

© Copyright 2016

Erin Winters

**Relationship between Comprehensive School Readiness Profiles
and Longer-Term Academic Growth**

Erin Winters

A thesis submitted in partial fulfillment
of the requirements
for the degree of

Master of Education

University of Washington

2016

Committee:

Elizabeth Sanders

Robert Abbott

Program Authorized to Offer Degree:

College of Education

University of Washington

Abstract

Relationship between Comprehensive School Readiness Profiles
and Longer-Term Academic Growth

Erin Winters

Chair of the Supervisory Committee:

Elizabeth Sanders, Assistant Professor

Educational Psychology

Understanding the links between school readiness and child outcomes is essential in helping educators understand crucial early childhood intervention points to help children adjust to their academic environments. Although cognitive indicators once defined school readiness, the prevailing research over the past 15 years indicates that social-emotional indicators play an important role. Whereas prior research has examined shorter-term links between social-emotional indicators and academic achievement, the present study seeks to extend the research to investigate the longer term effects. Using a subsample of data collected in a nationally representative longitudinal dataset (ECLS-K), the present study used latent growth curve modeling that incorporates school membership and complex survey sample weights on data from grades K through 8. Final results showed that the early conceptualization of kindergarten school readiness profiles that are inclusive of social-emotional skills do systematically contribute to longer-term math and reading growth through the end of eighth grade, but primarily this is simply the difference between one “comprehensive positive” profile and the other three profiles. It may be the case that the profiles are simply defining high vs. low risk, without distinction among the high risk profiles.

Relationship between Comprehensive School Readiness Profiles and Longer-Term Academic Growth

“School readiness” refers to the skills and/or abilities that, when mastered, allow a student to flourish academically and socially in school. Current educational policy and research places a heavy emphasis on cognitive skills and emergent literacy when discussing school readiness, but research by the National Educational Goals Panel (NEGP) (Kagan, Moore, & Bredekamp, 1998) has suggested a five-factor model of readiness, including: 1) Physical well-being and motor development, 2) Social-emotional development, 3) Language development, 4) Cognition and general knowledge, and 5) Approaches to learning. The present paper focuses on investigating the evidence linking the second and fourth factors to each other in the context of early school readiness.

Social-emotional Learning and Academic Outcomes

Social-emotional learning refers to the ways adults and children develop and apply the knowledge and skills needed to make and accomplish goals, experience and express empathy for others, develop positive relations, and process and manage their emotions (CASEL, 2013). In the past decade, there has been an increase in the perceived importance of social-emotional learning (SEL), with rigorous randomized trials showing that improvement of academic and behavioral outcomes for students involved in SEL interventions and curricula (Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011; Sklad, Diekstra, Ritter, Ben, & Gravesteyn, 2012). In fact, in their comprehensive meta-analysis of 213 school-based universal SEL programs, Durlak et al. (2011) found that students involved in SEL interventions had significantly improved social-emotional skills, attitudes, and behavior, as well as academic performance. Overall, they estimated that an 11-percentile point gain in academic achievement would be predicted for

students who participated in a SEL intervention compared to no intervention. There is also a body of research that provides evidence linking preschool children's behavior skills with academic outcomes: in their recent cross-sectional study of 61 preschool students, Doctoroff, Fisher, Burrows, and Edman (2016) found significant negative correlations between problem externalizing behavior and math skills as measured by the *Test of Early Mathematics Ability* (3rd Edition), as well as significant positive correlations between adaptive behavior and math skills.

These recent studies are consistent with earlier research on social-emotional skills. As part of a longitudinal study of Head Start participants, Bramlett, Scott, and Rowell (2000) found that parent and teacher ratings of social-emotional skills of 104 students at the beginning of first grade were predictive of reading and math performance at the end of first grade. Similarly, Jimerson, Egeland, and Teo (1999) found that scores on the Child Behavior Checklist-Teacher Report Form (CBCL-TRF) were predictive of academic trajectories. Specifically, Jimerson et al. created two composites (one from 1st through 3rd grades and one from 1st through 6th grades using the CBCL-TRF scores) that were then used (separately) to predict individual differences in expected academic trajectories at ages 12 and 16. Of the two composites, only the complete composite (using all grades 1-6) was significantly predictive of individual differences in math achievement at age 16. Generally, students whose teachers reported more problem behavior were predicted to be lower than expected, while students whose teachers reported fewer problem ratings were above the expected trajectory.

School Readiness within the NEGP Framework

The only systematic study on school readiness and child outcomes using the framework identified by the NEGP was conducted by Hair (2006). Specifically, the authors used the nationally representative Early Childhood Longitudinal Study-Kindergarten Cohort (ECLS-K)

dataset to develop a set of school readiness profiles integrating the five factors (for review, these were: 1) Physical well-being and motor development, 2) Social-emotional development, 3) Language development, 4) Cognition and general knowledge, and 5) Approaches to learning). Descriptive analysis of the five school readiness indices indicated that over 99% of students were considered “on track” on the *approaches to learning* factor. Because it was nearly constant (i.e., no variability), it was excluded from further analyses. Using standardized scale scores from parent, teacher, and child direct assessment items measured in kindergarten (other than “on track” items), a *k*-means cluster analysis was then employed to identify four school readiness profiles, as follows:

1. Comprehensive positive development (students scored at or above the mean on all 4 school readiness factors, which included 31% of the sample);
2. Social-emotional health strengths (students scored above average on the health scale, far above average on the social-emotional scale, but below the mean on the general knowledge and language development scale, which included 34% of the sample);
3. Social-emotional risk (students scored below average on all scales, but far below average on the social-emotional scale, which included 22% of the sample); and
4. Health risk (students scored below the mean on the health, general knowledge, and language development scales, but at or above average on the social-emotional scale, which included 13% of the sample).

Using these results, Hair et al. (2006) then tested the predictive utility of the profiles on end-of-first-grade outcomes; results showed that the students assigned to the comprehensive positive development profile performed best on both the reading and math outcomes, and that the social-emotional health strengths profile performed better than the social-emotional risk profile.

The health risk profile predicted the worst performance in reading and math. Importantly, in 2006, the ECLS-K data collection was not yet complete, and as such, the analyses by Hair et al. (2006) were limited to relationships between kindergarten readiness and first grade outcomes only.

Current Study

The current study replicates and extends the research by Hair et al. (2006) with the ECLS-K longitudinal dataset's later waves by:

1. Performing a latent growth curve analysis in a structural equation modeling framework to test whether longer term growth patterns are predicted by school readiness profiles;
2. Using school membership to account for non-independence of observations of students within the same school; and
3. Using complex survey weights to maintain the national representativeness of the data.

Methods

Dataset

Data collection. The present study uses extant data from the Early Childhood Longitudinal Study-Kindergarten Cohort (ECLS-K; Westat, 2009) study, which is a federally sponsored longitudinal, nationally representative dataset that surveyed public and private school kindergarten students and their parents, teachers, and school administrators starting in the fall of 1998. To reduce data collection costs, the ECLS-K utilized a clustered dual frame, multistage sampling design. Ultimately, 21,260 students from were identified for inclusion in the original study. Data collection continued in the spring of 1999 (end of kindergarten), fall of 1999 (fall of first grade), spring of 2000 (spring of first grade), and the springs of 2002 (third grade), 2004

(fifth grade) and 2007 (eighth grade). The fall 1999 sample only included a 30% sub-set of the original dataset.

