

Association between Children's Average, Current, and Past Peer Context  
and Their Own Level of Aggression

M.A.S. Comprehensive Project  
Stat 581

Kelly Rulison  
The Pennsylvania State University

## Association between Children's Average, Current, and Past Peer Context and Their Own Level of Aggression

### 1 Conceptual Overview

Affiliating with aggressive peers is one of the strongest predictors of children's own aggression (e.g., Coie & Dodge, 1998). This association likely reflects both selection and influence processes (e.g., Kandel, 1978). For example, aggressive children may choose to spend time with other aggressive youth (selection). In turn, affiliating with aggressive peers may lead children to become more aggressive over time (influence).

Disentangling selection and influence processes is complicated by the dynamic nature of children's peer networks: new friendships form and old friendships dissolve. However, despite the instability in the *identities* of particular network members, there can be stability in the *characteristics* of these peers (Neckerman, 1996). For example, a child may end a friendship with one aggressive peer and form a friendship with another aggressive peer: the identities of the peers have changed but the child's typical (i.e., averaged over time) peer context is the same.

Therefore, at any time the level of aggression within a child's peer network reflects two factors: (1) the child's general tendency to affiliate with aggressive peers; and (2) the child's deviation from that tendency at that time. A child's general tendency to have aggressive peers indicates between-person variation and is labeled "average peer context"; a child's deviation from that typical level of peer aggression at a particular assessment indicates within-person variation and is labeled "current peer context." Children's current behavior can also be impacted by their previous peers' behavior. A child's deviation from their average peer context at the *past* assessment also indicates within-person variation and is labeled "past peer context."

The goal of the current analyses is to explore the independent contributions of children's average, current, and past peer context to their own aggression. Because boys are generally more aggressive than girls (Coie & Dodge, 1998), the moderating effect of gender will also be explored.

#### 1.1 Research Questions:

1. Are individual differences in the tendency to affiliate with aggressive peers associated with a child's average aggression?
2. Are changes in current peer context associated with changes in a child's aggression?
3. Are changes in past peer context associated with changes in a child's aggression?
4. Are these associations similar for boys and girls?

#### 1.2 Statistical Questions

1. Is between-person variation in peer context associated with a child's average aggression?
2. Is within-person variation associated with changes in a child's aggression?

3. Is lagged within-person variation associated with changes in a child's aggression?
4. Are these associations moderated by gender?

## 2 Method

### 2.1 Participants

Participants were students at in a school district in a rural, working-class community. Current analyses focus on a sample of 427 students (193 girls, 234 boys) who were enrolled in 3<sup>rd</sup> (N = 146), 4<sup>th</sup> (N = 146), or 5<sup>th</sup> (N = 135) grade in Fall 2001. Students participated until they completed 7<sup>th</sup> grade. To describe students' peer contexts, data were collected from all eligible students at each assessment (92 - 94% of all enrolled students). All youth in the district attended a single elementary (K – 5<sup>th</sup> grades) and middle school (6<sup>th</sup> – 8<sup>th</sup> grades). In elementary school, there were six self-contained classes ( $M = 25$  students). In 6<sup>th</sup> grade, youth remained with peers from their homeroom class ( $M = 23$  students) for most of the day, switching teachers for different subjects. In 7<sup>th</sup> grade, youth switched peers and teachers for each subject. Consistent with the community population, 99% of the participants were Caucasian. The distribution of achievement scores for 5<sup>th</sup> graders at the school was similar to the statewide distribution, but the district had above average poverty and school dropout rates.

### 2.2 Procedure

Peer-nominations were obtained through a 45-minute group-administered survey in October and May of each school year. Teachers left the room during survey administration and completed a brief survey about each participant in their class. Two weeks prior to data collection, parents received a letter describing the study. Children participated only if they assented and if their parents did not return a form exempting them from the study.

### 2.3 Measures

#### *Response Variable*

*Teacher-rated Aggression.* Teachers rated the degree to which 32 statements were true for each student (1 = Does not apply at all to 5 = Applies very much). One scale, *Aggression*, described students' tendencies to display direct forms of aggressive or defiant behavior (fights; breaks rules; teases classmates; trouble accepting authority; harms others; median  $\alpha = .91$ ).

#### *Predictor Variables*

*Grade.* Students were followed from 3<sup>rd</sup> through 7<sup>th</sup> grade.

*Semester.* Data were obtained twice each school year: once in the fall and once in the spring. Semester was coded 0 for the Fall semester (generally late October) and 1 for the Spring semester (generally late April or early May).

