WOMEN IN SCIENCE: DISAGGREGATING BETWEEN- AND WITHIN-PERSON EFFECTS USING A NOVEL MODELING FRAMEWORK

Chelsea Burfeind

A thesis submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Master of Arts in the Department of Psychology (Quantitative Psychology)

Chapel Hill
2013

Approved by:
Patrick Curran
Daniel Bauer
Abigail Panter
ABSTRACT

CHELSEA BURFEIND: Women in Science: Disaggregating Between- and Within-Person Effects Using a Novel Modeling Framework (Under the direction of Patrick Curran)

Extant research indicates that females are less likely than males to pursue a degree in the fields of science, technology, engineering, and math (STEM), despite equal achievement in the subjects. This project aimed to examine the longitudinal relationship between interest and achievement in math in adolescents while implementing a novel modeling framework. The residualized autoregressive latent trajectory (RALT) was implemented to completely disaggregate between-and within-person effects, and a multiple-group model was utilized to test for gender differences in the RALT parameters. Males and females were shown to differ at the between-person level, with females reporting higher levels of math achievement and lower levels of interest in math than males. Within-person effects were not shown to vary systematically by gender. This project provided a thorough analysis of a substantive topic through the use of a new modeling framework.
# TABLE OF CONTENTS

**LIST OF TABLES**

Page vi

**LIST OF FIGURES**

Page vii

**CHAPTER 1: BACKGROUND AND PRIOR RESEARCH**

Page 1

1. Introduction

Page 1

1.1 Women in Science

Page 2

1.1.1 Prior research

Page 2

1.1.2 Methodological concerns

Page 4

1.2 The Longitudinal Study of American Youth (LSAY)

Page 5

1.3 Latent Curve Models

Page 7

1.3.1 Univariate models

Page 7

1.3.2 Bivariate models

Page 11

1.4 Between- and Within-Person Effects

Page 13

1.4.1 Disaggregating between- and within-person effects

Page 14

1.5 The Residualized Autoregressive Latent Trajectory (RALT) Model

Page 17

1.6 Cross-Classified Data Structures

Page 21

**CHAPTER 2: METHODS**

Page 23

2.1 Sample

Page 23

2.2 Measures

Page 23

2.2.1 Gender

Page 23

2.2.2 Interest in math

Page 24

2.2.3 Math achievement

Page 24

2.3 Patterns of Missing Data

Page 25

2.4 Intra-Class Correlation Coefficients (ICCs)

Page 25

2.5 Longitudinal Trends

Page 25
# LIST OF TABLES

1. Time-specific correlations, means, and standard deviations for achievement and interest in math

2. Univariate latent growth model of math achievement

3. Univariate RALT model for math achievement

4. Univariate latent growth model for interest

5. Univariate RALT model for interest in math

6. Means and regressions for an unconditional and bivariate RALT

7. Variance, covariances, and correlations among the growth parameters in bivariate RALT

8. Means and regression parameters for bivariate RALT model of math interest and achievement by gender

9. Variances, covariances, and correlations for a bivariate RALT model for males

10. Variances, covariances, and correlations for a bivariate RALT model for females

11a. Likelihood ratio tests of multiple-group analysis of a bivariate RALT for males and females

11b. Likelihood ratio difference tests between nested constraints of multiple-group analysis for bivariate RALT
LIST OF FIGURES

1. Unconditional, univariate, and linear latent growth curve model.................................8

2. Conditional univariate latent growth curve model with two time-invariant covariates...........10

3. Conditional univariate latent growth curve model with one time-varying covariate..........11

4. Unconditional, bivariate, and linear latent growth curve model.......................................12

5. Univariate latent growth curve model with a time-varying covariate and an intercept factor on the TVC to disaggregate between- and within-person effects. .................................................................................................................16

6. Unconditional, univariate, and linear residualized autoregressive latent trajectory model.....18

7. Bivariate and unconditional residualized autoregressive latent trajectory model...............20

8. Math achievement over time as a function of gender.........................................................27

9. Interest in math over time as a function of gender..............................................................28

10. Model-implied and sample means for the longitudinal trajectory of math achievement......34

11. Model-implied and sample means for the longitudinal trajectory of interest in math..........37

12. Model-implied trajectories for interest in math over time as a function of gender..............43

13. Model-implied trajectories for math achievement over time as a function of gender.........44

14. Normal probability and QQ plots for interest in math at each time point..........................47

15. Normal probability and QQ plots for math achievement at each time point......................48

16. Residual values versus predicted values for interest in math.............................................49

17. Residual values versus predicted values for math achievement........................................50

18. Residuals versus time plotted for four random subsamples of 100 subjects for math interest.................................................................................................................................51

19. Residuals versus time plotted for four random subsamples of 100 subjects for math achievement.................................................................................................................................52
CHAPTER 1: BACKGROUND AND PRIOR RESEARCH

Introduction

The latent growth curve (LGC) modeling framework encompasses a broad and flexible class of models that is widely-used by researchers. Because the vast majority of psychological research is longitudinal and multivariate in nature, it is often of interest to capture bidirectional relations between two or more variables over time, and the LGC provides a relatively straightforward way of doing just that. However, an additional modeling issue concerns the between- versus within-person effect distinction. Curran and Bauer (2011) noted that, though the preponderance of hypothesized theories concern within-person processes, the bulk of available methodologies for analyzing such theories deals with between-person processes.

The LGC can partially disaggregate between- and within-person effects, as will be shown in subsequent sections of this paper. It will be shown that it is indeed possible to fully parse apart the effects, but the methods of doing so are in their infancy and not widely known at this time. Additionally, further expansion is needed in order to completely disaggregate the between- and within-person effects when bidirectional relations between variables are theorized. This paper aims to expand on these methods in order to effectively analyze a substantive data set in a methodologically rigorous manner.

The introduction section will be set up as follows: the substantive issue at hand (women in science) will be established and explored, and the dataset to be analyzed will be introduced. Existing methods will be drawn from to create a model that is best suited to analyze the substantive hypotheses, and such methods will be utilized in the creation of a novel approach to analyze the data. The overarching goal is to implement a rigorous and empirical test of novel hypotheses through the use of a
newly developed method not yet applied in psychological research. The substantive issue that follows is
one that is both of theoretical interest to myself and many other researchers and is one that poses
several methodological issues in terms of the best way to analyze it.

1.1 Women in Science

The issue of women in the traditionally male-dominated fields of science, technology,
ingi eering, and mathematics (STEM fields) is one that puzzles substantive researchers and poses several
interesting analytic challenges. In addition to being an interesting and complex substantive issue, to date,
no existing analyses have fully dealt with and incorporated the slew of modeling issues that comes with
the complexity of the data. The following section introduces both the substantive and the methodological
issues that face researchers in the area of women and science.

1.1.1 Prior Research

The issue of women in the workplace has been the subject of much research and debate. Recently,
such research has focused on women's participation in the traditionally male fields of science, technology,
ingineering, and mathematics (STEM fields). The American Association of University Women (AAUW)
describes the situation: though women have made great strides in terms of rates of participation and
degrees earned in fields such as the social sciences, the arts, and the humanities, these strides are not as
apparent in STEM fields. Despite the fact that in elementary, middle, and high school, girls and boys take
roughly equal numbers of courses in math and science, and the fact that girls do as well as boys (Hyde et
al., 2008) women earn only 20% of Bachelor's degrees in the STEM fields. This number drops at the
graduate level and drops yet again at the transition to the workplace. In addition, girls report less interest
in STEM careers starting at a very early age, and this trend continues into adulthood (Lapan et al., 2000;
Turner et al., 2008).
In a report entitled “Why So Few?”, the AAUW sets out to determine why girls and women are so vastly underrepresented in STEM fields and why they report less interest than boys early on despite equal performance. Two reasons as to why girls are less enthusiastic about science and math came to fruition from the research highlighted in the report: first of all, the stereotype that men are superior when it comes to STEM fields continues to be highly prevalent, and women are susceptible to stereotype threat even at an implicit level. Secondly, girls’ lack of interest in STEM fields plays a large part in their going on to get degrees in other fields.

Though no IQ differences have been found between men and women with regards to STEM subjects, the stereotype that men are innately superior to women when it comes to science and math is still highly prevalent. Though men and women do differ in their strengths and weaknesses – men tend to be better at spatial orientation and visualization, and women tend to be better when it comes to verbal skills and memory - no evidence suggests that these differences are egregious enough to make one sex largely more competent at any given field (Lynn and Irving, 2004). Despite these findings, children are aware of the stereotypes and begin to express stereotypical beliefs as early as elementary school (Farenga and Joyce, 1999; Ambady et al., 2001). These negative stereotypes affect women’s and girl’s performance and aspirations in math and science through what is commonly known as “stereotype threat”, where a person purposefully avoids certain situations or activities to avoid fulfilling a given stereotype. Researchers posit that girls may attempt to reduce the likelihood that they will be judged negatively by saying they are not interested in and avoiding STEM fields (Steele and Aronson, 1995).

Girls also report less interest in pursuing STEM careers at very early ages. Starting from early adolescence, girls express less interest in math and science than boys do (Lapan et al., 2000; Turner et al., 2008) and in a survey conducted by WBGH, 24% of boys reported interest in an engineering career, compared to only 5% of girls. Even women and girls who excel in math and science often do not pursue
STEM fields in favor of degrees in the humanities, life sciences, and social sciences. The reverse is true for men (Lubinski and Benbow, 2006).

It can thus be seen that girls face a myriad of issues growing up with regards to interest and achievement in the STEM fields. Though prior research indicates that, in theory, individuals who are higher-achieving in given domains should express more motivation to pursue interests within that domain, it does not appear to be the case when it comes to women and math and science. This paper will explore the relationship between math achievement and interest in greater detail at both the between- and the within-person levels. Gender differences will be explored in order to further the understanding of exactly what is happening to females during adolescence that shapes them to be different from males, and shapes them to have less desire to pursue STEM fields than males.