For this study, we were interested in investigating the growth trajectories of first-time kindergarten students only, which limits the sample to 17,219 students that were identified as entering kindergarten for the first time in the initial year of data collection. Further, because we were interested in replicating and extending the analyses by Hair et al. (2006) but only had the publically available measure-level data (not item-level), the sample was further limited to only children who had sufficient measurements to be assigned a profile (as profiles were being used as predictors, not outcomes); hence the final sample used for analyses included $N = 4,130$ first-time kindergarten students with complete data on early kindergarten readiness indicators. Data employed for analysis were seven measurement occasions that began in the fall of kindergarten through the spring eighth grade.

Measures. The ECLS-K included teacher, parent, and school administrator surveys, as well as direct achievement tests for children throughout the study, and self-report measures in the final three years of data analysis. These included measures of perceived social-emotional skills, general health, and approaches to learning. Throughout the study, height and weight was measured and the child's BMI was calculated, based off of the CDC BMI guidelines.

The achievements tests in the first four rounds of data collection included language and literacy, mathematical thinking, and general knowledge. The later batteries included math, reading, and science inventories. Since developmentally appropriate measures were used as direct assessments throughout the study, Item Response Theory (IRT) scale scores were developed by the researchers to provide a consistent scale for longitudinal analysis of outcome data. Using IRT, scale scores were developed by the researchers to simulate the students taking

same scale throughout the study, as opposed to the different developmentally appropriate measures that were used. These IRT scaled scores provided by the researchers were used as the longitudinal outcome measures for reading and math in the present analysis.

Data Analysis

Software. Analyses were completed using *Mplus 7.2* (Muthén & Muthén, 1998-2014) and *SPSS 19* (SPSS Inc., 1989-2010).

Missing data. Please see Hair et al. (2006) for a complete description of missing outcome data for the entire sample. Despite not having a large percentage of the original sample available for analysis (due primarily to the lack of item-level data in the publically available dataset to create profile assignments for all students, as well as to dataset's own design to only measure 30% of the sample at early kindergarten), the use of longitudinal survey weights adjusts for systematic non-response issue keeping the results nationally representative.

In addition to the missingness by design as described above, for the present study's sample of 4,130 students, some outcome data were also missing at various measurement occasions. Complete outcome data for our sample was available from the first wave (beginning of kindergarten), but thereafter some students were missing data at one measurement occasion only; others were missing all data after kindergarten, and yet others had missing data in a few spots only. Indeed, there were 40 distinct missing data patterns found for this sample of students in reading, and 42 missing data patterns found among students in math, with the highest amount of missing data found at the beginning of first grade (3,085 were missing data) and second highest amount at end of eighth grade (2,261 were missing data, with just over 50% of these cases also missing data at the beginning of first grade). For the purpose of the present study, we treated the missing data across outcomes for the sample of 4130 students as "missing at random"

(MAR) and used *Mplus*' full information likelihood approach to utilize all available data to estimate the model parameters. This approach is preferred over listwise deletion or mean imputation because either of the former methods could either bias the variances and covariances or decrease statistical power (Schafer & Graham, 2002).

Latent growth curve analysis. This study used a structural equation modeling (SEM) framework for modeling growth, which is also known as "latent growth curve modeling" (LGCM). This type of modeling is employed for estimating mean patterns of growth in outcomes over time. LGCM typically assumes 1) at least three observations of a continuous outcome across subjects (at least two per subject), 2) scores at each time point measured with the same scale and units, and 3) time-structured data, where observations are assumed to have been tested on the same schedule (Kline, 2010), although this constraint can also be relaxed.

Structural equation modeling. Growth models in an SEM framework are multilevel models that take into account the unique relationship between repeated measures (Bauer, 2003; Kline, 2010). In SEM, scores are clustered within individuals, and are therefore not independent from each other. *Mplus*, along with other SEM software, automatically adjust for this non-independence during standard analysis. First, the initial growth model with only the repeated measures variables is estimated. This initial model includes the intercept, or mean initial value of the measure at baseline, and the linear slope, depicting the estimated mean change in the outcome variable per specified unit of time. If there are at least four measurement occasions, then a quadratic slope may also be tested. Additionally, the means, variances, and covariance can be estimated simultaneously, and can be separately tested against the null hypothesis (Muthén & Muthén, 2010).

Using the SEM framework, there are multiple ways to assess and compare model fit. The χ^2 significance test, where a significant χ^2 results indicates *poor* model fit, is a commonly used model fit assessment. However, the χ^2 is a poor choice for large samples in SEM (Kline, 2010). Instead, other options for assessing model fit include the Bayesian Information Criterion (BIC), Aikake's Information Criterion (AIC), Bentler's Comparative Fit Index (CFI), and the Standardized Root Mean Square Residual (SRMR). For both the AIC and BIC, a smaller number (closer to zero) is indicative of better fit. As a variation of the χ^2 statistic ($\chi^2 - 2df$), the AIC penalizes less parsimonious models, leading to a higher value and poorer fit. This statistic indicates which of the competing models is most likely to be replicated due to chance (Kline, 2010). The BIC includes parameter estimation information, but incorporates the log-likelihood of a model, with parameters only being included as they meet requirements that become more demanding as the sample size increases (Vrieze, 2012). For the CFI, values range from 0 to 1, with higher values indicating the relative improvement in fit of the tested model from the null model. The SRMR is a measure of the mean absolute correlation residual, and is the difference between the observed and predicted correlations. A SRMR close to zero indicates good model fit, while $SRMR \leq .08$ indicates acceptable model fit (Kline, 2010). The BIC, AIC, CFI and SRMR model fit statistics will be included in this analysis.

Time structure. Because the ECLS-K dataset was collected over the course of 8.5 years, it was important that time be appropriately coded for analysis, with our time unit code as years. The first intercept was coded at 1 for all time points; with the slope coded at 0 for fall of kindergarten, 0.5 for spring of kindergarten, 1.0 for fall of first grade, 1.5 for spring of first grade, 3.5 for spring of third grade, 5.5 for spring of fifth grade, and 8.5 for spring of eighth

grade. The intercept was purposefully considered as the first time point so that the profiles could be used to predict initial status as well as growth.

Model specification and selection. For both the reading and math measures, we specified a base growth model (Model 1) that freely estimated the variances and covariances among latent variables. If the model did not converge, then the model was re-specified to constrain the variance of the parameter of concern to 0. Following the base growth model, the school readiness profiles were added to the model as nominal growth predictors (Model 2). Then, the fall kindergarten school variable was added to the model to account for non-independence at the school level (Model 3). Finally, complex survey weights were added to the model to maintain the generalizability of the results, and to help account for any non-response bias (Model 4).

