, and also in meta-analytic applications (Feingold, 2017). Thus, future work on development of post hoc equations for CI estimation for time-specific GMA *d*s from quadratic models is needed. Moreover, such equations should incorporate bias corrections for use with small samples, as the findings from the Monte Carlo study found non-ignorable small sample bias in the *SEs* of time-varying GMA *d*s from quadratic models.

Standardized vs. unstandardized effect sizes for GMA.—Although the importance of effect sizes is now widely recognized (Cumming, 2013), it is not as well known that effect sizes can be either standardized or unstandardized (Kelley & Preacher, 2012). In a randomized study, Cohen's *d* for classical analysis and the GMA *d* (including the last time-specific GMA *d* from a quadratic GMA) both estimate the same parameter, namely, the standardized mean difference between independent groups at the end of the study.

Although standardized effect sizes are popular (Grissom & Kim, 2012), and especially useful in meta-analysis when the same constructs are measured differently across retrieved studies examining the same hypothesis (Borenstein et al., 2009; Schmidt & Hunter, 2014), they have often been criticized or deemed unnecessary (e.g., Baguley, 2009; Greenland, Schlesselman, & Criqui, 1986; Hayes, 2013). A seeming limitation of the current work would be that it addresses standardized but not unstandardized effect sizes for findings from quadratic GMA. However, that limitation is more apparent than real.

The unstandardized effect size that is the counterpart to Cohen's *d* in classical analysis is the mean difference between the two groups in raw score units of the dependent variable (Borenstein et al., 2009). Thus, the equivalent unstandardized effect size in quadratic GMA in a randomized study is the model-estimated raw score mean difference at a particular time point. The numerator in Equation 2 (and Equation 1 for linear GMA) estimates the same parameter as raw score mean differences in classical analysis at a given time point, and is the equation for the respective unstandardized effect size. (If the two groups are not equivalent at baseline, however, these model-estimated unstandardized effect sizes are the equivalent of adjusted mean differences in raw scores in an ANCOVA.) Note that the curves of the effect sizes plotted over time would be the same for unstandardized effect sizes and GMA *d*s.

Moreover, only minor modifications to the input statements for quadratic GMA in Appendices A and B would be needed for Mplus to produce respective time-varying unstandardized effect sizes. Specifically, the denominator in the each equation specified in MODEL CONSTRAINT would be eliminated, and labels applied to parameters for residual variances used to estimate *SD* in a single covariate model are unnecessary in inputs used to obtain unstandardized effect sizes. Without a need to estimate or specify *SD*, a one-step approach would be used to obtain unstandardized effects sizes from both single and multiple covariates models for both linear and quadratic GMA. Parameter bias in the time-varying unstandardized effects sizes and their CIs would be smaller than the bias in the respective standardized effect sizes because there would be no need to estimate an additional parameter (pooled sigma) that contributes to small sample bias.

Conclusions

This article presented a new conceptual model for effect sizes from quadratic GMA, and demonstrated the use of Mplus to obtain time-varying GMA δ s from such non-linear GMAs. The equations introduced can be used post hoc to compute these GMA δ s from linear and quadratic coefficients for slope differences obtained with any GMA software, or directly with specific programs. Irrespective of the approach used to obtain time-varying GMA δ s, the curve formed by plotting the GMA δ s over time effectively communicates the practical significance of changes in effect sizes over the course of the study, and may be preferable to the traditional display of the within-group curves of the model-estimated means of raw scores of the outcomes from non-linear GMA.

Acknowledgments

This work was supported by National Institutes of Health (NIH)/National Institute on Alcohol Abuse and Alcoholism (NIAAA) grant R01AA025069. The content is solely the responsibility of the author and does not necessarily represent the official views of the NIH or NIAAA.