1.1.2 Methodological concerns

With the aforementioned substantive concerns comes the issue of constructing a model that best corresponds to the theory at hand. Prior research paints an interesting picture of the issues females face while going through school in terms of their achievement in and motivation to pursue STEM fields. However, this research largely focuses on these processes on an average, or between-persons, level. While helpful in establishing what is happening at an overall level, this is only half of the picture. It is also extremely important to explore within-person processes in order to determine what is happening on the individual level. As will be explored in greater detail in later sections, cases can arise when between-person processes are misattributed as within-person processes due to the failure to completely disentangle the two sources of variability, which can lead to errors of inferences.

Additionally, when modeling constructs in school-age individuals over a period of many years, individuals normally experience change in their school environments. Typically, students transition from elementary to middle school, and then from middle school to high school. When nesting within schools in
considered, thus poses a further challenge, as typical methods of modeling hierarchical data that assume a single upper level of nesting are not fully sufficient. This also needs to be taken into consideration when dealing with such data structures. The following section describes the data set to be used in analyses.

1.2 The Longitudinal Study of American Youth (LSAY)

The data used here comes from the Longitudinal Study of American Youth (LSAY), which was a two-cohort longitudinal study conducted from 1987-1994. The nationally representative sample consisted of approximately 5,000 adolescents in 7th and 10th grade. The LSAY was originally designed to study the factors related to the development of interest and skills in math and science, which corresponds directly to my substantive questions. The younger cohort is utilized here, which consists of roughly 3,500 7th graders recruited from feeder schools. Beginning in the fall of 7th grade, students, their parents, and their teachers were assessed on a myriad of factors. Students were given achievement tests in math and science during the fall of each year, and scores were recorded. Questionnaires were given to assess students’ opinions on their interest, self-confidence, and attitudes toward math and science. Additionally, a rich body of information was gathered to assess many aspects of the students’ lives: friends, stereotypes, extracurricular activities, and many other variables, and a large number of school-level covariates are available that assess school atmosphere and emphasis on various aspects of academic and extracurricular activities.

In the following analyses, a total of six time points will be utilized, representing the fall of every school year starting in 7th grade and continuing through the students’ final year of high school. Students switched school between their 8th and 9th grade years to transition to high school, which needs to be accounted for in analyses. Drawing on the aforementioned theories and questions posed by the issue of women in STEM fields, the students’ interest and achievement in math will be modeled across time. Self-confidence in math, gender stereotypes, race, and sex will all be incorporated to examine their effects on
the relationship between the two constructs over time. Between- and within-person effects will be disaggregated to explore processes that are going on at each level.

Modeling such a data structure to match the questions of interest is complicated, and there is no existing straightforward method of doing so. Two major issues arise when considering the LSAY data set: the first is the modeling of relations between two constructs over time. Several modeling frameworks are in existence that focus on bivariate relations over time: namely, the autoregressive latent trajectory model (ALT) and the latent difference score model (LDS). The ALT model, first proposed by Curran and Bollen (2001), combines aspects of a latent growth curve model and a standard autoregressive model to simultaneously model growth processes and autoregressions among manifest variables. The LDS model, formulated by McArdle (2001), models the dynamics between variables over time; that is, it models how change in one variable can predict future change in itself and in other variables. Though both models are useful in certain contexts, neither is quite right in this case. The ALT model does not completely disaggregate between- and within-person effects and the LDS’ focus on dynamic relations between variables does not exactly parallel the theories described above.

It is thus necessary to create a new modeling framework that is consistent with theory and which can account for the multivariate, longitudinal, and cross-classified nature of this particular data set while successfully disaggregating between- and within-person effects. Drawing from a small body of extant literature, this paper will start with the most basic univariate and unconditional latent growth curve model and build to a model that incorporates all features of the data.
1.3 Latent Curve Models

1.3.1 Univariate models

The latent curve modeling framework (LCM) is a versatile and broad class of models that uses repeated measures to estimate a single underlying growth trajectory for each person across all time points. The equation for a linear univariate, unconditional LCM is

\[ y_{it} = \alpha_i + \lambda t + \beta_i t + \epsilon_{it} \]  

where \( \alpha_i \) and \( \beta_i \) are the random intercept and random slope for case \( i \), respectively, and \( \epsilon_{it} \) represents error. In the case of a linear trajectory model, \( \lambda_i = t - 1 \) for all \( t \). It is assumed that \( E(\epsilon_{it}) = 0 \) for all \( i \) and \( t \), \( COV(\epsilon_{it}, \epsilon_{is}) = 0 \) for all \( i \) and \( t \), \( E(\epsilon_{it}, \epsilon_{st}) = 0 \) for all \( t \) and \( i \neq s \), and \( E(\epsilon_{it}, \epsilon_{st}) = \sigma^2_{\epsilon_{it}} \) for each \( t \). Lastly, it is assumed that \( COV(\epsilon_{it}, \epsilon_{i,t+k}) = 0 \) for \( k \neq 0 \) so that the errors are not correlated over time, though it is possible to model autoregressive errors and that this assumption will in fact be relaxed later in the text (Curran and Bollen, 2001).

The LCM allows each individual to have a unique intercept and slope, which is expressed in the model equations by indexing each parameter with the subscript \( i \). The mean intercept and slope terms across all individuals can be expressed as

\[ \alpha_i = \mu_{\alpha} + \zeta_{\alpha i} \]  

\[ \beta_i = \mu_{\beta} + \zeta_{\beta i} \]

where, in each case, \( \mu \) represents the mean intercept and slope across all cases. The \( \zeta \) parameters represent random variability around the mean intercept and slope with means of zero and uncorrelated with \( \epsilon_{it} \). It is assumed that \( E(\zeta_{\alpha i}) = 0 \) and \( E(\zeta_{\beta i}) = 0 \), and \( COV(\zeta_{\alpha i}, \zeta_{\beta i}) = \psi_{\alpha\beta} \). Additionally,
\[ \text{COV}(\zeta_{\alpha}, \zeta_{\alpha}) = \psi_{\alpha\alpha} \] and \[ \text{COV}(\zeta_{\beta}, \zeta_{\beta}) = \psi_{\beta\beta}. \] Lastly, \[ \text{COV}(\zeta_{\alpha}, \epsilon_{it}) = \text{COV}(\zeta_{\beta}, \epsilon_{it}) = 0. \] Figure 1 is a path diagram for this model.

**Figure 1.** Unconditional univariate, and linear latent growth curve model.

The unconditional LCM describes the trajectory of a single repeated measure and summarizes the patterns of change in repeated measures in terms of the latent growth parameters. It is often useful to explore predictors of such parameters, and predictors of the repeated measures themselves, in order to see how these affect the trajectories. It is possible to add two types of covariates to an LCM: time-invariant covariates (TICs) whose values remain constant over the entire study duration (e.g., gender, race), and the slope and intercept terms are regressed on them, and time-varying covariates (TVCs) whose values can change over time (such as a person's day-to-day emotional state, or the number of hours of sleep each night), and the repeated measures themselves are regressed on the covariates. Once covariates are added to the model, it is called a conditional LCM.
Consider first the case of a univariate conditional LCM with two TICs. The level 1 equation is same as that for the unconditional LCM shown in equation 1, and the assumptions remain the same as well. The unconditional differs from the conditional in the level 2 equations, however,

\[ \alpha_i = \mu_\alpha + \gamma_{\alpha_1} x_{i1} + \gamma_{\alpha_2} x_{i2} + \xi_\alpha \]  

(4)

\[ \beta_i = \mu_\beta + \gamma_{\beta_1} x_{i1} + \gamma_{\beta_2} x_{i2} + \xi_\beta \]  

(5)

Where \( \mu_\alpha \) and \( \mu_\beta \) are the intercepts of the equations that predict the random intercepts and random slopes. The \( x_{i1} \) and \( x_{i2} \) are the two covariates that predict the random intercepts and slopes, and the \( \gamma \) variables are the covariate coefficients for the random slope and intercept equations. The interpretation of the \( \gamma \) coefficients is analogous to that of a regression equation; they give the expected change in the outcome variable given a one-unit change in the predictor variable, holding all other predictors constant.

Here the same assumptions hold as those given for the univariate and unconditional LCM, with the additional assumption in effect that the error terms for the latent slope and intercept factors are uncorrelated with the TIC. It should be noted that the intercept and slope variances are no longer variances of the random intercepts and random slopes themselves, but are now conditional variances on the covariates. A path diagram for the above model is shown in Figure 2.
Additionally, it is possible to incorporate TVCs, on which the measures themselves are typically regressed. Consider a model with one time-varying covariate, denoted $z_{it}$. In this case, the predictor is incorporated into the model at the first level. The equation for this is

$$y_{it} = \alpha_i + \beta_i \lambda_i + \gamma_i z_{it} + \epsilon_{it}$$

(6)

Where $\gamma_i$ represents the time-specific influence of the covariate on the repeated measures. All of the same assumptions hold that were given for the unconditional model, and additionally it is assumed that the error terms are uncorrelated with the predictors. The level 2 equations remain the same as those given in equations 2 and 3, and a path diagram for the univariate TIC model is shown in Figure 3.

---

**Figure 2.** Conditional univariate latent growth curve model with two time-invariant covariates.
1.3.2 Bivariate Models

It is a straightforward process to extend the univariate LCM to the bivariate case and essentially involves the simultaneous estimation of two univariate LCMs. The level 1 equations for an unconditional bivariate LCM are

\[ y_{it} = \alpha_{yi} + \lambda_i \beta_{yi} + \varepsilon_{yit} \]  
\[ x_{it} = \alpha_{xi} + \lambda_i \beta_{xi} + \varepsilon_{xit} \]

The assumptions are equivalent to those of the univariate LCM, with the addition of the assumption that the errors for each equation are uncorrelated with the random slopes and random intercepts for each variable. The equations for the random slopes and intercepts for the \(y\) variable are

Figure 3. Conditional univariate latent growth curve model with one time-varying covariate.
\[ \alpha_{yi} = \mu_{y\alpha} + \zeta_{y\alphai} \]  \hspace{1cm} (9)

\[ \beta_{yi} = \mu_{y\beta} + \zeta_{y\betai} \]  \hspace{1cm} (10)

and the equations for those for the x variable are

\[ \alpha_{xi} = \mu_{x\alpha} + \zeta_{x\alphai} \]  \hspace{1cm} (11)

\[ \beta_{xi} = \mu_{x\beta} + \zeta_{x\betai} \]  \hspace{1cm} (12)

where the disturbances are assumed to have means of zero. A path diagram for this model is shown in Figure 4.  