Results

Profile Membership

As previously mentioned, due to missingness on profile indicators in the publically available dataset, profile membership was assigned to 4,130 of the students from the original sample. Profiles were constructed using the “liberal” profile definitions put forth in the original Hair et al. (2006) study, resulting in 2,283 students (55%) assigned to the “Comprehensive Positive” (CP; reference group) profile, 868 (21%) assigned to the “Social-Emotional and Health Strengths” (SEHS) profile, 571 (14%) assigned to the “Social-emotional Risk” (SER) profile, and 408 (10%) assigned to the “Health Risk” (HR) profile. Despite following the prescription by Hair et al. (2006), the present study’s proportions did differ significantly from the classification proportions found in the original study (z -test $ps < .05$ except the social-emotional risk profile); these differences are likely to be due to the original study’s access to the restricted-use dataset which had item-level information available for all the kindergarten readiness indicator measures.

This said, the only proportional difference that differed greatly (more than 13%) was on the comprehensive positive profile classification, with 55% of the present study's sample was assigned to this profile whereas the earlier study only classified 30% of the sample.

Observed Measures

Means, standard deviations, and zero-order intercorrelations for the observed variables used for the present study are shown in Table 1.

Latent Growth Curve Analysis

Growth curve model fit indices are shown in Table 2. Figures 1 and 2 illustrate the path diagrams for the Reading and Math IRT Scale models, respectively. Figure 3 depicts the mean predicted trajectory growth curves by profile for each outcome in Panels A (Reading) and B (Math).

Results for Models of Reading. As outlined in the Methods section, four separate models were conducted in order to observe the model fit changes as more aspects were added to the model. Because of convergence issues surrounding the variance of the seventh reading measurement occasion (end of grade 8), all models had the variance of that time point constrained to zero. Table 2 shows us that for the initial reading growth model (Model 1), we do not have strong initial fit (CFI=.79, SRMR=.17). Note that CFIs of .95 or higher and SRMRs of .05 or lower are general rules of thumb for acceptable model fit. Nevertheless, given that growth models are far more constrained than typical structural equation models, the lack of ideal model fit was not surprising. For fit comparison to later model iterations, we'll be using the BIC. As seen in Table 2, Model 1 had a BIC = 156714. In Model 1, all latent variable means were statistically different than zero: intercept $M = 35.45$, linear slope $M = 30.18$, and quadratic slope

$M = -1.69$, indicating that the average initial score was 35.45, with an average growth of 30.18 points per year, and an average deceleration of 1.69 points per year.

Model 2, where we introduced the profile predictors, had a BIC of 154851, a $BIC_{\text{difference}} = 1863$, indicating an improvement in goodness of fit between Model 1 and Model 2. For Model 2, all latent variable means were statistically different than zero ($p < .001$) for the reference group, Comprehensive positive: intercept $M = 40.93$; linear slope $M = 33.40$; and quadratic slope $M = -1.99$. This model indicates a higher intercept and faster growth patterns for the reference group as Model 1 predicted for the full sample.

As seen in Table 2 Model 2 (see also Figure 3A), profiles were significant predictors of differences on the intercept, with all three of the SEHS, SER, and HR profiles scoring lower than the CP profile on early kindergarten reading skills. All profiles were also significant predictors of individual differences in linear slope, with all three showing significantly lower growth than the CP profile. Furthermore, all profiles were predictive of differences in quadratic slope, indicating that their growth was decelerating at a lower rate than the reference group.

As seen in Table 2 Model 3, where we included a random effect (school membership) at the intercept to account for non-independence of students within schools, showed a BIC = 156832, indicating no change in model fit as measured by the BIC when accounting for non-independence. Notably, the only change between this model and the prior model was an increase in the coefficient standard errors (leading to a decrease in z -test values).

The final model, Model 4, where we introduced the complex survey weights to maintain the generalizability of the results and account for any systematic non-response patterns, had a BIC = 231141 and $BIC_{\text{difference}} = -76290$, indicating a large *decrease* in model fit (as a low BIC

indicates better model fit). In Model 4, all latent variable means were statistically different than zero: intercept $M = 40.04$; linear slope $M = 33.26$; and quadratic slope $M = -2.01$.

These final results, after accounting for both school membership and sampling weights, show that the differences between the groups remained significant on the intercept (early kindergarten reading skills: CP had higher initial skills than the other three profiles), on linear slope (CP scores increased the most per year), and quadratic slope (SEHS, SER, and HR growth slows at a lesser rate than CP growth).

Results for Models of Math. As for math, four separate math models were run in order to observe the model fit changes as different parts were added to the model. Similar to the reading model, because of convergence issues surrounding the variance of the seventh math measurement occasion (end of grade 8), all models had the variance of that time point constrained to zero. As seen in Table 2 (Model 1), the model fit for the general growth model was better for the math outcome than it was for reading (CFI=.92, SRMR=.08). While CFI = .92 is not ideal (.95 or higher would be preferred), the SRMR = .08 falls within the acceptable range (although .05 or lower would be preferred). In Model 1, all latent variable means were statistically different than zero: intercept $M = 25.81$; linear slope $M = 24.99$; and quadratic slope $M = -1.35$. These results indicate that the average initial score was 25.81, with an average growth of 24.99 points per year, with growth decelerating by 1.35 points per year.

Model 2, where we introduced the profile predictors, had a BIC of 143502, a $BIC_{\text{difference}} = 2472$, indicating an improvement in goodness of fit between Model 1 and Model 2. For Model 2, all latent variable means were statistically different than zero for the reference group, Comprehensive positive: intercept $M = 31.32$; linear slope $M = 27.59$; and quadratic slope $M = -1.59$. As seen in Table 3 (and Figure 3B), profiles were significant predictors of differences on

the intercept (recall that the reference group is CP), which showed that the SEHS, SER, and HR profiles were significantly lower than the CP profile on early kindergarten math skills.

Additionally, they were also significant predictors of difference on linear slope and quadratic slope. Although CP scores grew fastest (linear slope) compared with the other profiles, the other profiles also had less deceleration than CP. Consistent with the findings for reading, the math trajectories are remarkable in that the CP profile is distinctively different from the other three profiles, which show very little difference among each other.

Model 3, where we accounted for school membership as a random factor, had $BIC = 143502$, a Model 2-Model 3 $BIC_{\text{difference}} = 0$, indicating no change in model fit when accounting for non-independence between Models 2 and 3. For Model 3, all latent variable means were statistically different than zero: intercept $M = 31.32$; linear slope $M = 27.59$; and quadratic slope $M = -1.59$. The only change was an increase in the standard errors, and subsequent decrease in z-scores, as seen in Table 3.

The final model, Model 4, where we introduced the complex survey weights to maintain the generalizability of the results and account for systematic non-response patterns, had $BIC = 231141$ and $BIC_{\text{difference}} = -76290$, indicating a decrease in model fit (as a low BIC indicates better model fit) from Model 3 to Model 4. In Model 4, all latent variable means were statistically different than zero for the reference group: intercept $M = 30.14$; linear slope $M = 27.87$; and quadratic slope $M = -1.65$.