*Cohort.* There were three cohorts of youth in this study. At the first assessment, the oldest cohort (cohort = 1) was in 5<sup>th</sup> grade, the middle cohort (cohort = 2) was in 4<sup>th</sup> grade, and the youngest cohort (cohort = 3) was in 3<sup>rd</sup> grade.

*Gender.* Gender was coded 0 for girls and 1 for boys.

*Number of Group Members:* To identify youth who were members of the same social group, Cairns' Social Cognitive Mapping approach (Cairns, Cairns, Neckerman, Gest, & Garipey, 1988) was used. At each assessment, students were asked, "Are there some kids in your class who hang around together a lot?" In middle school, students were allowed to nominate groups of kids in the entire grade. Students were given space to write nine groups of up to 10 peers; they were asked to write as many or as few groups as they wanted. Collecting peer nominations is an alternative to observing who each child spends time with: Past research indicates that the number of times two students are named in the same group correlates with the number of times that dyad is observed to interact (Gest, Farmer, Cairns, & Xie, 2003). Peer nominations were aggregated into an  $n \times n$  co-nomination matrix, with  $n$  = number of students in the class or grade. The off-diagonal cells in the matrix indicated the number of times each dyad was named in the same group; the diagonal cells indicated the number of times each child was named to any group. Principal Components Analysis with Varimax rotation was used to identify groups (i.e., "components") of three or more youth with highly correlated nomination profiles (i.e., they were often named by classmates as being in the same group). Youth were members of any group on which they loaded  $> .32$ , ensuring that they shared 10% of the variance in their nominations with the group. Students who loaded highly on multiple components were members of multiple groups. Students who did not load on any component were isolates. *Number of group members* was the number of different peers who were in one or more groups with the child.

*Group Members' Peer-nominated Aggression.* Students were given a roster and asked to nominate classmates who fit 12 different descriptors. In elementary school, children were only allowed to nominate peers in their own class; in middle school, they were allowed to nominate peers within their entire grade. The mean number of nominations received for "starts fights" and "hits and picks on others" (median  $r = 0.94$ ) was divided by the number of nominators in the class (in elementary school) and grade (middle school). The different number of nominators led the absolute size of these scores to vary considerably between elementary and middle school. Therefore, mean proportion of aggressive nominations was standardized within grade. At each assessment, the average standardized peer-nominated aggression of a student's group members was calculated. Group members' aggression was treated as missing for youth who were not in any group at a given assessment. Fewer than 4% of students were not in a group at any given assessment.

Table 1 summarizes the variables included in the current analyses. The levels for categorical variables and possible range for continuous variables are noted.

*Table 1. Summary of Variables*

Variable Description	Variable Name	Variable Type	Level or Possible Range
<i>Predictor Variables</i>			
Gender	SexF	Categorical	0 = Boy 1 = Girl
Cohort in study	CohortID	Categorical	1 = Youngest, 2 = Middle, 3 = Oldest
Grade	Grade	Continuous	3rd – 7th grade
Semester	SemesterF	Categorical	0 = Fall, 1 = Spring
Number of group members	Sz_gf	Continuous	0 - 31
Group Members' Average Aggression	gfavzppagg	Continuous	$Z \approx -3 - +3$
<i>Response Variable</i>			
Teacher-rated Aggression	tagg	Continuous	1.00 - 5.00

## 2.4 Study Design

### 2.4.1 Data Collection Design

Data were collected using a cohort-sequential longitudinal design. In this type of research design, data are obtained from multiple cohorts of participants over time. The current study focuses on three cohorts of students: At Time 1 (October 2001), the youngest cohort was in 3<sup>rd</sup> grade, the middle cohort was in 4<sup>th</sup> grade, and the oldest cohort was in 5<sup>th</sup> grade. All three cohorts of students were then followed until they completed 7<sup>th</sup> grade, which was May 2004 for the oldest cohort, May 2005 for the middle cohort, and May 2006 for the youngest cohort. Table 2 summarizes when data were obtained from each cohort.

*Table 2. Data Collection Summary*

Cohort	Time 1	Time 2	Time 3	Time 4	Time 5	Time 6	Time 7	Time 8	Time 9	Time 10
	Oct. '01	May '02	Oct. '02	May '03	Oct. '03	May '04	Oct. '04	May '05	Oct. '05	May '06
Youngest	3rd	3rd	4th	4th	5th	5th	6th	6th	7th	7th
Middle	4th	4th	5th	5th	6th	6th	7th	7th	-	-
Oldest	5th	5th	6th	6th	7th	7th	-	-	-	-

Researchers use cohort-sequential designs for several reasons:

First, these designs provide replications to test whether effects are unique to a particular group of individuals or whether they are more robust and generalizable across multiple groups of people. In the current study, it is possible to test whether peer context is similarly associated with child aggression across all three cohorts. When results are replicated, it suggests that the results are not solely dependent on a particular, unique, group of students.