*Figure 4. Unconditional, bivariate, and linear latent growth curve model.*

Once again, it is often useful to consider predictors in the case of the bivariate LCM and the addition of such covariates is a process very similar to that described above in the univariate. As before, in the case of a TIC, the level 1 equations for the bivariate model remain the same as those given in Equations 7 and 8. The predictors are incorporated into the level 2 equations,
\[ \alpha_{iy} = \mu_{\alpha y} + \gamma_{q\alpha i} x_{iq} + \zeta_{\alpha y i} \] (13)

\[ \beta_{iy} = \mu_{\beta y} + \gamma_{q\beta i} x_{iq} + \zeta_{\beta y i} \] (14)

And

\[ \alpha_{ix} = \mu_{\alpha x} + \gamma_{q\alpha i} x_{iq} + \zeta_{\alpha x i} \] (15)

\[ \beta_{ix} = \mu_{\beta x} + \gamma_{q\beta i} x_{iq} + \zeta_{\beta x i} \] (16)

All previously stated assumptions hold, with additional assumptions that the covariates for each variable are uncorrelated with the intercept, slope, and error terms of the other variable.

1.4 Between- and Within-Person Effects

An important issue in both substantive and quantitative literature alike is that of the between-and within-person effects distinction. The vast majority of psychological research focuses on within-person processes; for example, the relationship between exercise and positive affect (Pendo and Dahn, 2005). Curran and Bauer (2011) note that while most studies posit within-person processes, the majority of research conducted to evaluate such processes often involves the collection and analysis of between-person data. They point out the need to disentangle between- and within-person effects, as errors of inference can arise from failing to do so. Most commonly, group-level results are misattributed to individuals in what is called the ecological fallacy (Robinson, 1950). If the direction and magnitude of the between- and within-person effects is the same, then errors in interpretation are less likely. Unfortunately, between- and within-person processes can operate simultaneously and in opposite directions, leading to a probable misattribution of the effects.
Disentangling between- and within-person effects begins with TVCs, which represent a mean shift in the repeated measures themselves as a function of the values of the TVCs within each individual. A single TVC can vary both between individuals (on an overall average level), and within individuals (across time). Though methods of disaggregating such effects are well-established and relatively straightforward in the multilevel model (MLM) framework, it has been shown that these methods do not work within the SEM framework and that other novel methods must be utilized (Curran et al., in press).

1.4.1 Disaggregating between- and within-person effects

Within the MLM, the between- and within- person effects can simultaneously be parsed apart by incorporating the person-centered TVC as a level 1 predictor and the person mean as a level 2 predictor. The main effect of the person-mean-centered TVC at level 1 represents the within person effect; the incorporation of the person mean at level 2 represents a direct estimate of the between-person effect (Raudenbush & Bryk, 2002). Unfortunately, this straightforward method does not apply within the SEM framework. Curran et al. (in press) points out that the SEM is not estimable under standard maximum likelihood when using a person-mean-centered TVC because TVCs are ipsative, resulting in a singular sample covariance matrix. An ipsative measure is defined when the sum of a set of items for a given individual is the same for all individuals in a sample. Because the expected value of a person-mean-centered TVC is zero for all individuals, it is said to be ipsative. Additionally, it is not possible to obtain a between-effect in the SEM through the use of the person-mean as a predictor due to the linear dependency between the person-mean and the TVCs in the covariance matrix.

Up until recently, no models were developed that were able to fully disaggregate between- and within-person effects. Though the inclusion of a TVC within the SEM provides an accurate estimate of the within-person effect, there is no information available about the between-person effect. Fortunately, recent research has resulted in methods by which the between- and within-person effects can be
disaggregated within the SEM framework. Curran et al. (in press) starts by highlighting the importance of the covariance structure between the TVCs and the random intercept. Because this structure enters directly in the model formulation, the random intercept effect is estimated net the TVC effects, and the TVC effects are estimated net the effects of the random intercept. Thus, the TVCs capture the within-person effect because they are unique to the influence of the random intercept. The saturated covariance structure between the TVCs and the random intercept “absorbs” the omitted predictor.

Within the SEM, it is possible to disentangle the between- and within-person effects through the use of a latent variable methodology to explicitly incorporate information about within-person sampling variability. Within the MLM, the person-mean used to predict the random intercept is an exogenous manifest measure and is assumed to be error-free (Raudenbush & Bryk, 2002). Since this cannot be utilized in the SEM, it cannot be assumed to be error free. Fortunately, it is possible to model measurement error within the SEM through the use of multiple indicator latent factors. For a given TVC, the total variance can be broken down into three components: a mean, the true score variance and residual variance. A TVC can then be expressed as such,

\[ z_i = \mu_{\alpha_i} + \zeta_{\alpha_i} + \epsilon_{\alpha_i} \]  \hspace{1cm} (17)

Where \( \mu_{\alpha_i} \) is a latent random intercept for the TVC, \( \zeta_{\alpha_i} \) represents the true score variance, and \( \epsilon_{\alpha_i} \) represents residual, or unexplained, variance. It is assumed that \( E(z_i) = \mu_{\alpha_i} \), \( \text{var}(z_i) = \text{var}(\zeta_{\alpha_i} + \epsilon_{\alpha_i}) = \psi_{\alpha_i\alpha_i} + \sigma^2_{\epsilon} \). The variance of the residual variance (\( \epsilon_{\alpha_i} \)) is thus \( \sigma^2_{\epsilon} \), which is an estimate of within-person residual variance.

To incorporate this between-person residual variability into the model, a random intercept latent factor for the TVC (as seen in equation 17) is defined. The random intercept is then used as a predictor of
the random intercept and random slope of the outcome $y$. Additionally, because the goal is to obtain simultaneous estimates of the between- and within-person effects, the regression of $y_{it}$ on $z_{it}$ must be retained. However, doing so results in an aggregate effect of the TVC on the outcome. To circumvent this, $y_{it}$ will be regressed on the residuals of $z_{it}$, specifically $\varepsilon_{z_{it}}$, which is the within-person residual variance. The motivation for this is the fact that the time-specific residuals represent the deviation of each observation from the true mean of the TVC. The final path diagram for this is shown in Figure 5.

![Path Diagram](image)

*Figure 5. Univariate latent growth curve model with a time-varying covariate and an intercept factor on the TVC to disaggregate between- and within-person effects.*
It can thus be seen that this parameterization cleanly and unambiguously disaggregates the between- and within-person effects. The between-effects are represented by the regression of the random slope and random intercepts of the outcome variable on the latent intercept for the TVC (denoted $\gamma_\alpha$ and $\gamma_\beta$ in Figure 5). The within-effects are represented by the regression of the repeated measures on the residual values of the TVC (denoted $\gamma$ in Figure 5).

1.5 The Residualized Autoregressive Latent Trajectory (RALT) Model

Using the ideas highlighted above, I aimed to extend the concepts utilized to a model which includes bidirectional relations between variables, which will be denoted the residualized autoregressive latent trajectory (RALT) model. Though not explicitly a TVC model, it will draw on the between- and within-effects distinction that was made clear in the use of the TVC model to disaggregate the effects. The model is more commonly known as an LCM with autoregressive disturbances and was first introduced by Chi and Reinsel (1989). A path diagram for a univariate, unconditional RALT model is shown below in Figure 6.
Figure 6. Unconditional, univariate, and linear residualized autoregressive latent trajectory model.

The corresponding equations are

\[ y_{it} = \alpha_i + \lambda_i \beta_i + \delta_{it} \]  \hspace{1cm} (18)

\[ \delta_{it} = \rho \delta_{i,t-1} + \epsilon_{\delta_{it}} \]  \hspace{1cm} (19)

where \( \alpha \) represents the random intercept, \( \beta \) represents the random slope, \( \delta \) is the error term, \( \rho \) is the effect of the previous disturbance on the current one, and \( \epsilon_{\delta_{it}} \) is the residual term. Because it is necessary to further break down the residual variance terms in subsequent equations, the level-1 error terms are now represented by \( \delta \). In keeping with the between- and within-person effect distinction, both the random slope and random intercept term consist of a mean plus individual variation around the mean, as shown in equations 9 and 10 in the case of the bivariate LCM described above. The same assumptions hold as in the LCM, with the additional assumptions that \( COV(\epsilon_{\delta_{it}}, \alpha_i) = 0 \) and \( COV(\epsilon_{\delta_{it}}, \beta_i) = 0 \). Here, the between-person effects are represented by the mean levels of the slope and intercept and the covariance between the two. Within-person effects are represented by the autoregression of the disturbances of the variable on previous values.

This disaggregation is easier to see and more relevant in the bivariate case, which allows for a myriad of relations between the residuals of both variables. The level 1 equations for an unconditional bivariate RALT are

\[ y_{it} = \alpha_{yi} + \lambda_{yi} \beta_{yi} + \delta_{yi} \]  \hspace{1cm} (20)

\[ z_{it} = \alpha_{zi} + \lambda_{zi} \beta_{zi} + \delta_{zi} \]  \hspace{1cm} (21)
Where the $\delta$ variables represent the composite residual terms. It is important to examine the breakdown of the residual terms, as that is where the disaggregation of between- and within-effects will take place.