These final results, after accounting for both school membership and sampling weights, show that the differences between three of the groups remained significant on the intercept (early kindergarten math skills: CP had higher initial skills than all profiles), and SEHS and SER remained significant predictors of difference for both linear and quadratic slopes.

Discussion

Although the Hair et al. (2006) study found that their derived school readiness profiles predicted end-of-grade 1 outcomes using multiple linear regression, they did not control for school membership or sample survey weights in their analyses (and their results were limited to end of grade 1). The present study extends the prior research by using school readiness profiles to predict outcomes through the end of grade 8, and appropriately accounts for both school membership as well as complex sample survey weights. The results showed that profiles were significant predictors of differences on the intercept for both reading and math outcomes, indicating, as expected, that school readiness profile membership in fall of kindergarten does predict academic performance in fall of kindergarten. More importantly, model estimates consistently showed that three of the profiles also differed from the comprehensive positive profile on linear and quadratic growth estimates for both outcomes, even after the incorporation of survey weights.

This said, the plots of predicted values (Figure 3, Panels A and B) show that the comprehensive positive profile had a very different trajectory than the other three profiles. Indeed, the other three profiles were nearly indistinguishable from each other, which begs the question of the construct validity of the social-emotional and health strengths profile, at least in terms of predicting academic outcomes.

Limitations

Methodologically it would be prudent to also center the intercept on the final timepoint (end of grade 8), rather than the first (beginning of grade K), to predict final status of these students. However, the present study was interested in establishing whether the profiles would be

predictive of the initial timepoint for the subset sample used (as a check on the validity of the profile construction).

With respect to the dataset's limitations in particular: the age of the dataset itself is also a concern. The ECLS-K dataset is an older dataset- data collection began in 1998, and concluded in 2007. Since the inception of this longitudinal study, considerably more research has been done on social-emotional learning, and it may be possible to better measure these aspects of children's development. Furthermore, the U.S. landscape has changed dramatically since study planning began (e.g., the age of electronic devices in the classroom), and as such there may be fundamental differences in the population as a whole that may be important to attend to. In a similar vein, this dataset lacks a longitudinal social-emotional outcome. Given the paucity of research on longer-term studies of social-emotional outcomes, a systematic study of social emotional growth trajectories in concert with academic trajectories seems warranted to better establish baseline norms.

In regards to the use of the profiles established by Hair et al. (2006): the inclusion of social-emotional skills as a core factor of school readiness may be partly to blame for the non-significant links between the HR profiles and math growth after the inclusion of complex survey weights for generalizability. Social-emotional skills are widely regarded as a difficult construct to measure, due to the varied cultural expectations of children at different developmental stages. Without raters that are cross-culturally competent students' social-emotional skills may be underestimated based on differing cultural lenses (Garner, Mahatmya, Brown, & Vesely, 2014). Even with the

Of equal concern is the large amount of missing data used to assign the profile membership. Because the original researchers had access to the restricted item-level data (Hair,

2006), they were able to better compute index scores for each of the school-readiness facets when students had partial index scores. Due to missing data, this study was only able to assign profiles to 4,130, as opposed to the total number of first time kindergarten students (17,219) achieved by Hair et al. (2006).

Last but not least, another potential limitation of the present research is the reliance of Hair et al.'s (2006) results using a *k*-means cluster analysis instead of a latent class model to determine profile membership (McLachlan & Peel, 2000). Latent class analysis empirically tests patterns in the association between data points, and classifies cases using a probability (model-based) maximum likelihood estimation method (Kline, 2010). *K*-means cluster analysis, on the other hand, assigns cases to classes based on an iterative process where observations are assigned to the closest mean.

Future Research Directions

The present study has found additional evidence to support a systematic link between school readiness defined by Hair et al. (2006) with longer-term academic outcomes. However, it may be the case that the profiles used are simply defining high vs. low risk children, without any real distinction among the high risk profiles.

Taking the limitations of the present study as a starting point, future research should first investigate the potential school readiness profile designations that would be identified from a number of academic and social-emotional indicators using a latent class analysis approach in SEM in concert with the longer-term outcomes. Additionally, it would be interesting to explore the appropriateness of complex survey weights when generalizability is not a concern of the researcher.

Importantly, the National Center for Educational Statistics is five years into a new ECLS cohort. This dataset promises to include a longitudinal social-emotional outcome, with data collection occurring via a self-report measure in third, fourth, and fifth grades. Social-emotional measures will be collected on both teachers and parents in kindergarten through second grades. Those data might be used to examine how growth in social emotional skills (rather than just at an early timepoint) relates to growth in academic skills. Ultimately, those analysis results could be fruitfully used to open the discussion for ongoing social-emotional interventions that support students' longer term academic outcomes.

References

- Bauer, D. J. (2003). Estimating Multilevel Linear Models as Structural Equation Models. *Journal of Educational and Behavioral Statistics*, 28(2), 135-167.
doi:10.3102/10769986028002135
- Bramlett, R. K., Scott, P., & Rowell, R. K. (2000). A comparison of temperament and social skills in predicting academic performance in first graders. *Special Services in the Schools*, 16(1-2), 147-158.
- CASEL. (2013). CASEL Guide :Effective Social and Emotional Learning Programs Secondary Edition. In S. Collaborative for Academic, and Emotional Learning (Ed.).
- Doctoroff, G. L., Fisher, P. H., Burrows, B. M., & Edman, M. T. (2016). Preschool children's interest, social–emotional skills, and emergent mathematics skills. *Psychology in the Schools*, 53(4), 390-403. doi:10.1002/pits.21912
- Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The Impact of Enhancing Students' Social and Emotional Learning: A Meta-Analysis of School-Based Universal Interventions. *Child Development*, 82(1), 405-432.
doi:10.1111/j.1467-8624.2010.01564.x
- Garner, P. W., Mahatmya, D., Brown, E. L., & Vesely, C. K. (2014). Promoting Desirable Outcomes Among Culturally and Ethnically Diverse Children in Social Emotional Learning Programs: a Multilevel Heuristic Model. *Educational Psychology Review*, 26(1), 165-189. doi:10.1007/s10648-014-9253-7
- Hair, E. H., Tamara, Terry-Humen, Elizabeth; Lavelle, Bridget; Calkins, Julia. (2006). Children's school readiness in the ECLS-K: Predictions to academic, health, and social outcomes in