Second, cohort-sequential designs can be an efficient way to examine a large age interval within a shorter period of time. For example, in the current study, it is possible to describe developmental trajectories from 3<sup>rd</sup> through 7<sup>th</sup> grade, even though only some

of the participants were assessed when they were in 3<sup>rd</sup> grade. In fact, data collection could have ended after Time 6 while still allowing this grade range to be examined.

Finally, cohort-sequential designs make it possible to test for generational differences in how variables are related to each other. For example, are peers more influential for 5<sup>th</sup> graders in 2000 than they were for 5<sup>th</sup> graders in 1960? Although interesting, this use of cohort-sequential designs is not possible to analyze with the current data.

## 2.4.2 Data Analytic Design

Before analyzing data from a cohort-sequential research design, it is necessary to consider how to represent time. In some cases, time-in-study (i.e., Time 1, Time 2, Time 3...) might be the best option. For example, in a study examining the impact of an intervention given between Time 1 and Time 2, time-in-study may be the most appropriate representation of time: Time 1 is a pre-test, occurring prior to the intervention; Time 2 is an immediate post-test, occurring after the intervention; the remaining times are longitudinal follow-ups to examine whether the intervention effects fade over time. In this type of study, the relationship to the intervention is likely more important than the fact that data was collected from multiple cohorts of students.

Alternatively, in other studies it may be appropriate to realign the data using a different metric of time, such as grade or age. For example, at Time 5 and Time 6 in the current study, the three cohorts of students were in 5<sup>th</sup>, 6<sup>th</sup>, and 7<sup>th</sup> grade. Children's peer experiences may be very different across these three grades. In this school district, 5<sup>th</sup> graders were in self-contained classrooms with a single teacher and the same group of classmates the entire day. Their interactions with students in other classrooms were minimal. In 6<sup>th</sup> grade, students were in middle school and remained with the same classmates for most of the day, but they had different teachers for different subjects. There were more opportunities for cross-classroom peer interactions than in 5<sup>th</sup> grade. Finally, in 7<sup>th</sup> grade, students switched peers and teachers for every class.

In the current study, grade, rather than time-in-study, is a more appropriate metric of time. Table 3 summarizes how the data was realigned prior to analyses.

*Table 3. Data Analytic Summary*

<b>Cohort</b>	<b>Fall</b>	<b>Spring</b>	<b>Fall</b>	<b>Spring</b>	<b>Fall</b>	<b>Spring</b>	<b>Fall</b>	<b>Spring</b>	<b>Fall</b>	<b>Spring</b>
Youngest	3rd	3rd	4th	4th	5th	5th	6th	6th	7th	7th
Middle	-	-	4th	4th	5th	5th	6th	6th	7th	7th
Oldest	-	-	-	-	5th	5th	6th	6th	7th	7th

An example of the SAS syntax to realign the data is provided in Figure 2 below. Figure 2a shows a part of the data before it is realigned (when the variables are listed by time in the study rather than grade). Figure 2b shows the data after it is realigned. For example, student 6110 is in the oldest cohort (cohort 1). At Time 1, the oldest cohort was in fall semester of 5<sup>th</sup> grade. In the realigned data set, the value of pp1aggr for student 6110 is now under ppaggr5F. Student 6110 has no data for ppaggr3F –

ppaggr4S, as no data was collected from this student prior to 5<sup>th</sup> grade. Student 4613 is in the youngest cohort (cohort 3). At Time 1, the youngest cohort was in the fall semester of 3<sup>rd</sup> grade. In the realigned data set, the value of pp1aggr for student 4613 is now under ppaggr3F.

```

/* Realign the data so that variables are given by grade rather than wave
   To do this, variables are renamed separately by cohort
   For example, cohort 3 is in the fall of 5th grade at time 5 whereas
   cohort 2 is in the fall of 5th grade at time 3 */

DATA Work.Example2; SET Work.Example2;

IF cohort=3 THEN DO;
ppaggr3F = pp1aggr; ppaggr3S = pp2aggr;
ppaggr4F = pp3aggr; ppaggr4S = pp4aggr;
ppaggr5F = pp5aggr; ppaggr5S = pp6aggr;
ppaggr6F = pp7aggr; ppaggr6S = pp8aggr;
ppaggr7F = pp9aggr; ppaggr7S = pp10aggr; END;

IF cohort=2 THEN DO;
ppaggr4F = pp1aggr; ppaggr4S = pp2aggr;
ppaggr5F = pp3aggr; ppaggr5S = pp4aggr;
ppaggr6F = pp5aggr; ppaggr6S = pp6aggr;
ppaggr7F = pp7aggr; ppaggr7S = pp8aggr; END;

IF cohort=1 THEN DO;
ppaggr5F = pp1aggr; ppaggr5S = pp2aggr;
ppaggr6F = pp3aggr; ppaggr6S = pp4aggr;
ppaggr7F = pp5aggr; ppaggr7S = pp6aggr; END; RUN;