In this example, the residuals will be allowed to have both autoregression and cross-regression terms, and the equations are thus

$$
\delta_{yt} = \rho_{\delta_{yt}\delta_{y_{t-1}}} \delta_{y_{t-1}} + \rho_{\delta_{yt}\delta_{z_{t-1}}} \delta_{z_{t-1}} + \epsilon_{\delta_{yt}}
$$

(22)

$$
\delta_{zt} = \rho_{\delta_{zt}\delta_{y_{t-1}}} \delta_{y_{t-1}} + \rho_{\delta_{zt}\delta_{z_{t-1}}} \delta_{z_{t-1}} + \epsilon_{\delta_{zt}}
$$

(23)

and it is assumed (in addition to the assumptions given for the univariate RALT) that all error terms for one variable are uncorrelated with the latent slope and intercept factors for the other variable,

$$
COV(\epsilon_{\delta_{yt}}, \alpha_{zt}) = COV(\epsilon_{\delta_{zt}}, \beta_{yt}) = COV(\epsilon_{\delta_{zt}}, \alpha_{zt}) = COV(\epsilon_{\delta_{zt}}, \beta_{yt}) = 0.
$$

A path diagram for this model is shown in Figure 7.
Figure 7. Bivariate and unconditional residualized autoregressive latent trajectory model.

Now it can be seen that the between- and within-person effects have been completely disaggregated. The between-person effects are represented by the means, variances, and covariances of the random slopes and random intercepts, and the within-person effects are represented by the autoregressions and
cross-regressions among residual terms. This model provides a way to cleanly estimate within- and between-person effects, and can now effectively model within-person processes without the contamination of between-effects. It is thus possible to draw conclusions at the within-person level.

The above model will be utilized in analyzing the data from the LSAY. However, this only covers half of the analytic problem presented at the beginning of the paper. The other half lies in the cross-classified data structure that arises from students switching schools. The given model assumes independence of observations when, in actuality, students are nested within the cross-classification of schools over time. For illustration purposes, it is necessary to briefly move into the multilevel modeling (MLM) framework, though the eventual goal is to incorporate this into the SEM.

1.6 Cross-Classified Data Structures

Within the standard MLM, it is assumed that observations are nested within higher-level units, and that from this nesting arises a potentially problematic correlation between observations within given units. For example, students may be nested in schools (Aitkin and Longford, 1986), or repeated measures may be nested within individual subjects (Laird & Ware, 1982). It can be assumed that students within a given classroom will be more correlated with each other than they will be with students in another classroom, or that repeated observations within individuals will be more correlated with each other than they will be with those of another individual. Failure to account for this nesting can result in biased standard errors, which in turn can result in errors of interpretation (Raudenbush & Bryk, 2002). As described in previous sections, this is accounted for by allowing for random variability around fixed effects. This is easily done within standard software programs and is relatively straightforward in terms of interpretation.

However, within this standard modeling procedure that is so prominent, it is typically the case that lower-level units are only nested within one higher-level unit. For example, students are only nested
within a single school or a single classroom. Though less common, there also exist many cases exist where units at the same level of a hierarchy are simultaneously nested within more than one factor. There are many situations in which students could potentially be nested within both schools and neighborhoods. In the longitudinal case, students switch classrooms and/or schools as the years progress, resulting in a more complicated nesting structure. Here, students are not nested within schools; they are now nested within cells of the level 3 cross-classification. Such cross-classified data structures pose issues in model estimation, and failure to account for the cross-classification can result in biased random effects estimates (Raudenbush, 1993).

Consider the case of the data analyzed in this paper. If students did not switch schools over time, the nesting structure would be a relatively straightforward one: time nested within students who are in turn nested within schools. This three-level nesting structure can be easily analyzed using standard statistical software packages (such as SAS’ PROC MIXED). Due to the switching of schools, this method cannot be utilized here and a standard MLM will not suffice in completely explicating the nature of the data. Methods need to be utilized to account for this cross-classified nesting structure within the SEM framework, though models are not yet well-developed.

In sum, a myriad of methodological issues are present that need to be accounted for. First, a main goal of this paper is to disaggregate the between- and within-person effects, and measures need to be taken to ensure that they are completely parsed apart. Second, the data are longitudinal and bivariate, so the appropriate over-time relations between the variables need to be modeled. Lastly, the cross-classifications needs to be considered in the analyses. The following section proposes an analytic plan that will incorporate these three main issues.
CHAPTER 2: METHODS

2.1 Sample

Students were sampled from 50 nationally representative schools that served as feeder schools to larger high schools. Within each school, up to 60 students were randomly selected, and, in total, 2,640 students participated. The sample was roughly even with regards to gender, with 52 percent of the sample being male. The sample was 70% White, 11% Black, 9% Hispanic, 3% Asian, 1% Native American, and 6% other or unavailable. The US was divided into four sampling regions (northeast, north central, south, and west) and students were sampled from urban, rural, and suburban areas according to the proportions in each region. Students entered the study at the beginning of the year in seventh grade, and data collection continued two times per year until one year post-high school graduation. A total of six time points will be considered in my analyses, representing the fall of each year from seventh grade through high school graduation. Because the final time point was when the students were out of high school, and because it did not assess students on the variables of interest, only the school years will be considered here. Students switched schools after the second time point, representing the move from junior high or middle school to high school.

2.2 Measures

Variables used are gender, a math test achievement score, and a measure of interest in math.

2.2.1 Gender

Females were coded as 0 and males were coded as 1.
2.2.2 Interest in math

Math interest was created from items that were recorded on a 5 item Likert-type scale with a score of 1 corresponding to "strongly disagree" and 5 corresponding to "strongly agree". To create the interest in math variable, three items were summed: "I enjoy math", "math is useful in everyday problems", and "I will use math often as an adult."

2.2.3 Math achievement

The math achievement score was created using items from the 1986 National Assessment of Educational Progress (NAEP) that covered a wide variety of topics, including algebra, statistics, and geometry. To ensure growth in test performance over the years, the test items differed across both grades and students within grades. To avoid floor and ceiling effects, and to maximize the amount of information gathered about math achievement, a student received either an easy, medium, or hard test form based on their performance at the last testing time. Though some items differed across test forms, the tests had many items in common so that tests could be equated.

The tests scores were then IRT scored and scaled to have a mean of 50 and a standard deviation of 10. Achievement scores were imputed by the original researchers prior to the public release of the data set. Only achievement scores were imputed, and it was done in all cases except in those where students dropped out of school or when four or more time points were missing. Imputation was done by either regressing the missing time point on the two previous time points, by computing the mean of the score immediately following and immediately prior to the missing time point, or by regressing the missing time point on earlier time points than the ones immediately prior to the missing one.
2.3 Patterns of Missing Data

Over 75 percent of cases were missing at least one math achievement score over the six year period. The majority (62 percent) of these cases were only missing one single time point, and less than seven percent of missing cases were due to school dropout.

2.4 Intra-Class Correlation Coefficients (ICCs)

To examine the degree of nesting within schools, intra-class correlation coefficients (ICCs) were calculated for each variable at the first time point using the formula

\[
ICC = \frac{\tau_{00}}{\tau_{00} + \sigma^2}
\]  

(24)

The ICC represents the degree of correlation across individuals in a group. Here, it represents the degree of dependence within schools. The ICC for math achievement at the first time point is .16 and the ICC for interest is .10. This indicates that between ten and twenty percent of the variance in each measure is due to between-groups differences.

2.5 Longitudinal Trends

Table 1 shows the time-specific correlations for each variable with itself and with other variables, as well as time-specific means and standard deviations for each variable.
Table 1.

Time-specific correlations, means, and standard deviations for achievement and interest in math.

<table>
<thead>
<tr>
<th></th>
<th>Act1</th>
<th>Act2</th>
<th>Act3</th>
<th>Act4</th>
<th>Act5</th>
<th>Act6</th>
<th>Act7</th>
<th>Act8</th>
<th>Act9</th>
<th>Act10</th>
<th>Act11</th>
<th>Act12</th>
<th>Int1</th>
<th>Int2</th>
<th>Int3</th>
<th>Int4</th>
<th>Int5</th>
<th>Int6</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>30</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>60</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>90</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>120</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Note: **p<.01; *p<.05. Ach = math achievement; Int= interest in math.

The highlighted regions along the diagonal show autocorrelations of a variable with itself; it can be seen that in all cases, correlations among variables at each time point generally tend to decay as two time points get further from each other.

In order to examine the general trends over time for each variable and how they varied by gender, means at each time point were calculated by gender, and the plots are shown below in Figures 8 and 9.
Figure 8. Math achievement over time as a function of gender.
Figure 9. Interest in math over time as a function of gender.

As indicated by the figures, patterns of gender differentiation in each variable are as predicted by theory. Figure 8 shows achievement in math over time by gender. It can be seen that females appear to do slightly better on math achievement tests. Figure 9 shows mean levels of interest in math over time, and males show higher degrees of interest in math relative to females. This pattern of findings (females are lower on interest in math despite equal or better performance than males) is consistent with theory and other empirical findings.
2.6 Hypotheses to Test

There were four main hypotheses that the aforementioned plan of analysis tested. First, I hypothesized that there would be significant within-person variability in the trajectories of change over time in math interest and achievement. I expected the trajectories to be linear in shape, but considered alternative functional forms should the need arise. Next, I hypothesized that between-person variability would covary; that is, I predicted that an individual’s starting point and rate of change over time would be related such that a higher initial score on a variable would be positively related to a steeper slope for that same variable. Third, I expected to see unique prospective within-person relations over time. I predicted that elevations in either variable at a specific time point would be predictive of subsequent elevations. Lastly, I predicted that between- and within-person processes would vary as a function of gender. I considered this last hypothesis to be exploratory in nature, as the specific nature of this relation was unknown.

2.7 Analytic Plan

My analyses consisted of starting with univariate growth models and gradually building up to my final model. I first fit unconditional, univariate growth curves to both interest in math and math achievement to establish the functional form for each process on its own and to ensure that the most basic model is appropriate before moving on to more complex models.