- first grade. *Early Childhood Research Quarterly*, 21(4), 431-454.
doi:10.1016/j.ecresq.2006.09.005
- Jimerson, S., Egeland, B., & Teo, A. (1999). A longitudinal study of achievement trajectories: Factors associated with change. *Journal of Educational Psychology*, 91(1), 116-126.
doi:10.1037/0022-0663.91.1.116
- Kagan, S. L., Moore, E., & Bredekamp, S. (1998). *Reconsidering Children's Early Development and Learning Toward Common Views and Vocabulary: National Education Goals Panel*: DIANE Publishing.
- Kline, R. B. (2010). *Principles and Practice of Structural Equation Modeling, Third Edition*: Guilford Publications.
- Muthén, L. K., & Muthén, B. O. (2010). *Mplus: Statistical Analysis with Latent Variables: User's guide* (Version 7 ed.). Los Angeles, CA: Muthén & Muthén.
- Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, 7(2), 147-177. doi:10.1037/1082-989X.7.2.147
- Sklad, M., Diekstra, R., Ritter, M. D., Ben, J., & Gravesteyn, C. (2012). Effectiveness of school-based universal social, emotional, and behavioral programs: Do they enhance students' development in the area of skill, behavior, and adjustment? *Psychology in the Schools*, 49(9), 892-909. doi:10.1002/pits.21641
- Vrieze, S. I. (2012). Model selection and psychological theory: A discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). *Psychological Methods*, 17(2), 228-243. doi:10.1037/a0027127

Westat (2009). Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), Combined User’s Manual for the ECLS-K Eighth-Grade and K–8 Full Sample Data Files and Electronic Codebooks (NCES 2009–004). National Center fo

Appendix A
Input Instructions for Reading Growth Model 4

Title: ReadingIRT Profile Cluster Weights Model4
 DATA: FILE is "C:\ErinWinters\Updated Analyses\FinalDataTestR.dat";
 FORMAT IS FREE;
 Variable:
 NAMES ARE
 CHILDDID
 Weights
 School
 Profile
 Read1
 Read2
 Read3
 Read4
 Read5
 Read6
 Read7
 Math1
 Math2
 Math3
 Math4
 Math5
 Math6
 Math7
 SEH
 HR
 SER;
 USEVARIABLES ARE
 Read1 Read2 Read3 Read4 Read5 Read6 Read7 SEH HR SER;
 !Nominal are SEH HR SER;
 IDVARIABLE IS CHILDDID;
 CLUSTER IS School;
 Weight is WEIGHTS;
 MISSING ARE ALL (-99);
ANALYSIS:
TYPE IS COMPLEX;
ESTIMATOR IS MLR;
 OUTPUT: STANDARDIZED MODINDICES CINTERVAL SAMPSTAT TECH1 TECH4;
 SAVEDATA:
 FILE IS what;
 FORMAT IS F13.3;
 SAVE = FSCORES;
 RESULTS ARE Reading Growth Profile Cluster Weights Model4.dat;
 PLOT:
 TYPE IS PLOT1;

```
TYPE IS PLOT2;  
TYPE IS PLOT3;  
SERIES IS Read1(0) Read2(.5) Read3(1) Read4(1.5)  
Read5(3.5) Read6(5.5) Read7(8.5);  
MODEL:  
i s q | Read1@0 Read2@.5 Read3@1 Read4@1.5 Read5@3.5  
Read6@5.5 Read7@8.5;  
[i s q];  
Read7@0;  
i;  
s;  
q;  
i s q on SEH HR SER;
```

Appendix B
Input Instructions for Math Growth Model 4

Title: Math IRT Profile Cluster Weights Model4
 DATA: FILE is "C:\ErinWinters\Updated Analyses\FinalDataTestR.dat";
 FORMAT IS FREE;
 Variable:
 NAMES ARE
 CHILDDID
 Weights
 School
 Profile
 Read1
 Read2
 Read3
 Read4
 Read5
 Read6
 Read7
 Math1
 Math2
 Math3
 Math4
 Math5
 Math6
 Math7
 SEH
 HR
 SER;
 USEVARIABLES ARE
 Math1 Math2 Math3 Math4 Math5 Math6 Math7 SEH HR SER;
 !Nominal are SEH HR SER;
 IDVARIABLE IS CHILDDID;
 CLUSTER IS School;
 Weight is WEIGHTS;
 MISSING ARE ALL (-99);
ANALYSIS:
TYPE IS COMPLEX;
ESTIMATOR IS MLR;
 OUTPUT: STANDARDIZED MODINDICES CINTERVAL SAMPSTAT TECH1 TECH4;
 SAVEDATA:
 FILE IS what;
 FORMAT IS F13.3;
 SAVE = FSCORES;
 RESULTS ARE Math Growth Profile Cluster Weights Model4.dat;
 PLOT:
 TYPE IS PLOT1;

TYPE IS PLOT2;
TYPE IS PLOT3;
SERIES IS Math1(0) Math2(.5) Math3(1) Math4(1.5)
Math5(3.5) Math6(5.5) Math7(8.5);
MODEL:
i s q | Math1@0 Math2@.5 Math3@1 Math4@1.5 Math5@3.5
Math6@5.5 Math7@8.5;
[i s q];
Math7@0;
i;
s;
q;
i s q on SEH HR SER;

Table 1.

Descriptive Statistics and Zero-Order Correlations among Profiles and Outcomes

Measure	<i>M</i>	(<i>SD</i>)	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.
<i>Profiles</i>																			
1. Social-Emotional & Health Strengths (SHE)	0.21	(0.41)	--																
2. Social-Emotional Risk (SER)	0.14	(0.35)	-.21	--															
3. Health Risk (HR)	0.10	(0.30)	-.17	-.13	--														
<i>Reading IRT Scale Scores</i>																			
4. Reading Fall Gr K	35.49	(11.01)	-.33	-.26	-.15	--													
5. Reading Spring Gr K	46.97	(14.87)	-.28	-.23	-.16	.86	--												
6. Reading Fall Gr 1	53.85	(18.05)	-.31	-.22	-.15	.79	.90	--											
7. Reading Spring Gr 1	78.30	(24.37)	-.29	-.27	-.17	.70	.77	.84	--										
8. Reading Spring Gr 3	128.31	(27.54)	-.31	-.28	-.16	.55	.58	.62	.73	--									
9. Reading Spring Gr 5	152.48	(25.26)	-.28	-.27	-.16	.50	.53	.56	.67	.84	--								
10. Reading Spring Gr 8	173.20	(25.97)	-.25	-.22	-.19	.43	.43	.44	.53	.72	.76	--							
<i>Math IRT Scale Scores</i>																			
11. Math Fall Gr K	26.17	(9.20)	-.38	-.34	-.20	.74	.69	.67	.68	.63	.57	.51	--						
12. Math Spring Gr K	36.79	(12.35)	-.32	-.30	-.21	.66	.70	.68	.68	.63	.59	.53	.83	--					
13. Math Fall Gr 1	43.81	(14.02)	-.33	-.27	-.20	.61	.62	.69	.68	.65	.61	.53	.79	.83	--				
14. Math Spring Gr 1	62.05	(17.96)	-.28	-.29	-.20	.55	.58	.60	.67	.64	.61	.54	.71	.78	.82	--			
15. Math Spring Gr 3	99.59	(24.52)	-.29	-.30	-.18	.52	.55	.55	.64	.73	.69	.63	.69	.74	.78	.78	--		
16. Math Spring Gr 5	125.42	(24.04)	-.27	-.29	-.19	.48	.49	.51	.58	.67	.71	.66	.62	.67	.70	.72	.86	--	
17. Math Spring Gr 8	143.83	(20.97)	-.28	-.25	-.19	.45	.45	.44	.53	.64	.66	.72	.57	.61	.63	.65	.78	.84	--

Note. $N = 4,130$ students. For both the Reading and Math outcomes, different achievement measures were used throughout the study. IRT-based scale scores were developed to provide a continuous, longitudinal outcome. Comprehensive Positive (CP) profile is reference group across the three profiles (all dummy coded). All correlations are significant at the .001 level.