```

Figure 1. Sample SAS Syntax to Realign Data by Grade Rather Than Time-in-Study

a)

scanid	cohort	pp1aggr	pp2aggr	pp3aggr	pp4aggr	pp5aggr	pp6aggr	pp7aggr	pp8aggr	pp9aggr	pp10aggr
6110	1	0.114	0.125	0.008	0.004	0.02	0.021	.	.	.	.
4613	3	0.114	0.022	0	0	0	0	0	0	0	0
4609	3	0.114	0.304	0.194	0.389	0.091	0.079	0.003	0	0.007	0.011
5104	2	0.121	0.15	0.053	0.167	0.081	0.039	0.052	0.035	.	.

b)

scanid	cohort	ppaggr3F	ppaggr3S	ppaggr4F	ppaggr4S	ppaggr5F	ppaggr5S	ppaggr6F	ppaggr6S	ppaggr7F	ppaggr7S
6110	1	.	.	.	.	0.114	0.125	0.008	0.004	0.02	0.021
4613	3	0.114	0.022	0	0	0	0	0	0	0	0
4609	3	0.114	0.304	0.194	0.389	0.091	0.079	0.003	0	0.007	0.011
5104	2	.	.	0.121	0.15	0.053	0.167	0.081	0.039	0.052	0.035

Figure 2. Subset of the (a) Original Data and (b) Data After it is Realigned

### 3 Analytic Approach

One of the assumptions of the general linear model is that the residuals are independent (i.e., uncorrelated). When data are from the same person, however, this assumption is typically violated: people tend to be more similar to themselves than they are to other people. For example, if 10 people are measured on three occasions, we do

not have 30 completely independent observations; measurements within a person are likely to be correlated. In other words, the effective sample size is less than 30 (how much less depends on just how highly correlated observations are within a person). An analysis that ignores the correlations among the observations tends to underestimate the standard errors, which in turn can lead to the wrong conclusions.

In the current study, students' aggression and peer group context were measured on multiple occasions (Range: 2-10 occasions,  $M = 7.2$  occasions). Therefore, to analyze the relationship between students' peer group context and their aggression, an analytic approach that can account for the non-independence of these observations is needed.

### 3.1 Repeated Measures ANOVA

One approach that is often used to account for correlated errors is repeated measures ANOVA (Hedeker & Gibbons, 2006). The model for a multiple-sample repeated measures ANOVA is given by:

$$y_{hij} = \mu + \gamma_h + \tau_j + (\gamma\tau)_{hj} + \pi_{i(h)} + e_{hij}$$

Here, the response variable  $y$  for person  $i$  in group  $h$  at time  $j$  is a function of the overall mean of variable  $y$  ( $\mu$ ), a fixed effect of being in group  $h$  ( $\gamma_h$ ), a fixed effect due to time  $j$  ( $\tau_j$ ), a fixed group by time interaction  $(\gamma\tau)_{hj}$ , a random effect for subject  $i$  nested in group  $h$  ( $\pi_{i(h)}$ ), and a random error for person  $i$  in group  $h$  at time  $j$ . In repeated measures ANOVA, the non-independence of observations within a person is captured by the random effect of subject.

In the most restrictive case of repeated measures ANOVA, compound symmetry is assumed (Hedeker & Gibbons, 2006): compound symmetry exists when the variance of  $y$  is the same at each assessment and when the covariances between observations at different assessments are also equal. The slightly less restrictive assumption of sphericity is met if the variances of the differences between observations at any two assessments are equal. Unfortunately, sphericity is unlikely to hold in longitudinal data because observations that are closer in time tend to be more highly correlated. Furthermore, if groups undergo different treatments or change differently over time, the variance in  $y$  is likely to increase. It is possible to relax the assumption of sphericity by specifying more general variance-covariance structures. However, when these alternative structures are used, balanced data is required: in other words, all participants must be observed at the same assessments and participants who are missing data at any assessments are discarded.