Once the growth models had been fitted and examined, I expanded the models and fit a univariate, unconditional RALT model to each outcome, which allowed me to examine the nature of the residuals and the between- and within-person processes for interest and achievement in math. Next, I fit an unconditional, bivariate RALT model to the data to examine the between- and within-person processes and to examine how the two variables interact at both levels. Covariances among the growth parameters
account for the between-person relations between starting point and rate of change in both variables, and the autoregression and cross-regression parameters make up the within-person processes.

In order to determine whether gender differences arise in the relationship between interest and achievement, I implemented a multiple group model to investigate how gender affects both the between-person and the within-person processes. I tested for differences among slope and intercept values, the covariances between slopes and intercepts, and the autoregression and cross-regressions between the residuals for each variable.

Because the cross-classification is not the core focus of the substantive or analytic challenges highlighted earlier, and because no extant models within the SEM framework were available to utilize, it was considered a nuisance variable and in all analyses the nesting at the first time point was accounted for through the use of the CLUSTER statement in MPlus. Additionally, a sensitivity analysis was attempted using a dummy code for each school that entered into the model as TICs.

2.8 Model Evaluation

To assess model fit, I considered a variety of fit statistics, including the chi-square test statistic, the Tucker-Lewis Index (TLI), the Incremental Fit Index (IFI), and the root-mean square error of approximation (RMSEA). Because my sample size was large, and because the chi-square test statistic was highly significant due to the excess statistical power that this entails, I supplemented the chi-square test with the other fit indices to get a global idea of the goodness of fit. In cases where a nonlinear trend was suspected, I used likelihood ratio tests (LRTs) to determine whether or not any higher-order polynomial trends significantly improved model fit, as the models were nested.
2.9 Evaluation of Assumptions

To ensure that the assumptions of homoscedasticity and normality of residuals all hold, I created and examined residual variables for each individual. To test for homoscedasticity, I plotted the residuals versus time and the residuals versus predicted values and examined the plots for any trends. To test for normality of residuals, I created a normal probability plot for each variable at each time point.

The preceding sections lay out an analytic plan that accounts for the bivariate, longitudinal nature of the data while disaggregating the between- and within-person effects and controlling for cross-classification. The plan of study also answers the substantive issues at hand of gender differences in the relationship between math achievement and interest in math over time through the use of a multiple group analysis framework. Thus, both the methodological and the substantive issues are accounted for in a novel framework that has not yet been explored in great detail.
CHAPTER 3: RESULTS

Results

The analytic process involved first fitting univariate LGC and RALT models to each construct separately to establish functional form and assess the general model fit. Additionally, differences in model fit between the LCG and the RALT was examined for each variable. The models were then combined and a bivariate RALT was fit. Lastly, a multi-group RALT was fit to the model to examine gender differences in the model parameters. To control for nesting, school membership at the first time point was accounted for in the analyses using the CLUSTER statement in Mplus, and standard errors were adjusted for subsequently.

3.1 Univariate Models for Math Achievement

3.1.1. Latent growth curve model

The initial analyses involved fitting univariate LGC and RALT models to each construct. Achievement in math was analyzed first, and an intercept-only LGC model was fit. This model proved to be a very poor fit to the data, $\chi^2(19)=6971.242$, $p<.001$, CFI=.477, TLI=.587, RMSEA=.343. Next, a model with a linear slope was fit, and the model fit the data better, but fit statistics still indicated relatively poor fit, $\chi^2(16)=686.863$, $p<.001$, CFI=.95, TLI=.953, RMSEA=.116. This model did fit the data significantly better than the null model ($\chi^2(3)=3143$, $p<.001$), but a better fit was to be desired. A quadratic model was fit to the data, and the model appeared to fit the data well, $\chi^2 (12)=124.864$, $p<.001$; CFI=.992, TLI=.989, RMSEA=.045. Visual inspection of the data and a likelihood ratio test between the linear and quadratic
models ($\chi^2(4)=281, p<.05$) indicated that a quadratic model was the best fit for the math achievement variable. Results of this are shown in Table 2.

Table 2.

Univariate latent growth model of math achievement

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>S.E.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variances</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\psi_{\alpha\alpha}$</td>
<td>92.377**</td>
<td>2.873</td>
</tr>
<tr>
<td>$\psi_{\beta_1\beta_1}$</td>
<td>13.786**</td>
<td>1.045</td>
</tr>
<tr>
<td>$\psi_{\beta_0\beta_0}$</td>
<td>.454**</td>
<td>.039</td>
</tr>
<tr>
<td>Covariances (Correlations)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\psi_{\alpha\beta}$</td>
<td>-3.011 (-.084)*</td>
<td>1.363</td>
</tr>
<tr>
<td>$\psi_{\alpha\beta}$</td>
<td>0.867 (.134)**</td>
<td>.256</td>
</tr>
<tr>
<td>$\psi_{\beta_1\beta_0}$</td>
<td>-2.256 (-.902)**</td>
<td>.195</td>
</tr>
<tr>
<td>Means</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu_\alpha$</td>
<td>50.259**</td>
<td>.184</td>
</tr>
<tr>
<td>$\mu_\beta_1$</td>
<td>4.355**</td>
<td>.102</td>
</tr>
<tr>
<td>$\mu_\beta_0$</td>
<td>-.348**</td>
<td>.021</td>
</tr>
</tbody>
</table>

Note: *$p<.05$, **$p<.01$

It can be seen that there are significant means, variances, and covariances for all three growth factors. There is a positive linear slope and a negative quadratic slope, indicating that achievement is growing over time at a positive rate, but that the acceleration slows over time. There are also significant relationships between the parameters; a higher intercept is correlated with a lower rate of change and a lower rate of deceleration. Additionally, a steeper linear slope is correlated with a higher rate of deceleration. Figure 10 shows the model-implied means plotted with the observed sample means. It can be seen that the two lines are very similar to one another, further indicating that this model is a good fit to the data and adequately captures the functional form.
Figure 10. Model-implied and sample means for the longitudinal trajectory of math achievement.

3.1.2 RALT model

Next, a univariate RALT model with equality constraints on the autoregression parameters was fit to the achievement variable, and the results are given below in Table 3.
Table 3.

Univariate RALT model for math achievement

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>S.E.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variances</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\psi_{\alpha}$</td>
<td>85.096**</td>
<td>4.131</td>
</tr>
<tr>
<td>$\psi_{\beta_1}$</td>
<td>8.234**</td>
<td>1.769</td>
</tr>
<tr>
<td>$\psi_{\beta_0}$</td>
<td>.245**</td>
<td>.058</td>
</tr>
<tr>
<td>Covariances (Correlations)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\psi_{\alpha_0}$</td>
<td>.458 (.017)</td>
<td>2.151</td>
</tr>
<tr>
<td>$\psi_{\alpha_0}$</td>
<td>.461 (.01)</td>
<td>.337</td>
</tr>
<tr>
<td>$\psi_{\beta_1\beta_0}$</td>
<td>-1.236 (-.87)**</td>
<td>.303</td>
</tr>
<tr>
<td>Means</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu_{\alpha}$</td>
<td>50.276**</td>
<td>.184</td>
</tr>
<tr>
<td>$\mu_{\beta_1}$</td>
<td>4.356**</td>
<td>.102</td>
</tr>
<tr>
<td>$\mu_{\beta_0}$</td>
<td>-.348**</td>
<td>.021</td>
</tr>
<tr>
<td>Regressions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rho_{1,t-1}$</td>
<td>.244**</td>
<td>.036</td>
</tr>
</tbody>
</table>

Note: *p<.05, **p<.01

The data fit the model well, $\chi^2 (11)=61.659$, p<.001; CFI=.996, TLI=.995, RMSEA=.038. Additionally, a likelihood ratio test between the unconditional LCM and the unconditional RALT for achievement indicates that the RALT is a significantly better fit to the data, $\chi^2 (1)=31.6$, p<.001.

Here, the covariances between the intercept term and the two slopes that were initially significant in the univariate LCM for achievement become insignificant. There is a significant and positive autoregression between the residuals from an individual’s math achievement at one time point and their achievement at the time point immediately following that, thus indicating that higher residual levels of achievement are predictive of subsequent higher residual levels of achievement net the effects of the growth parameters.
3.2 Univariate Models for Interest in Math

3.2.1 Latent growth curve model

A univariate LGC was also fit to the interest variable. An intercept-only model was initially fit to the data, and this model proved to be an insufficient fit, $\chi^2 (19)=1006.95$, $p<.001$; CFI=.711, TLI=.772, RMSEA=.129. Next, a model with a linear slope was fit, and this model fit the data well, $\chi^2 (16)=117.002$, $p<.001$; CFI=.97, TLI=.972, RMSEA=.045. The linear model also showed a significant improvement in fit when a likelihood ratio test was performed between that and the null model, $\chi^2 (3)=445$, $p<.001$. The fit statistics for the linear model indicate an improvement in model fit as well. A quadratic model was also fit, and this model also fit the data well, $\chi^2 (12)=40.301$, $p<.001$; CFI=.992, TLI=.99, RMSEA=.028. A likelihood ratio test between the linear and quadratic models indicated that the quadratic model was a better fit to the data, $\chi^2 (4)=38$, $p<.001$. However, upon inspection of parameter estimates, it was noted that the quadratic parameter was not significant when considered both as a fixed effect and allowed to have random effects. Thus, the linear model was retained, as the nonsignificant quadratic factor added nothing of substantive interest to the model. Results of the univariate model for interest are shown in Table 4.