Appendix A: Quadratic GMA with Multiple Covariates

This is an example of calculation of time-varying GMA δ s from a quadratic GMA with multiple covariates

```
DATA: FILE IS ex6.10.dat;
VARIABLE: NAMES ARE y11-y14 x1 x2 a31-a34;
DEFINE: IF (x1 GE -.073) THEN x1=1;
        IF (x1 LT -.073) THEN x1=0;
MODEL:
    i s q | y11@0 y12@1 y13@2 y14@3;
    i s q ON x1 x2;
    y11 ON a31;
    y12 ON a32;
    y13 ON a33;
    y14 ON a34;
    s on x1(bs);
    q on x1(bq);
MODEL CONSTRAINT:
    new(dt4);
    dt4 = ((bs*3) + (bq *9))/1.478;
    new(dt3);
    dt3 = ((bs*2) + (bq *4))/1.478;
    new(dt2);
    dt2 = ((bs*1) + (bq *1))/1.478;
OUTPUT:
    SAMPSTAT CINTERVAL;
```

Note. GMA = Growth Modeling Analysis. Bold type indicates an addition or modification to the input statement for Example 6.10 in the Mplus user's guide.

Appendix B: Quadratic GMA with Single Covariate (Treatment)

This is an example of calculation of time-varying GMA ds from a quadratic GMA with a single covariate

```
DATA: FILE IS ex6.10.dat;
VARIABLE: NAMES ARE y11-y14 x1 x2 a31-a34;
USEVARIABLES=Y11-Y14 X1;
DEFINE: IF (X1 GE -.073) THEN X1=1;
        IF (X1 LT -.073) THEN X1=0;
MODEL: i s q | y11@0 y12@1 y13@2 y14@3;
        i s q ON X1;
        s ON X1(BS);
        q ON X1 (BQ);
        i(v1);
        y11-y14(r1-r4);
MODEL CONSTRAINT:
        new(dt4);
        dt4 = ((bs * 3)+(bq * 9))/1.478;
        new(dt3);
        dt3 = ((bs * 2)+(bq * 4))/1.478;
        new(dt2);
        dt2 = ((bs * 1)+(bq * 1))/1.478;
OUTPUT: SAMPSTAT CINTERVAL;
```

Note. GMA = Growth Modeling Analysis. When estimating *SD* from mean residuals, replace 1.478 with $\sqrt{v1 + r1/4 + r2/4 + r3/4 + r4/4}$. When estimating *SD* from y11 residuals, replace 1.478 with $\sqrt{v1 + r1}$. Italicized moel statements needed only when estimating *SD* from one of the methods using residual variances rather than when the *SD* is specified as 1.478. Bold type indicates an addition or modification to the input statement for Example 6.10 in the Mplus user's guide.

Appendix C: Mplus Input Statement for a Monte Carlo Study Evaluating Time-Varying GMA ds, SEs and CIs from a Quadratic Growth Model

```
MONTECARLO:
        names = y1-y4 x;
        cutpoints= x(0);
        nobs = 250;
        nreps = 10000;
MODEL POPULATION:
        [x@0]; x@1;
        !uncorrelated by default
        i s q | y1@0 y2@1 y3@2 y4@3;
        [y1-y4@0];
        y1-y4*.5;
        [i*s*.2 q*.1];
```

```

i*.25; s*.09; q*.04; i with s*0;
i on x*.0;
s on x*.2999;
q on x*-.0333;
MODEL:
i s q | y1@0 y2@1 y3@2 y4@3;
[y1-y4@0];
y1-y4*.5;
[i*0 s*.2 q*.1];
i*.25; s*.09; q*.04; i with s*0;
i on x*.0;
s on x*.2999;
q on x*-.0333;
s ON x*.2999 (bs);
q ON x*-.0333 (bq);
i (v1);
y1-y4 (r1-r4);
MODEL CONSTRAINT:
new(dt4*.6928);
dt4 = ((bs*3)+(bq*9))/sqrt(v1 + r1/4 + r2/4 + r3/4 + r4/4);
new(dt3*.5388);
dt3 = ((bs*2)+(bq*4))/sqrt(v1 + r1/4 + r2/4 + r3/4 + r4/4);
new(dt2*.3079);
dt2 = ((bs*1)+(bq*1))/sqrt(v1 + r1/4 + r2/4 + r3/4 + r4/4);
OUTPUT:
tech9;