### 3.2 Multilevel Modeling (MLM)

An alternative approach is to use a multilevel modeling (MLM) strategy,<sup>1</sup> which treats time as nested within individuals (Hedeker & Gibbons, 2006; Singer & Willet, 2003; Snijders & Bosker, 1999). As with repeated measures ANOVA, MLM can account for

---

<sup>1</sup> Multilevel models have also been referred to as hierarchical linear models, random coefficients or effects models, mixed-effects regression models, and variance component models (Hedeker & Gibbons, 2006).



the non-independence of errors that occurs when multiple observations of a variable are obtained from the same person. The primary difference is in *how* the non-independence is specified. In ANOVA models, the random effect for subjects captures differences between participants in terms of their overall mean. Within an ANOVA model, however, any between-person differences in patterns of change over time cannot be disentangled from the overall error term.

In contrast, within a multilevel model, participants can differ from each other not only in their overall mean but also in how they change across time. These variations are modeled by adding random subject effects into a standard regression model, in which time is treated as a continuous variable. Rather than relying on a single intercept and slope to describe how the outcome changes over time, participants are assigned their own subject-specific intercept and slope. In other words, students may differ in their initial level of the outcome and they may differ in how they change over time: some students may increase in the outcome, other students may increase at a faster rate, others may not change at all, and still others may decrease in the outcome over time.

One advantage of multilevel models is that because time is treated as continuous, they can easily accommodate unbalanced data in which students are not all measured at the same assessments (Hedeker & Gibbons, 2006; Singer & Willet, 2003). As noted above, the current study utilized a cohort-sequential design. Therefore, some students did not have data in 3<sup>rd</sup> or 4<sup>th</sup> grade (see Table 1). In addition, not all students were present at every wave; at any given assessment, some students were absent, had moved, or were exempt. Using MLM, it was possible to use data from all 427 students. Students contributed information for effect estimation only at times when they participated. Although all of the students in the current study participated at two or more assessments, an MLM analysis could even include data from students who only participated at one assessment: these students would contribute to the estimation of between-person effects, but not within-person effects.

### 3.2.1 Specifying the Multilevel Model

In multilevel modeling, the within-person effects are referred to as Level 1 effects and the between-person effects are referred to as Level 2 effects. A general multilevel model is given by:

#### Level 1 Model:

$$\text{Response Variable}_{ti} = \beta_{0i} + \beta_{1i}(\text{time}) + e_{ti}$$

#### Level 2 Model:

$$\beta_{0i} = \gamma_{00} + \zeta_{0i}$$

$$\beta_{1i} = \gamma_{10} + \zeta_{1i}$$

In the Level 1 model, the response variable for person  $i$  at time  $t$  is a function of a person-specific intercept,  $\beta_{0i}$ , a person-specific effect of time,  $\beta_{1i}$ , and a random residual

error,  $e_{it}$ . In turn, the person-specific intercept is a function of an overall grand mean,  $\gamma_{00}$ , and a student's random deviation from that mean,  $\zeta_{0i}$ ; the person-specific slope is a function of the overall slope for that population,  $\gamma_{10}$ , and a student's random deviation from that slope,  $\zeta_{0i}$ .

Essentially, each person is assigned their own regression line: the Level-2 random errors describe how the person-specific regression lines vary from the population-level regression line and the Level-1 random errors describe how each person varies from their own regression line at a specific assessment. Additional variables can be added to the model at either level: Time-varying variables can be added at Level 1 to try to reduce the residual error; they attempt to explain why people are off their predicted regression lines (e.g., Are people more aggressive at assessments when they have fewer friends?) Time-invariant (or person-level) variables can be added at Level 2 to try to reduce the intercept or slope variance; they attempt to explain why a person's regression line is different from the population regression line (e.g., Do boys start off more aggressive than girls? Do boys increase in aggression faster than girls?)

### 3.2.2 Group-mean centering

Multilevel models to analyze longitudinal data are being used more often; however, many of these applications do not take full advantage of the strengths of this analytic approach. In particular, many applications of MLM include time-varying covariates, but do not try to disentangle the between-person and within-person effects of these covariates. For example, someone may receive a high score on depression at a particular point in time because they are generally a depressed person (a between-person effect). Alternatively, a normally upbeat person may receive a high score on depression because a close friend recently died (a within-person effect). These effects may be independently associated with a person's outcome (additive effects) or they may interact (e.g., within-person variations in depression may only matter for those who are not generally depressed).