Table 4.
Univariate latent growth model for interest.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>S.E.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variances</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\psi_{aa}$</td>
<td>2.772**</td>
<td>.128</td>
</tr>
<tr>
<td>$\psi_{bb}$</td>
<td>.153**</td>
<td>.011</td>
</tr>
<tr>
<td>Covariances (Correlations)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\psi_{ab}$</td>
<td>-.242 (-.371)**</td>
<td>.032</td>
</tr>
<tr>
<td>Means</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu_{a}$</td>
<td>11.645**</td>
<td>.039</td>
</tr>
<tr>
<td>$\mu_{b}$</td>
<td>-.240**</td>
<td>.012</td>
</tr>
</tbody>
</table>

Note: *$p<.05$, **$p<.01$
It can be seen from the table that all growth parameters were significant. At the initial time point, students have an average interest value of 11.6, and that value drops and average of -.24 units over time. There is also a negative covariance between the intercept and slope factors; students show higher levels of interest initially tend to show steeper declines in their levels of interest, and vice versa. Figure 11 shows the model-implied means plotted with the estimated sample means. The model appears to fit the data well.

![Figure 11. Model-implied and sample means for the longitudinal trajectory of interest in math.](image)

3.2.2 RALT model

The latent growth model was an adequate fit to the data for interest in math, but the RALT model could potentially further improve the fit, and an unconditional, univariate RALT model with equality constraints on the autoregression parameters was fit to the data. The model fit the data well, $\chi^2$
A likelihood ratio test was conducted between the univariate LCM for interest in math and the univariate RALT for interest in math and the RALT was shown to be a significantly better fit to the data, $\chi^2 (1)=35$, $p<.001$. Results from the univariate RALT are shown in Table 5.

**Table 5.**

Univariate RALT model for interest in math

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>S.E.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Variance</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\psi_{\alpha \alpha}$</td>
<td>2.172**</td>
<td>.167</td>
</tr>
<tr>
<td>$\psi_{\beta \beta}$</td>
<td>.099**</td>
<td>.014</td>
</tr>
<tr>
<td><strong>Covariances (Correlations)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\psi_{\alpha \beta}$</td>
<td>-.107 (-.371)**</td>
<td>.041</td>
</tr>
<tr>
<td><strong>Means</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu_{\alpha}$</td>
<td>11.637**</td>
<td>.039</td>
</tr>
<tr>
<td>$\mu_{\beta}$</td>
<td>-.236**</td>
<td>.012</td>
</tr>
<tr>
<td><strong>Regressions</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rho_{t,t-1}$</td>
<td>.158**</td>
<td>.020</td>
</tr>
</tbody>
</table>

*Note: *$p<.05$, **$p<.01$*

Of interest here is the significant autoregression between temporal residuals. It can be seen that there is a significantly positive predictive relationship between the residuals at one time point and the residuals at the immediately following time point. Thus, higher residual values on interest in math are predictive of subsequent residual values, holding constant the effects of the growth parameters.

### 3.3 Bivariate RALT

While these significant results are interesting, it is of substantive theoretical interest to examine the cross-regressions between the two variables to examine the within-person effects that characterize the relationship between interest and achievement in math. To do so, a series of bivariate RALT models
were fit to the data. The RALT models separated the residual variances from the latent growth models highlighted above into unique and specific variances, and the relationships between residuals across both variables were examined. A path diagram from Figure X shows the final RALT model fit. Initially, a bivariate RALT model with autocorrelations but no cross-regressions was fit to the data. The model appeared to fit the data well, $\chi^2 (50)=95.549$, p<.001; CFI=.997, TLI=.995, RMSEA=.017. Because the cross-regressions were also of significant substantive interest, a model was fit that allowed for cross-regressions among the residuals of the variables. The model fit the data well, $\chi^2 (48)=92.187$, p<.001; CFI=.997, TLI=.996, RMSEA=.017. A likelihood ratio test between the two models shows no significant improvement in fit between the two models, $\chi^2 (2)=2.07$, p=.355. Though not a significant fit, the results of the final model are shown below in order to examine the cross-regressions. Means and regressions are shown in Table 6, while variances, covariances, and correlations among the growth parameters are shown in Table 7. It should be noted that the individual residuals across construct and within time were covaried within the model as well, though it is not reported in the table.

Table 6.
Means and regressions for an unconditional and bivariate RALT

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>S.E.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_{\alpha_c}$</td>
<td>50.517**</td>
<td>.596</td>
</tr>
<tr>
<td>$\mu_{\beta_{t-1}}$</td>
<td>4.427**</td>
<td>.226</td>
</tr>
<tr>
<td>$\mu_{\beta_{t-0}}$</td>
<td>-3.56**</td>
<td>.045</td>
</tr>
<tr>
<td>$\mu_{\alpha_t}$</td>
<td>11.683**</td>
<td>.073</td>
</tr>
<tr>
<td>$\mu_{\beta_{t}}$</td>
<td>-.232**</td>
<td>.021</td>
</tr>
<tr>
<td>Regressions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rho_{t-1, t-0}$</td>
<td>.265**</td>
<td>.105</td>
</tr>
<tr>
<td>$\rho_{t-1, t-0}$</td>
<td>.168**</td>
<td>.033</td>
</tr>
<tr>
<td>$\rho_{\delta_{t-1}, \delta_{t-1}}$</td>
<td>.048</td>
<td>.071</td>
</tr>
<tr>
<td>$\rho_{\delta_{t}, \delta_{t-1}}$</td>
<td>-.007</td>
<td>.013</td>
</tr>
</tbody>
</table>

Note: *p<.05, **p<.01
Table 7.

Variance, covariances, and correlations among the growth parameters in bivariate RALT

<table>
<thead>
<tr>
<th></th>
<th>( \mu_{\alpha m} )</th>
<th>( \mu_{\beta m} )</th>
<th>( \mu_{\alpha \text{ach}} )</th>
<th>( \mu_{\beta \text{ach}} )</th>
<th>( \mu_{\beta \text{ach}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \mu_{\alpha m} )</td>
<td>2.114 (.286)**</td>
<td>-.125</td>
<td>.190**</td>
<td>.034</td>
<td>.075</td>
</tr>
<tr>
<td>( \mu_{\beta m} )</td>
<td>-.047</td>
<td>.065 (.026)**</td>
<td>-.071</td>
<td>.073</td>
<td>.073</td>
</tr>
<tr>
<td>( \mu_{\alpha \text{ach}} )</td>
<td>2.515**</td>
<td>-.168</td>
<td>83.18 (9.4)**</td>
<td>.113</td>
<td>.092</td>
</tr>
<tr>
<td>( \mu_{\beta \text{ach}} )</td>
<td>.121</td>
<td>.046</td>
<td>2.510</td>
<td>5.95 (4.96)**</td>
<td>-.847**</td>
</tr>
<tr>
<td>( \mu_{\beta \text{ach}} )</td>
<td>.047</td>
<td>.008</td>
<td>.358</td>
<td>-.886**</td>
<td>.184 (.162)**</td>
</tr>
</tbody>
</table>

Note: *p<.05, **p<.01. Variances and their standard errors are on the diagonal, covariances are in the bottom off-diagonal, and correlations are in the top off-diagonal.

Despite the relatively good fit, there appear to be no significant cross-regressions between variables. Models were fit that allowed all cross-and autoregression parameters to be free, and models were fit that fixed the parameters to equality. Between those, no differences in significance were obtained; the likelihood ratio test showed no significant differences between the models, \( \chi^2 (8)=3.2, p>.05 \), thus, the models with fixed parameters are presented here for parsimony.

3.4 Multiple-Group Models

Lastly, a series of multiple-group models were fit to the data to examine gender differences in the means, variances, covariances, cross-regressions, and autoregressions for math interest and achievement. An initial baseline model was fit to the data in which all parameters were free to vary across gender. Means and variances for males and females are reported in Table 8, and two separate tables of variances, covariances, and correlations are given in Tables 9 and 10.
### Table 8.
Means and regression parameters for bivariate RALT model of math interest and achievement by gender

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>S.E.</th>
<th>Estimate</th>
<th>S.E.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Means</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu_{\text{ach}}$</td>
<td>50.879**</td>
<td>2.71</td>
<td>49.627**</td>
<td>.244</td>
</tr>
<tr>
<td>$\mu_{\beta_{\text{ach}}}$</td>
<td>4.047**</td>
<td>.152</td>
<td>4.653**</td>
<td>.134</td>
</tr>
<tr>
<td>$\mu_{\beta_{Q_{\text{ach}}}}$</td>
<td>-.290**</td>
<td>.032</td>
<td>-.407**</td>
<td>.028</td>
</tr>
<tr>
<td>$\mu_{\text{int}}$</td>
<td>11.562**</td>
<td>.055</td>
<td>11.707**</td>
<td>.055</td>
</tr>
<tr>
<td>$\mu_{\beta_{\text{int}}}$</td>
<td>-.232**</td>
<td>.017</td>
<td>-.244**</td>
<td>.020</td>
</tr>
<tr>
<td><strong>Regressions</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rho_{\text{ach}, \text{int}}$</td>
<td>.282**</td>
<td>.055</td>
<td>.192**</td>
<td>.045</td>
</tr>
<tr>
<td>$\rho_{\text{ach}, \text{int}^{-1}}$</td>
<td>.130**</td>
<td>.027</td>
<td>.187**</td>
<td>.028</td>
</tr>
<tr>
<td>$\rho_{\beta_{\text{ach}}, \beta_{\text{int}}}$</td>
<td>-.022</td>
<td>.057</td>
<td>.045</td>
<td>.048</td>
</tr>
<tr>
<td>$\rho_{\delta_{\text{ach}}, \delta_{\text{int}}^{-1}}$</td>
<td>-.009</td>
<td>.008</td>
<td>-.012</td>
<td>.009</td>
</tr>
</tbody>
</table>

*Note: *$p<.05$, **$p<.01$.