Table 2.

Model Fit Indices

	Number Parameters Estimated	Log Likelihood Value	Model Fit Indices				
			CFI	SRMR	BIC	AIC	BIC Difference
<i>Reading IRT Scale¹</i>							
<i>Latent Growth Curve Model²</i>							
Model 1	15	-78290	.79	.17	156714	156613	--
Model 2	24	-77363	.79	.14	154851	154775	1863
Model 3	24	-77363	.82	.14	154851	154775	0
Model 4	24	-115509	.75	.12	231141	231065	-76290
<i>Math IRT Scale¹</i>							
<i>Latent Growth Curve Model²</i>							
Model 1	15	-72948	.92	.08	145974	145926	--
Model 2	24	-71689	.93	.06	143502	143426	2472
Model 3	24	-71689	.94	.06	143502	143426	0
Model 4	24	-106106	.89	.05	212335	212259	-68833

¹For both the Reading and Math outcomes, different achievement measures were used throughout the study. IRT-based scale scores were developed to provide a continuous, longitudinal outcome.

²For both outcome models, variance of the final timepoint was constrained to 0 to avoid impermissible negative residuals.

Table 3.

Growth Model Results for Grade K-8 Reading¹

Parameters ²	Model 1: Growth Only				Model 2: Growth + Profiles				Model 3: Adjust for Schools				Model 4: Adjust for Weights			
	<i>Coeff</i>	<i>SE</i>	<i>Z</i>	<i>p</i>	<i>Coeff</i>	<i>SE</i>	<i>Z</i>	<i>p</i>	<i>Coeff</i>	<i>SE</i>	<i>Z</i>	<i>p</i>	<i>Coeff</i>	<i>SE</i>	<i>Z</i>	<i>p</i>
<i>Intercept (Baseline Score)</i>	35.45	0.21	172.21	<.001	40.93	0.27	151.48	<.001	40.93	0.33	125.42	<.001	40.04	1.98	20.23	<.001
Soc-Emo & Health Strengths (SEH)					-12.35	0.27	-45.28	<.001	-12.35	0.32	-38.48	<.001	-11.02	1.75	-6.31	<.001
Social-Emotional Risk (SER)					-12.61	0.29	-43.14	<.001	-12.61	0.34	-37.61	<.001	-12.09	1.66	-7.31	<.001
Health Risk (HR)					-10.29	0.43	-23.95	<.001	-10.29	0.45	-22.77	<.001	-9.35	1.89	-4.96	<.001
<i>Linear Slope (Growth per Year)</i>	30.18	0.18	165.14	<.001	33.40	0.21	156.51	<.001	33.40	0.27	124.47	<.001	33.26	0.74	44.90	<.001
Soc-Emo & Health Strengths (SEH)					-6.93	0.38	-18.35	<.001	-6.93	0.42	-16.37	<.001	-7.03	2.20	-3.20	.001
Social-Emotional Risk (SER)					-8.01	0.46	-17.49	<.001	-8.01	0.47	-17.10	<.001	-9.75	2.08	-4.70	<.001
Health Risk (HR)					-7.39	0.57	-13.05	<.001	-7.39	0.59	-12.57	<.001	-8.66	2.25	-3.85	<.001
<i>Quadratic Slope (Rate of Growth)</i>	-1.69	0.02	-80.37	<.001	-1.99	0.03	-80.21	<.001	-1.99	0.03	-64.23	<.001	-2.01	0.08	-25.01	<.001
Soc-Emo & Health Strengths (SEH)					0.64	0.04	14.61	<.001	0.64	0.05	13.18	<.001	0.74	0.23	3.26	.001
Social-Emotional Risk (SER)					0.75	0.05	14.11	<.001	0.75	0.05	13.81	<.001	1.02	0.23	4.52	<.001
Health Risk (HR)					0.64	0.07	9.65	<.001	0.64	0.07	9.27	<.001	0.82	0.20	4.03	<.001

¹ Reading Outcome= IRT Scaled Score of reading measures administered throughout study.

² Linear growth curve model in structural equation modeling framework used (MLR estimator): intercept (time0) = Fall Kindergarten, time1 = Spring Kindergarten, time2 = Fall 1st Grade, time3=Spring 1st Grade, time4 = Spring 3rd Grade, time5= Spring 5th grade, time6= Spring 8th Grade; *Mplus 7.1* used to model growth. Comprehensive Positive (CP) profile set as reference group for profile predictions (each of other profiles dummy coded).

Table 4.

Growth Model Results for Grade K-8 Math¹

Parameters ²	Model 1: Growth Only				Model 2: Growth + Profiles				Model 3: Adjust for Schools				Model 4: Adjust for Weights			
	<i>Coeff</i>	<i>SE</i>	<i>Z</i>	<i>p</i>	<i>Coeff</i>	<i>SE</i>	<i>Z</i>	<i>p</i>	<i>Coeff</i>	<i>SE</i>	<i>Z</i>	<i>p</i>	<i>Coeff</i>	<i>SE</i>	<i>Z</i>	<i>p</i>
<i>Intercept (Baseline Score)</i>	25.81	0.15	176.15	<.001	31.32	0.18	175.56	<.001	31.32	0.23	135.34	<.001	30.14	0.59	51.23	<.001
Soc-Emo & Health Strengths (SEH)					-12.06	0.22	-55.00	<.001	-12.06	0.26	-46.03	<.001	-11.24	0.98	-11.51	<.001
Social-Emotional Risk (SER)					-13.09	0.23	-56.23	<.001	-13.09	0.27	-47.86	<.001	-12.33	1.06	-11.60	<.001
Health Risk (HR)					-11.01	0.33	-33.67	<.001	-11.01	0.36	-30.45	<.001	-9.02	1.86	-4.85	<.001
<i>Linear Slope (Growth per Year)</i>	24.99	0.14	184.76	<.001	27.59	0.16	174.15	<.001	27.59	0.20	138.27	<.001	27.87	0.53	52.58	<.001
Soc-Emo & Health Strengths (SEH)					-5.11	0.32	-16.24	<.001	-5.11	0.35	-14.56	<.001	-4.61	1.39	-3.33	.001
Social-Emotional Risk (SER)					-6.74	0.37	-18.22	<.001	-6.74	0.37	-18.17	<.001	-6.61	1.63	-4.07	<.001
Health Risk (HR)					-6.50	0.48	-13.41	<.001	-6.50	0.49	-13.32	<.001	-5.26	2.93	-1.80	.072
<i>Quadratic Slope (Rate of Growth)</i>	-1.35	0.01	-93.84	<.001	-1.59	0.02	-92.22	<.001	-1.59	0.02	-73.46	<.001	-1.65	0.06	-29.70	<.001
Soc-Emo & Health Strengths (SEH)					0.47	0.03	13.95	<.001	0.47	0.04	12.76	<.001	0.46	0.11	4.06	<.001
Social-Emotional Risk (SER)					0.65	0.04	16.70	<.001	0.65	0.04	16.23	<.001	0.64	0.18	3.68	<.001
Health Risk (HR)					0.61	0.05	11.92	<.001	0.61	0.05	11.91	<.001	0.47	0.26	1.83	.068

¹ Math Outcome= IRT Scaled Score of reading measures administered throughout study.