Between-person and within-person effects of a time-varying covariate can be disentangled by using group-mean centering (Raudenbush & Bryk, 2002). This approach, also known as within-person centering (Singer & Willet, 2003), subtracts students' own cross-time mean from their score at each assessment. Group-mean centering and the inclusion of the cross-time (between-person) mean at Level 2 ensures that time-varying effects only reflected within-person variability across time, over and above any stable (unobserved) individual differences and allowed us to estimate how changes in peer context were related to changes in adjustment (see Kim, McHale, Crouter, & Osgood, 2006 for a similar application).

Figure 3 illustrates what can be gained from group-mean centering. Group members' aggression was measured for two students across 10 assessments. The time-varying values for group members' aggression are shown with red squares for student 1 and with green triangles for student 2. For example, at the first assessment, the average aggression of student 1's group members was  $Z = 1.0$  and the average aggression of student 2's group members was  $Z = -0.4$ . Group members' aggression varies across

assessments for both students. However, on *average* student 1 affiliates with highly aggressive groups ( $Z = .9$ ) and on *average* student 2 affiliates with non-aggressive groups ( $Z = -0.2$ ).

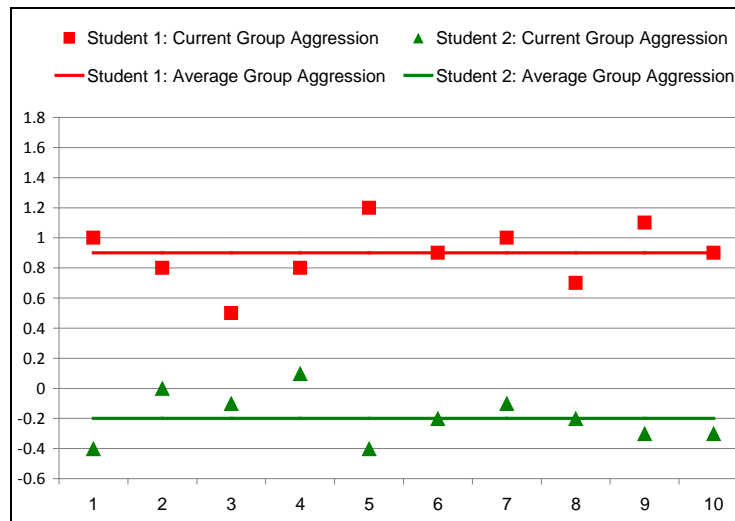


Figure 3. Separating Between-Person from Within-Person Peer Context Effects

Table 4 shows part of a data set corresponding to the data shown in Figure 3. At Time = 1, student 1 affiliated with highly aggressive group members (GroupAgg = 1). Within the data alone, it is unclear whether the value is high because student 1 *usually* affiliates with aggressive groups or whether this value is an anomaly for this student. Group-mean centering allows us to separate these two effects into a between-person effect (i.e., BPGroup) and a within-person effect (i.e., WPGroupLag0). The between-person effect represents the student's average group aggression. Because average group aggression is a time-invariant predictor; (i.e., it is a person-level effect), the value of BPGroup for student 1 is always 0.9. In contrast, the within-person effect indicates how the student's current group aggression deviates from the student's average group aggression. Current group aggression is a time-varying predictor; positive values indicate that a student is affiliating with *more aggressive group members than usual*. For example, at Time 1, student 1's group members are slightly more aggressive than usual (WPGroupLag0 = +0.1) whereas at Time 2, student 1's group members are slightly less aggressive than usual (WPGroupLag0 = -0.1).

Table 4. Example Data Set Showing Variables after Group-Mean Centering

Student	Time	GroupAgg	BPGroup	WPGroupLag0	WPGroupLag1
1	1	1	0.9	0.1	
1	2	0.8	0.9	-0.1	0.1
1	3	0.5	0.9	-0.4	-0.1
1	4	0.8	0.9	-0.1	-0.4
1	5	1.2	0.9	0.3	-0.1
1	6	0.9	0.9	0	0.3
1	7	1	0.9	0.1	0

After group-mean centering, it is also possible to create a variable to indicate whether a student's group members were more or less aggressive than usual at the *past* assessment. Past group aggression (i.e., WPGroupLag1) is a time-varying predictor. Because we do not know the extent to which student 1's group was aggressive before the first assessment, WPGroupLag1 is treated as missing at Time 1.