```

References

- Aiken LS, West SG, & Millsap RE (2008). Doctoral training in statistics, measurement, and methodology in psychology: Replication and extension of Aiken, West, Sechrest, and Reno's (1990) survey of Ph.D. programs in North America. *American Psychologist*, 63, 32–50. [PubMed: 18193979]
- Badour CL, Flanagan JC, Gros DF, Killeen T, Pericot-Valverde I, Korte KJ, Allan NP, & Back SE (2017). Habituation of distress and craving during treatment as predictors of change in PTSD symptoms and substance use severity. *Journal of Consulting and Clinical Psychology*, 85, 274–281. [PubMed: 28221062]
- Baguley T (2009). Standardized or simple effect size: What should be reported? *British Journal of Psychology*, 100, 603–617. [PubMed: 19017432]
- Benichou J, & Gail MH (1989). A delta method for implicitly defined random variables. *The American Statistician*, 43, 41–44.
- Bollen KA, & Curran PJ (2006). *Latent curve models: A structural equation perspective*. Hoboken, NJ: Wiley.
- Bollen KA, & Stine RA (1992). Bootstrapping goodness-of-fit measures in structural equation models. *Sociological Methods & Research*, 21, 205–229.
- Borenstein M, Hedges LV, Higgins JPT, & Rothstein HR (2009). *Introduction to meta-analysis*. New York: Wiley.

- Cheung MW (2009). Comparison of methods for constructing confidence intervals of standardized indirect effects. *Behavior Research Methods*, 41, 425–438. [PubMed: 19363183]
- Chorpita BF, Daleiden EL, Park AL, Ward AM, Levy MC, Cromley T, ... & Krull JL (2017). Child STEPs in California: A cluster randomized effectiveness trial comparing modular treatment with community implemented treatment for youth with anxiety, depression, conduct problems, or traumatic stress. *Journal of Consulting and Clinical Psychology*, 85, 13–25. [PubMed: 27548030]
- Cumming G (2013). *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis*. New York, NY: Routledge.
- Efron B, & Tibshirani R (1993). *An introduction to the bootstrap*. Boca Raton, FL: Chapman & Hall.
- Feingold A (2009). Effect sizes for growth-modeling analysis for controlled clinical trials in the same metric as for classical analysis. *Psychological Methods*, 14, 43–53. [PubMed: 19271847]
- Feingold A (2013). A regression framework for effect size assessments in longitudinal modeling of group differences. *Review of General Psychology*, 17, 111–121. [PubMed: 23956615]
- Feingold A (2015). Confidence interval estimation for standardized effect sizes in multilevel and latent growth modeling. *Journal of Consulting and Clinical Psychology*, 83, 157–168. [PubMed: 25181028]
- Feingold A (2017). Meta-analysis with standardized effect sizes from multilevel and latent growth models. *Journal of Consulting and Clinical Psychology*, 85, 262–266. [PubMed: 28068118]
- Feingold A, Washburn IJ, Tiberio SS, & Capaldi DM (2015). Changes in the associations of heavy drinking and drug use with intimate partner violence in early adulthood. *Journal of Family Violence*, 30, 27–34. [PubMed: 25678737]
- Felder JN, Epel E, Lewis JB, Cunningham SD, Tobin JN, et al. (2017). Depressive symptoms and gestational length among pregnant adolescents: Cluster randomized control trial of CenteringPregnancy® plus group prenatal care. *Journal of Consulting and Clinical Psychology*, 85, 574–584. [PubMed: 28287802]
- Goldstein H (2011). *Multilevel statistical models* (4th ed.). Hoboken, NJ: Wiley.
- Goodnight JA, Bates JE, Holtzworth-Munroe A, Pettit GS, Ballard RH, et al. (2017). Dispositional, demographic, and social predictors of trajectories of intimate partner aggression in early adulthood. *Journal of Consulting and Clinical Psychology*, 85, 950–965. [PubMed: 28703604]
- Greenland S & Schlesselman JJ, & Criqui MH. (1986). The fallacy of employing standardized coefficients and correlations as measures of effect. *American Journal of Epidemiology*, 123 203–208. [PubMed: 3946370]
- Grimm KJ, Ram N, & Estabrook R (2016). *Growth modeling: Structural equation and multilevel modeling approaches*. New York: Guilford Press.
- Grissom RJ, & Kim JJ (2012). *Effect sizes for research: Univariate and multivariate application* (2nd ed.). New York: Routledge.
- Gueorguieva R, & Krystal JH (2004). Move over ANOVA: Progress in analyzing repeated-measures data and its reflection in papers published in the *Archives of General Psychiatry*, 61, 310–317. [PubMed: 14993119]
- Hayes AF (2013). *Introduction to mediation, moderation and conditional process analysis: A regression-based approach*. New York: Guilford Press.
- Hedeker D, & Gibbons RD (2006). *Longitudinal data analysis*. Hoboken, NJ: Wiley.
- Hedges LV, Pustejovsky JE, & Shadish WR (2012). A standardized mean difference effect size for single case designs. *Research Synthesis Methods*, 3, 224–239. [PubMed: 26062165]
- Hox JJ, Moerbeek M, & van de Schoot R (2010). *Multilevel analysis: Techniques and applications*. New York: Routledge.
- Jöreskog KG & Sörbom D (2006). *LISREL 8.80 for Windows*. Lincolnwood, IL: Scientific Software International, Inc.
- Kelley K, & Preacher KJ (2012). On effect size. *Psychological Methods*, 17, 137–152. [PubMed: 22545595]
- Kendall M, & Stuart A (1977). *The advanced theory of statistics: Volume 1* (4th edition). New York: MacMillan.