² Linear growth curve model in structural equation modeling framework used (MLR estimator): intercept (time0) = Fall Kindergarten, time1 = Spring Kindergarten, time2 = Fall 1st Grade, time3=Spring 1st Grade, time4 = Spring 3rd Grade, time5= Spring 5th grade, time6= Spring 8th Grade; *Mplus 7.1* used to model growth. Comprehensive Positive (CP) profile set as reference group for profile predictions (each of other profiles dummy coded).

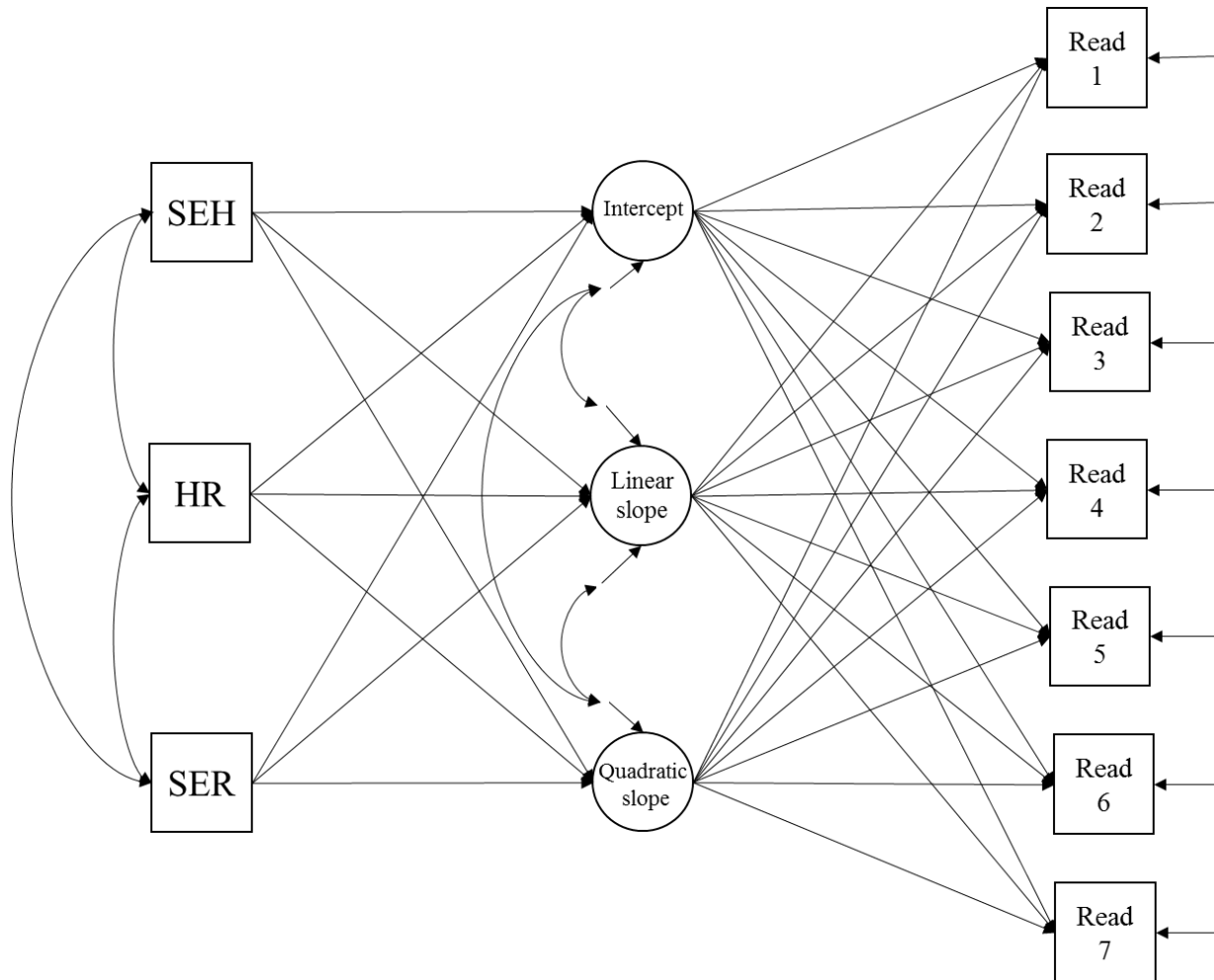


Figure 1. *Path Diagram of Model 2 (Profile Included) for Reading*

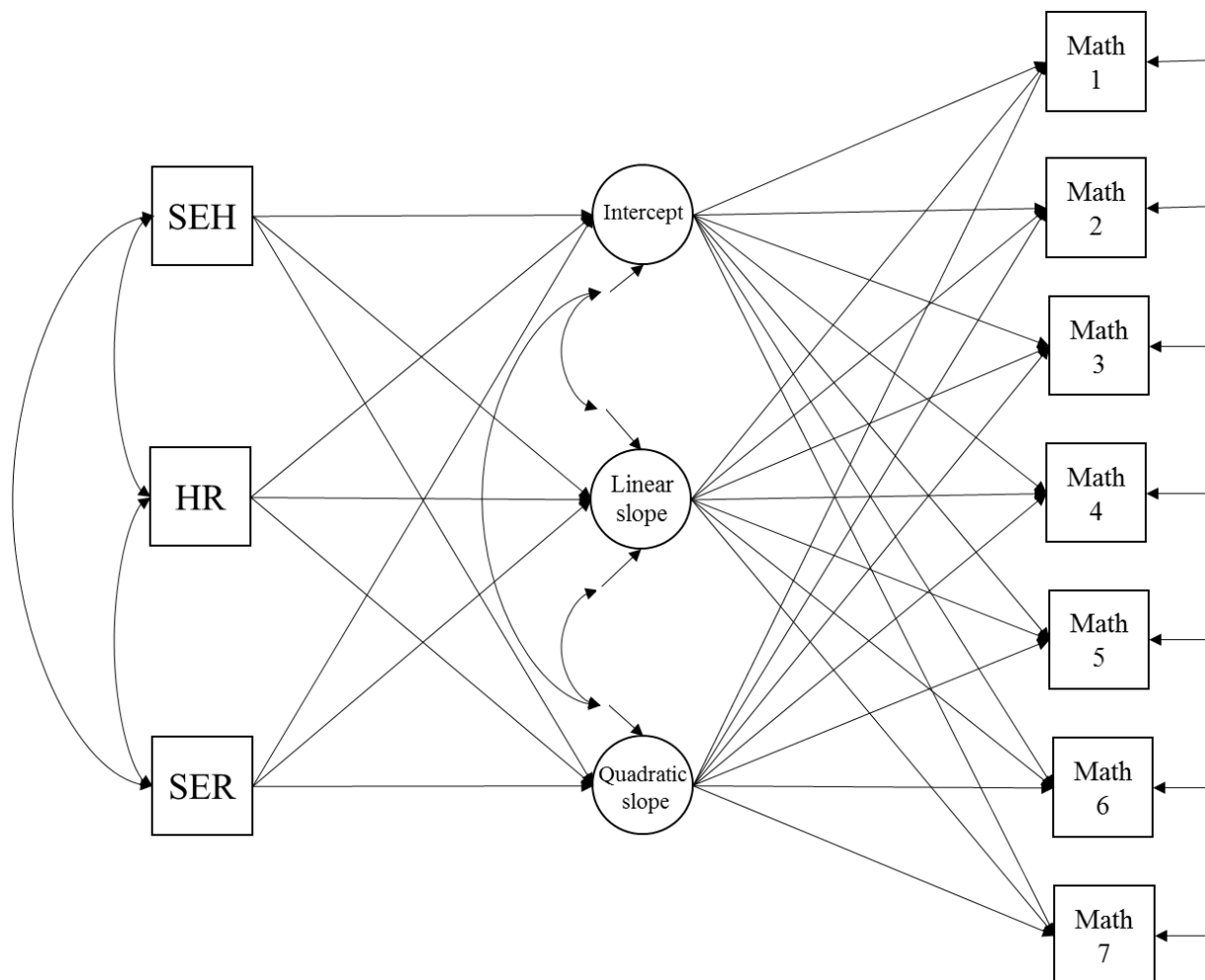
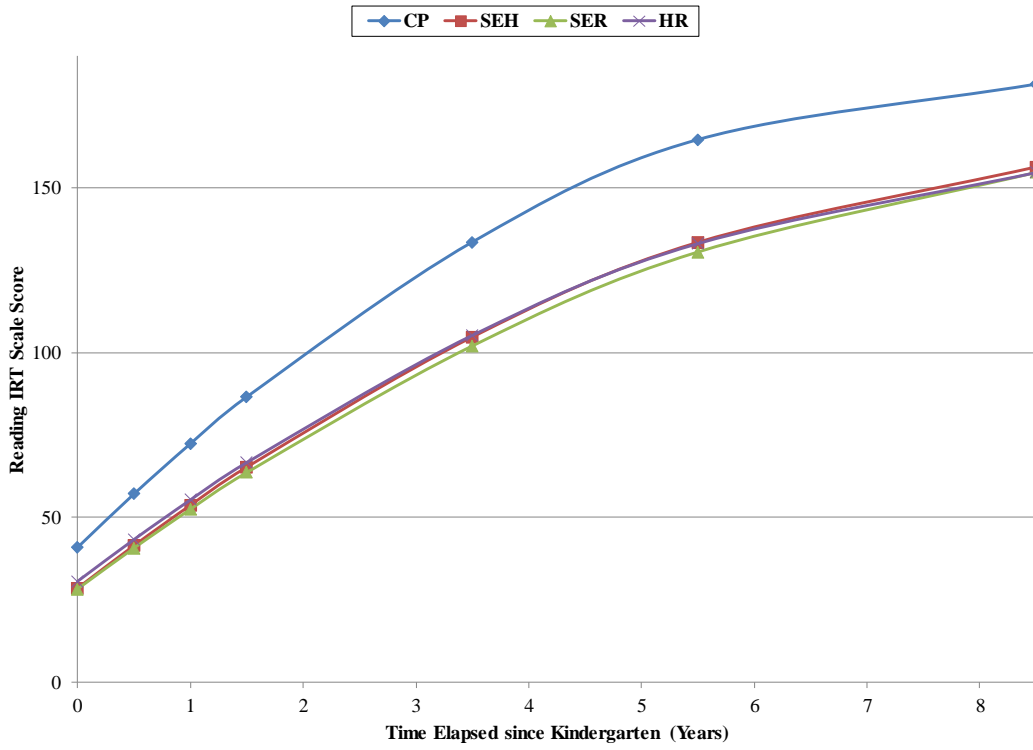


Figure 2. Path Diagram of Model 2 (Profile Included) for Math

Panel A



Panel B

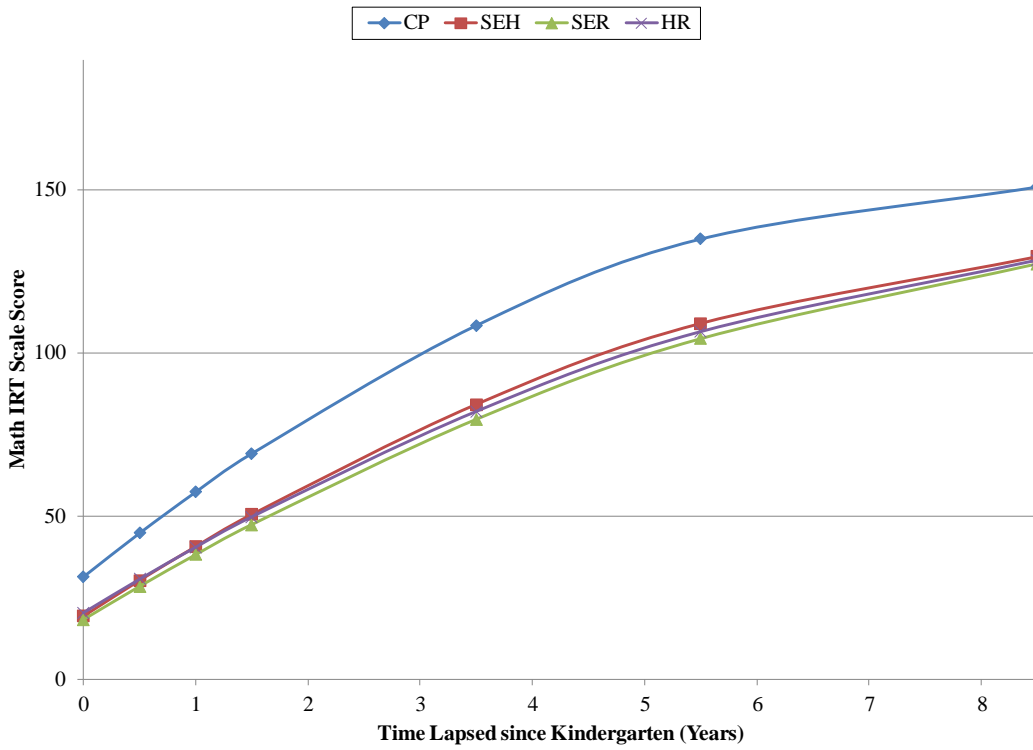


Figure 3. Predicted Growth for Profiles (Model 2) for Reading (Panel A) and Math (Panel B)