Once these new variables are created, the independent contributions of average, current, and past group aggression on a student's own aggression can be evaluated. If students who typically affiliate with aggressive groups (e.g., student 1) are *on average* more aggressive than students who typically affiliate with non-aggressive groups (e.g., student 2), then we would expect the coefficient corresponding to BPGroup to be positive. If students are more aggressive at assessments when they have more aggressive group members than usual, then we would expect the coefficient corresponding to WPGroupLag0 to be positive. Similarly, if students are more aggressive at assessments when they had more aggressive group members than usual at the *past* assessment, then we would expect the coefficient corresponding to WPGroupLag1 to be positive.

### 3.2.3 Summary

In sum, given the current design and hypotheses, a multilevel modeling strategy was selected as the better approach for the current study. In particular, the current data were from a cohort-sequential longitudinal study, which means that there was some missing data by design (i.e., the middle cohort was not observed in 3<sup>rd</sup> grade; the oldest cohort was not observed in 3<sup>rd</sup> or 4<sup>th</sup> grade). There was also some missing data due to students being absent, being exempt, or moving from the school. Repeated measures ANOVA approaches that do not assume compound symmetry (which is unlikely to hold in the current study) exclude participants who have any missing data. Complete case analysis would result in a considerably smaller, and likely biased, sample (i.e., only those students in the youngest cohort who participated in all 10 assessments would be included). Multilevel modeling takes better advantage of the full data set, by including students in the estimation of effects whenever they had data. In addition, the multilevel modeling framework allows questions about how between-person and within-person peer group context independently (or interactively) contribute to a student's individual aggression.

## 4 Exploratory Data Analysis

### 4.1 Preparing for Longitudinal Data Analyses

Cross-sectional analyses and some ANOVA analytic approaches utilize a person-level data set (Hedeker & Gibbons, 2006; Singer & Willet, 2003). In a person-level data set, the data for each student is found on a single row; repeated measurements are represented as separate variables (i.e., columns). For example, aggression in the fall of 3<sup>rd</sup> grade is one variable and aggression in the spring of 3<sup>rd</sup> grade is another variable. In a person-period data set, however, each student has as many rows as there are assessments and repeated measurements are represented as different observations of

a single variable. For example, row 1 for each student would have their aggression in the fall of 3rd grade and row 2 would have their aggression for the spring of 3rd grade. Time-invariant variables are constant across rows for each student. (e.g., if a student was a boy (sex=1), then each row under the sex variable for that student would equal 1).

Figure 4 below illustrates what a person-level data set looks like compared to a person-period data set. In the person-level data set, there are 10 variables to represent each student's data pattern across the study (datpat3F through datpat7S). In the person-period data set, however, data pattern is represented as a single variable. Each row contains data for a different assessment. For example, the first row for student 4801 has data for the fall of 3<sup>rd</sup> grade (grade = 3, semester = 0) and the second row has data for the spring of 3<sup>rd</sup> grade (grade = 3.5, semester = 1).

As noted above, in MLM, a participant can have incomplete data on some occasions but still be included in the analysis. Setting up the person-period data set allows us to see how this occurs. For example, in Figure 4, student 4523 did not participate in the spring of 5<sup>th</sup> grade (his data pattern for this assessment, datpat5S = 0, indicates that he was not enrolled at the school). In repeated measures ANOVA, the entire row for this student would be removed. In an MLM analysis, however, only that assessment is removed from the analysis (the row corresponding to grade = 5.5 for student 4523 in the person-period data set).