- Kuljanin G, Braun MT, & DeShon RP (2011). A cautionary note on modeling growth trends in longitudinal data. *Psychological Methods*, 16, 249–264. [PubMed: 21517180]
- Lau RS, & Cheung GW (2012). Estimating and comparing specific mediation models in complex latent variable models. *Organizational Research Methods*, 15, 3–16.
- MacKinnon DP, Lockwood CM, & Williams J (2004). Confidence limits for the indirect effect: Distribution of the product and resampling methods. *Multivariate Behavioral Research*, 39, 99–128.
- Marcoulides KM, & Khojasteh J (2018). Analyzing longitudinal data using natural cubic smoothing splines. *Structural Equation Modeling*, 25, 965–971.
- McArdle JJ, & Nesselroade JR (2014). *Longitudinal data analysis using structural equation models*. Washington, DC: American Psychological Association
- Muthén B, & Muthén LK (2000). Integrating person-centered and variable-centered analyses: Growth mixture modeling with latent trajectory classes. *Alcoholism: Clinical and Experimental Research*, 24, 882–891.
- Muthén BO, Muthén LK, & Asparouhov T (2016). *Regression and mediation analysis using Mplus*. Los Angeles, CA: Muthén & Muthén.
- Muthén LK, & Muthén BO (2002). How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling*, 4, 599–620.
- Muthén LK, & Muthén BO (2017). *Mplus user's guide* (8th ed). Los Angeles, CA: Muthén & Muthén.
- Parra-Cardona JR, Bybee D, Sullivan CM, Rodríguez MMD, Tams L, & Bernal G (2017). Examining the impact of differential cultural adaptation with Latina/o immigrants exposed to adapted parent training interventions. *Journal of Consulting and Clinical Psychology*, 85, 58–71. [PubMed: 28045288]
- Preacher KJ, Wichman AL, MacCallum RC, & Briggs NE (2008). *Latent growth modeling*. Los Angeles, CA: Sage.
- Raudenbush SW, & Bryk AS (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed). Thousand Oaks, CA: Sage.
- Rosseel Y (2012). *lavaan: An R package for structural equation modeling*. *Journal of Statistical Software*, 48, 1–36.
- SAS Institute Inc. (2011). *SAS/STAT® 9.3 user's guide*. Cary, NC: SAS Institute Inc.
- Schmidt FL, & Hunter JE (2014). *Methods of meta-analysis: Correcting error and bias in research findings* (3rd ed). Thousand Oaks, CA: Sage.
- Stice E, Rohde P, Shaw H, & Gau JM (2017). Clinician-led, peer-led, and internet- delivered dissonance-based eating disorder prevention programs: Acute effectiveness of these delivery modalities. *Journal of Consulting and Clinical Psychology*, 85, 883–895. [PubMed: 28425735]