### Table 9.
Variances, covariances, and correlations for a bivariate RALT model for males

<table>
<thead>
<tr>
<th>Parameter</th>
<th>$\mu_{\text{ach}}$</th>
<th>$\mu_{\beta_{\text{ach}}}$</th>
<th>$\mu_{\text{int}}$</th>
<th>$\mu_{\beta_{\text{int}}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_{\text{int}}$</td>
<td>2.01 (.24)**</td>
<td>-.177</td>
<td>1.511**</td>
<td>-.065</td>
</tr>
<tr>
<td>$\mu_{\beta_{\text{int}}}$</td>
<td>-.080</td>
<td>.101 (.021)**</td>
<td>-.125</td>
<td>.163</td>
</tr>
<tr>
<td>$\mu_{\text{ach}}$</td>
<td>1.838**</td>
<td>-.342</td>
<td>73.88 (4.44)**</td>
<td>-.157**</td>
</tr>
<tr>
<td>$\mu_{\beta_{\text{ach}}}$</td>
<td>-.257</td>
<td>.145</td>
<td>-1.360**</td>
<td>7.82 (1.88)**</td>
</tr>
<tr>
<td>$\mu_{\beta_{Q_{\text{ach}}}}$</td>
<td>.044</td>
<td>.003</td>
<td>.520</td>
<td>-1.198**</td>
</tr>
</tbody>
</table>

*Note: *$p<.05$, **$p<.01$. Variances and their standard errors are on the diagonal, covariances are in the bottom off-diagonal, and correlations are in the top off-diagonal.
Table 10.

Variances, covariances, and correlations for a bivariate RALT model for females

<table>
<thead>
<tr>
<th></th>
<th>( \mu_a )</th>
<th>( \mu_{\beta_a} )</th>
<th>( \mu_{\text{ach}} )</th>
<th>( \mu_{\beta_{ach}} )</th>
<th>( \mu_{\beta_{ach}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \mu_{a_m} )</td>
<td>2.34 (.230)**</td>
<td>-.288**</td>
<td>.233**</td>
<td>.202**</td>
<td>-.154</td>
</tr>
<tr>
<td>( \mu_{\beta_m} )</td>
<td>-.138**</td>
<td>.099 (.019)**</td>
<td>-.102</td>
<td>.144</td>
<td>-.005</td>
</tr>
<tr>
<td>( \mu_{\text{ach}} )</td>
<td>3.479**</td>
<td>-.312</td>
<td>95.01 (7.36)**</td>
<td>.086</td>
<td>.066</td>
</tr>
<tr>
<td>( \mu_{\beta_{ach}} )</td>
<td>.911**</td>
<td>.134</td>
<td>2.468</td>
<td>8.69 (3.14)**</td>
<td>- .866**</td>
</tr>
<tr>
<td>( \mu_{\beta_{ach}} )</td>
<td>-.117</td>
<td>.001</td>
<td>.322</td>
<td>-1.271**</td>
<td>.25 (.098)**</td>
</tr>
</tbody>
</table>

Note: *p<.05, **p<.01. Variances and their standard errors are on the diagonal, covariances are in the bottom off-diagonal, and correlations are in the top off-diagonal.

Additionally, separate model-implied trajectories for males and females on interest and achievement in math are shown in Figures 12 and 13.
Figure 12. Model-implied trajectories for interest in math over time as a function of gender.
Figure 13. Model-implied trajectories for math achievement over time as a function of gender.

The next model fixed all growth parameter means to equality across groups. Next, means, variances, and covariances were fixed to equality. To test for equivalence of the within-group processes, a model was fit that fixed all means, variances, covariances, and autoregressions to equality. Lastly, means, variances, covariances, and all regression parameters were set to equality. Results of the likelihood ratio tests are shown below in Tables 11.a and 11.b.
Table 11.a.

Likelihood ratio tests of multiple-group analysis of a bivariate RALT for males and females

<table>
<thead>
<tr>
<th>Model</th>
<th>( \chi^2 ) value</th>
<th>df</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Baseline</td>
<td>129.18</td>
<td>84</td>
</tr>
<tr>
<td>2</td>
<td>( \mu_{\alpha}^B = \mu_{\alpha}^G ) and ( \mu_{\beta}^B = \mu_{\beta}^G )</td>
<td>156.86</td>
<td>89</td>
</tr>
<tr>
<td>3</td>
<td>Model 2 and ( \psi_{\alpha\alpha}^B = \psi_{\alpha\alpha}^G ); ( \psi_{\beta\beta}^B = \psi_{\beta\beta}^G ); ( \psi_{\alpha\beta}^B = \psi_{\alpha\beta}^G )</td>
<td>206.73</td>
<td>104</td>
</tr>
<tr>
<td>4</td>
<td>Model 3 and ( \rho_{\delta_{\alpha\alpha}}^B = \rho_{\delta_{\alpha\alpha}}^G ); ( \rho_{\delta_{\beta\beta}}^B = \rho_{\delta_{\beta\beta}}^G ); ( \rho_{\delta_{\alpha\beta}}^B = \rho_{\delta_{\alpha\beta}}^G )</td>
<td>208.55</td>
<td>106</td>
</tr>
<tr>
<td>5</td>
<td>Model 4 and ( \rho_{\delta_{\alpha\alpha}}^B = \rho_{\delta_{\alpha\alpha}}^G ); ( \rho_{\delta_{\beta\beta}}^B = \rho_{\delta_{\beta\beta}}^G ); ( \rho_{\delta_{\alpha\beta}}^B = \rho_{\delta_{\alpha\beta}}^G )</td>
<td>223.57</td>
<td>108</td>
</tr>
</tbody>
</table>

Note: Superscript B indicates male; superscript G indicates female

Table 11.b.

Likelihood ratio difference tests between nested constraints of multiple-group analysis for bivariate RALT

<table>
<thead>
<tr>
<th>Model Comparison</th>
<th>Difference (abs)</th>
<th>df</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 vs. 1</td>
<td>27.68</td>
<td>5</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>3 vs. 2</td>
<td>86.36</td>
<td>15</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>4 vs. 3</td>
<td>1.82</td>
<td>2</td>
<td>.403</td>
</tr>
<tr>
<td>5 vs. 4</td>
<td>15.02</td>
<td>2</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

It can be seen that there are significant differences between an entirely free model and one with parameter means fixed to equality and between a model with fixed means and a model with fixed means, variances, and covariances. No differences arise when the autoregression parameters are fixed to equality, but the model fit improves when the cross-regression parameters are free to vary. This indicates that males and females differ in terms of their means, variances, and covariances for interest and achievement in math, but do not differ in terms of their autoregression parameters. Because the cross-regression parameters are not significant themselves, the fact that they are significantly different between males and females is of little utility, as the values of the cross-regression parameters are miniscule themselves. Thus, males and females differ between groups when it comes to interest and achievement in math, but not within groups.
In sum, there appear to be no significant cross-regressions in any of the models tested. Results of a multiple-group model indicate that males and females differ in terms of their means, variances, and covariances in the bivariate RALT, but not in terms of autoregression parameters. It was indicated that males and females did differ significantly on their cross-regression parameters, but, as the parameters were insignificant, this does not mean much here.

### 3.5 Sensitivity Analysis

In order to determine whether or not the cross-classification of students within schools affected the model enough that simply controlling for nesting at the first time point was insufficient, I attempted a sensitivity analysis. I created dummy codes for each school and entered them into the final bivariate RALT model as TICs. There were a total of 50 dummy codes, and the results of the full model with all dummy variables in the model proved to be intractable for unknown reasons.

### 3.6 Evaluation of Assumptions

In order to determine whether or not the cross-classification of students within schools affected the model enough that simply controlling for nesting at the first time point was insufficient, I attempted a sensitivity analysis. I created dummy codes for each school and entered them into the final bivariate RALT model as TICs. There were a total of 50 dummy codes, and the results of the full model with all dummy variables in the model proved to be intractable for unknown reasons.

There were several model assumptions that needed to hold in order for the models to justify fitting the models to the data. To test for the assumption of normality of residuals, I created normal probability and QQ plots for the residuals for each variable at each time point. The plots for interest in math are shown in Figure 14, and the plots for math achievement are shown in Figure 14.
Figure 14. Normal probability and QQ plots for interest in math at each time point.
Figure 15. Normal probability and QQ plots for math achievement at each time point.
Both plots show that the assumption of normality of residuals was not severely violated for either variable at either time point.

To test for the assumption of homoscedasticity, or equality of variance over time, I plotted the residual values versus the predicted values and the residual values versus time for each variable and examined the results. Figure 16 shows residuals versus predicted values for interest in math.

![Residual values versus predicted values for interest in math.](image)

*Figure 16. Residual values versus predicted values for interest in math.*

The slight curvature of the fitted line indicates a possible omitted predictor or non-linearity, but nothing egregious enough to discount the model. The likelihood ratio test between a linear and quadratic model for math interest did show a significant improvement in model fit, but the quadratic parameter itself was not significant. It was thus omitted from the model; the inclusion of the fact may improve the look of the
residuals versus predicted plot, but the quadratic factor added nothing substantive to the model. The plot for residuals versus predicted for math achievement is shown in Figure 17.

![Figure 17. Residual values versus predicted values for math achievement.](image)

No issues are apparent, given the flat line in the plot.

Lastly, to further test for homoscedasticity, residuals versus time were plotted for four random samples of 100 subjects for each variable. Plots for math interest are shown in Figure 18, and plots for achievement are shown in Figure 19. The plots show no apparent trends in residual values over time.
Figure 18. Residuals versus time plotted for four random subsamples of 100 subjects for math interest.
Figure 19. Residuals versus time plotted for four random subsamples of 100 subjects for math achievement.

Taking the two sets of plots that test for homoscedasticity, the residuals appear to be fairly equal over time.

In sum, the model assumptions of normality of residuals and homoscedasticity appear to hold in the model.
Discussion

The preceding analyses aimed to test gender differences in the longitudinal relationship between interest in math and math achievement while simultaneously disaggregating between-and within-person effects. Though the majority of psychological research and theories consider within-person processes, the preponderance of available methodologies for analyzing such data focus solely on between-person processes. Misattributing between-person processes as within-person processes can lead to errors of inference, and my project focused on expanding a new methodology to avoid such consequences. The RALT model was introduced as a novel method within the SEM framework that can concurrently model between- and within-person effects.