Table 1

SEs and CIs for Time-Varying GMA ds from Quadratic GMA, as Function of Estimation Methods and Time Points in Illustrative Analysis

<u>Covariates in Example 6.10 Dataset Included in the Quadratic GMA</u>												
	x1 only						All 3 Covariates					
	SD1			SD2			SD3			SD1		
	SE	95% CI		SE	95% CI		SE	95% CI		SE	95% CI	
Delta Method												
T4	.134	.766,	1.291	.138	.758,	1.299	.155	.756,	1.365	.105	.660,	1.071
T3	.104	.536,	.942	.106	.531,	.948	.118	.530,	.994	.084	.460,	.788
T2	.068	.263,	.530	.069	.261,	.532	.075	.261,	.556	.058	.222,	.448
Standard Bootstrap												
T4	.133	.747,	1.289	.136	.752,	1.285	.161	.762,	1.411	.105	.671,	1.076
T3	.100	.536,	.930	.103	.545,	.957	.121	.548,	1.029	.080	.473,	.784
T2	.066	.271,	.515	.067	.276,	.528	.076	.272,	.573	.055	.224,	.447
Residual Bootstrap												
T4	.131	.766,	1.282	.135	.773,	1.307	.153	.779,	1.383	.096	.689,	1.061
T3	.100	.547,	.936	.102	.550,	.949	.116	.565,	1.028	.080	.460,	.775
T2	.066	.261,	.527	.067	.266,	.530	.074	.269,	.574	.056	.225,	.442

Note. $N=500$. GMA = growth modeling analysis, *SE* = standard error, CI = confidence interval, *SD1* = 1.748, *SD2* = *SD* estimated with y_{11} residual variance, *SD3* = *SD* estimated with mean of all y (y_{11} - y_{14}) residual variances. CIs for time-varying GMA ds from the single covariate model (x1 only) cannot be compared with respective CIs from the multiple covariates model (using 3 covariates) because point estimates differ between the two types of models.

Table 2

Monte Carlo Analyses of Points Estimates for the GMA Effect Size for the Group Difference in Slopes for a Quadratic GMA as a Function of Sample Size, Effect Magnitude, and Time Point

Monte Carlo Results			Bias Estimates	
<i>N</i>	<i>δ</i>	<u>Estimate</u>	<u>Raw</u>	<u>Percent</u>
T4				
100	.3464	.3847	.0383	11.06
100	.6928	.7619	.0691	9.97
150	.3464	.3663	.0199	5.74
150	.6928	.7288	.0360	5.20
250	.3464	.3564	.0100	2.89
250	.6928	.7111	.0183	2.64
500	.3464	.3504	.0040	1.15
500	.6928	.7006	.0078	1.13
T3				
100	.3081	.3397	.0316	10.26
100	.5388	.5912	.0524	9.73
150	.3081	.3236	.0155	5.03
150	.5388	.5652	.0264	4.90
250	.3081	.3156	.0075	2.43
250	.5388	.5520	.0132	2.45
500	.3081	.3108	.0027	.88
500	.5388	.5443	.0055	1.02
T2				
100	.1925	.2115	.0190	9.87
100	.3079	.3372	.0293	9.52
150	.1925	.2015	.0090	4.68
150	.3079	.3223	.0144	4.68
250	.1925	.1968	.0043	2.23
250	.3079	.3150	.0071	2.31
500	.1925	.1940	.0015	.78
500	.3079	.3107	.0028	.91

Note. GMA = growth modeling analysis.

Table 3

Monte Carlo Analyses of the Standard Errors and Coverage for the Effect Sizes (GMA ds) as a Function of and Sample Size, Effect Magnitude, and Time Point for Three SE and CI Estimation Methods

Monte Carlo Results							Bias Estimates							
<u>N</u>	<u>δ</u>	<u>SD</u>	<u>Avg</u>			<u>Coverage</u>			<u>Percent</u>			<u>Confidence Interval</u>		
			Delta	BTSP	RESB	Delta	BTSP	RESB	Delta	BTSP	RESB	Delta	BTSP	RESB
T4														
100	.3464	.6550	.7705	1.3571	1.1180	.962	.985	.982	17.63	107.19	70.69	.012	.035	.032
100	.6928	.7269	.9081	1.4109	1.3417	.964	.982	.979	24.93	94.10	84.58	.014	.032	.029
150	.3464	.4659	.4671	.6647	.7875	.957	.976	.973	.26	42.67	69.03	.007	.026	.023
150	.6928	.4906	.4897	.7823	.7871	.959	.974	.972	.18	59.46	60.44	.009	.024	.022
250	.3464	.3444	.3383	.4083	.3734	.952	.966	.962	1.77	18.55	8.42	.002	.016	.012
250	.6928	.3546	.3478	.4558	.4274	.952	.966	.961	1.92	28.54	20.53	.002	.016	.011
500	.3464	.2376	.2350	.2467	.2388	.950	.954	.953	1.09	3.83	.51	.000	.004	.003
500	.6928	.2431	.2407	.2533	.2430	.951	.957	.953	.99	4.20	.04	.001	.007	.003
T3														
100	.3081	.4343	.5366	.8770	.7308	.967	.984	.982	23.56	101.93	68.27	.017	.034	.032
100	.5388	.4932	.6335	.9997	.9004	.963	.979	.979	28.45	102.70	82.56	.013	.029	.029
150	.3081	.2972	.2955	.4668	.5211	.958	.977	.973	.57	57.07	75.34	.008	.027	.023
150	.5388	.3184	.3146	.5403	.5436	.960	.975	.971	1.19	69.69	70.73	.010	.025	.021
250	.3081	.2159	.2117	.2649	.2363	.953	.965	.962	1.95	22.70	9.45	.003	.015	.012
250	.5388	.2244	.2199	.3124	.2880	.955	.967	.961	2.01	39.22	28.34	.005	.017	.011
500	.3081	.1486	.1468	.1538	.1491	.951	.955	.952	1.21	3.50	.34	.001	.005	.002
500	.5388	.1533	.1518	.1603	.1525	.953	.957	.952	.98	4.57	.52	.003	.007	.002
T2														
100	.1925	.2710	.3353	.5361	.4481	.964	.984	.981	23.73	97.82	65.35	.014	.034	.031
100	.3079	.2995	.3824	.6303	.5436	.965	.981	.979	27.68	110.45	81.50	.015	.031	.029
150	.1925	.1842	.1829	.3088	.3255	.959	.976	.973	.71	67.64	76.71	.009	.026	.023
150	.3079	.1946	.1925	.3296	.3332	.960	.974	.971	1.08	69.37	71.22	.010	.024	.021
250	.1925	.1332	.1313	.1638	.1466	.955	.966	.964	1.43	22.97	10.06	.005	.016	.014
250	.3079	.1372	.1352	.1928	.1783	.955	.968	.964	1.46	40.52	29.86	.005	.018	.014
500	.1925	.0917	.0910	.0945	.0926	.953	.958	.954	.76	3.05	.98	.003	.008	.004

Monte Carlo Results										Bias Estimates					
<i>N</i>	<i>δ</i>	<i>SD</i>	Avg			Coverage			Percent			Confidence Interval			
			Delta	BTSP	RESB	Delta	BTSP	RESB	Delta	BTSP	RESB	Delta	BTSP	RESB	
500	.3079	.0939	.0934	.0981	.0940	.954	.959	.955	.53	4.47	.11	.004	.009	.005	

Note. Avg = average *SEd* across replications, Coverage = 95% coverage for time-varying *ds*, CI = 95% confidence interval, Delta = delta method, BTSP = standard bootstrap, RESB = residual bootstrap.