There were four main hypotheses that I tested. First, I hypothesized that there would be significant between-person variability in the trajectories of change over time in math interest and achievement. Next, I hypothesized that between-person variability would covary; that is, I predicted that an individual's starting point and rate of change over time would be related such that a higher initial score on a variable would be positively related to a steeper slope across constructs. Third, I expected to see unique prospective within-person relations over time. I predicted that elevations in either variable at a specific time point would be predictive of subsequent elevations. Lastly, I predicted that between- and within-person processes would vary as a function of gender. I used a series of LCMs and RALT models to test these hypotheses.
4.1 Univariate Models

Univariate LCM and RALT models were applied to each variable. Results of the latent curve model for math achievement indicated that math achievement systematically increased over time, though this increase decelerated over time as well. Higher intercept values were also correlated with a lower rate of change and lower rate of deceleration, and steeper slopes were correlated with a higher rate of deceleration. Thus, high-achieving individuals reported lesser increases in terms of their achievement relative to their lower-achieving peers, but the change in the rate of change of their achievement is also slower. Substantively, it can be concluded that the higher-achieving subjects tend to be more stable in their achievement over time, showing less change than lower achievers. This could be due to a ceiling effect; individuals close to the top in terms of achievement do not have much room for improvement.

A univariate RALT for math achievement indicated that higher residual values, or deviations from the underlying trajectory, for achievement significantly predict subsequent elevations in residual values. This is the within-person effect and indicates that, at the individual level, if an individual is higher than the mean achievement at one time point relative to their mean trajectory, they will be higher at the next time point. Higher-achieving individuals tend to be higher-achieving at each time point. This suggests that, as an individual does better and better in terms of their mean trajectory, they will do better and better and the next time point. The effect carries over from one time point to the next.

Univariate models were also applied to the interest in math variable. Results of the LCM indicated that interest decreased over time at a constant rate. Additionally, there was a negative covariance between the intercept and slope factors. Taken together, this indicates that students with higher initial levels of interest shower steeper declines in their level of interest over time. This could be due to students’ coming into the study in the seventh grade having never taken a true math course and only having experienced relatively easy math concepts. As they advance, the potential for difficulties and frustrations
in math may lead them to become less interested as the concepts get harder. However, data were not available to test this hypothesis.

The RALT model for interest in math also indicated that increased residual values are predictive of subsequent elevations at the within-person level, indicating that students who have higher levels of interest relative to their underlying tend to have higher levels of interest across all time points.

4.2 Bivariate RALT

The bivariate RALT model was applied to the LSAI data set to determine the relationship between interest and achievement in math. The means, variances, and covariances among the growth parameters are representative of between-person effects. Within constructs, results were identical to those reported in the univariate models. Across constructs, there was one significant positive covariance between the intercept factors for interest and achievement in math, which indicates that higher initial values of achievement are related to higher initial levels of interest in math.

The within-person results from the bivariate RALT generally did not support my hypotheses. Though the model fit the data well overall, the main parameters of interest—the cross-regressions among the disturbance terms—were nonsignificant, indicating that elevated levels in the residuals of one variable do not significantly predict subsequent elevated residual levels in the other. The autoregressions on the residuals of the variables on themselves were significant, indicating that higher residual values from the means on a variable significantly predict immediately following elevations from the overall mean, net the effects of the growth parameters. In a substantive sense, these residual regressions represent the within-person processes, and students who show higher levels of interest and achievement at one time point are likely to show the same elevated levels at subsequent time points. The significant autocorrelations indicate that at an individual level, a higher deviation from the mean slope and intercept values is predictive of later deviations as well.
In sum, initial levels of achievement were related to initial levels of interest at the between-person level. At the within-person level, individuals who were higher than their mean trajectories at one time point were more likely to show elevations from their underlying trajectory at subsequent time points, but elevations in one variable were not predictive of subsequent elevations in the other variable.

4.3 Multiple-Group Model

A multiple-group model was also implemented to test for gender differences in the relationship between interest and achievement in math over time, as it was hypothesized that males and females would differ in terms of how the constructs were related. Results showed that males and females differed in terms of their means, variances, and covariances in the bivariate RALT. In terms of means, females appear to be higher on initial achievement and lower on initial interest than males. They show less growth over time in achievement, though this growth rate decelerates at a lower rate than males. Thus, females show lower overall gains in achievement, but their rate of change remains more stable than males. In terms of growth in interest, females’ interest declines slightly more rapidly over time, though an inspection of the model-implied trajectories shown in Figure 13 and the raw means shown in Figure 9 indicates that females still show lower interest levels than males at the final time point in the study. These results support the findings highlighted in my introduction section; females show lower levels of interest in math than males despite slightly higher levels of achievement (Lapan et al., 2000; Turner et al., 2008).

In terms of covariances, females show a slightly different pattern of significant covariances than males. Both groups had significant positive covariances between the linear and quadratic slope terms for math achievement and the intercept terms for achievement and interest. This implies that, for both groups, if a subject initially shows rapid improvement on the math achievement test, their improvement tends to level off more quickly over time, and that students who are initially more interested in math also show higher scores on the math achievement test. Females, however, showed a significant negative
covariance between the slope term for interest and the intercept term for interest, indicating that higher initial values on interest are related to more rapid declines in interest over time. Females also showed a significant positive covariance between the linear slope factor for achievement and the intercept term for interest, indicating that higher levels of interest are correlated with higher levels of achievement. These results are interesting, as they suggest that females who were initially more interested in math show the greatest decreases in interest over time, and that higher levels of interest are indicative of higher levels of achievement over time. Taken together, these results support the extant empirical research conclusions highlighted in the introduction: females who are more interested in math tend to be higher-achieving in math, but those with higher levels of interest also tend to show the greatest decrease in interest (Lubinski and Benbow, 2006).

My results indicate that there are some significant between-person differences in terms of the longitudinal relationship between interest and achievement in math with regards to males and females. The latter part of the multiple-group model focused on within-person differences, and the autoregressions and cross-regressions were examined to explore any gender differences. Males and females were not different in terms of the autoregression parameters; that is, there are no significant differences between the parameters themselves, indicating that the within-person effects for that process are not different. The cross-regression parameters, however, were significantly different, but because the parameters themselves were not significant in any model fit, this is a relatively meaningless finding. This implies that, while males and females differ significantly from one another, the across-construct regressions do not significantly differ from zero. It appears that males and females differ in terms of their between-person effects, but not in terms of within-person effects.

Taken together, these results indicate that, though math achievement increases over time, interest in math decreases over time. Individuals who show greater strides in terms of their performance
on the achievement tests at first show lesser increases at latter time points. Females report less interest in math despite earning higher scores on the achievement tests, and females who reported higher initial interest in math show faster decreases in their interest levels and higher initial achievement scores. This indicates that students show increases in math achievement over time despite reporting diminishing interest. This trend is even more salient in females. At the within-person level, students who are more interested than their peers in math or who show higher levels of achievement tend to have higher levels across all time points.

4.4 Potential Limitations

Potential limitations of this study include the need to link the trajectories to real-world outcomes, potential mediators and moderators, the inability to model the cross-classification of the data, and time scaling. In order to fully evaluate the relationship between interest and achievement in math, the trajectories need to be linked to real-world outcomes, such as college success or career choice. Because those outcomes were not available here, it is unclear if either variable ended up being predictive of subjects’ career choices or satisfaction. The ultimate substantive goal of my project was to explore what happens in adolescence that leads females to be less interested in a STEM career, and while I found that they were, I was unable to determine what career subjects actually chose.

There are a number of potential mediators and moderators that would be interesting to consider. Parental support and encouragement to pursue STEM fields would most definitely affect interest and achievement, as would peers’ attitudes and support. For females, the amount of gender stereotyping they perceive would also be an interesting potential factor to add as a moderator into the model. Additionally, socioeconomic status (SES) probably plays a role here, as lower SES students may have less available resources to foster an interest in a STEM career and enable them to pursue such careers. Teacher quality and ability (or how qualified and able the students perceive the teachers) is also likely to play a role here,
as students who respect and like their teachers are more likely to go to them for help and view them as positive role models.

Because the sensitivity analysis proved to be computationally intractable, the cross-classification within the structure of the data was not accounted for. This could potentially result in biased standard errors. As stated earlier in the text, failure to control for cross-classification may result in consistent parameter estimates but biased standard errors. This could lead to errors of inference with regards to statistical significance.

Lastly, the time scale and length of time between time points could be a factor here as well. Because an entire year passed between each measurement occasion, the relationship between interest and achievement could be lessened due to the long time elapsed. Students may have had time to forget the effects of a particularly low or high achievement score, and the effects of that on their interest are probably less salient. It would be interesting to apply this to a shorter time span to determine whether or not the effects change.

4.5 Future Directions

There are a number of potential future directions for the RALT model and the substantive issue explored here. The first involves conducting simulation studies to explore the RALT model in greater detail. It would also be helpful to use other substantive data sets to see if the model fits better in other situations; as mentioned earlier, a shorter time span would be ideal in order to explore the relations at a more immediate level. Additionally, it would be ideal to find another data set for the analysis of the relationship between interest and achievement in math and also allows for real-world outcomes to be linked to the trajectories.
An interesting potential future study could involve an intervention, where programs are implemented to increase students’ interest in math. It would then be possible to look for subsequent increases in achievement and have a better understanding of whether or not increased interest is predictive of increased achievement, and vice versa.

4.6 Conclusions

My project analyzed gender differences in the between- and within-person effects in the relationship between interest and achievement in math. Unique contributions included the implementation of a novel modeling framework that fully disaggregated between- and within-person effects and the exploration of both processes as a function of gender. Though my hypotheses were not fully supported, the RALT model still provides a unique and novel way to fully parse apart between- and within-person effects. Doing this circumvents any possible errors of inference due to the misattribution of within-effects as between-effects (and vice versa) and is the first approach within the SEM framework that permits such a complete disaggregation. The relationship between interest and achievement in math as a function of gender was also explored in a way not yet done, and the between- and within-person effects were explored at great lengths in a new way.
REFERENCES