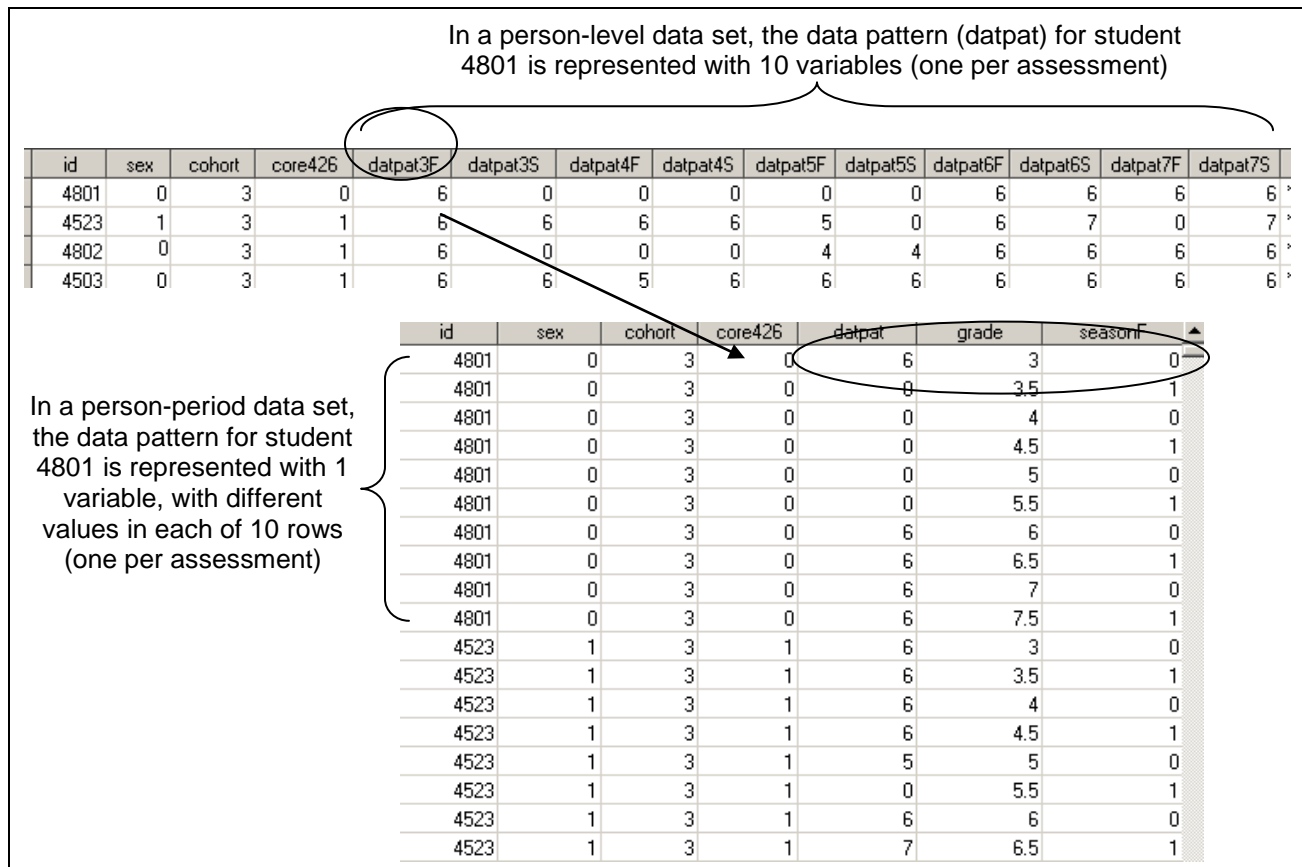


Figure 4. Comparing Person-Level and Person-Period Data Sets

## 4.2 Examining Missing Data

In longitudinal data analyses, once the data set is set up properly, the next step is to examine patterns of missing data. If the sample changes drastically over time, then inferences about relationships among variables can be affected. In the current study, students could be missing at a particular assessment for several reasons: they might have had moved, be temporarily suspended (or sent to an alternative school), or exempt (i.e., their parents had sent a consent form back asking that their child not participate). Most of the missing data in the current study was due to the latter two reasons (temporarily suspended or exempt); few students moved permanently.

Table 5 summarizes the participation rates for the core sample of students at each assessment. Because of the cohort-sequential design of the study, participation rates are reported separately by cohort. Because gender is expected to moderate the relationships between having aggressive group members and children's adjustment, participation rates are also reported separately by gender. Participation rates ranged from 76% (Fall of 5<sup>th</sup> grade for girls in the youngest cohort) to 98% (Spring of 4<sup>th</sup> grade for girls in the middle cohort; Fall of 4<sup>th</sup> grade for boys in the middle cohort). Overall, 67% percent (N = 286) of the sample participated at all waves (6-10 waves, depending on the cohort) and 95% (N = 406) participated at four or more waves.

*Table 5. Participation Rates at Each Assessment as a Function of Gender and Cohort*

Gender	Cohort	N	3rd		4th		5th		6th		7th	
			Fall	Spring	Fall	Spring	Fall	Spring	Fall	Spring	Fall	Spring
Girls	Youngest	71	67 (0.94)	68 (0.96)	57 (0.80)	55 (0.77)	54 (0.76)	60 (0.85)	58 (0.82)	59 (0.83)	57 (0.80)	57 (0.80)
	Middle	64	-	-	61 (0.95)	63 (0.98)	62 (0.97)	62 (0.97)	59 (0.92)	61 (0.95)	60 (0.94)	59 (0.92)
	Oldest	58	-	-	-	-	53 (0.91)	56 (0.97)	54 (0.93)	52 (0.90)	49 (0.84)	49 (0.84)
Boys	Youngest	75	73 (0.97)	72 (0.96)	69 (0.92)	68 (0.91)	68 (0.91)	66 (0.88)	68 (0.91)	63 (0.84)	62 (0.83)	60 (0.80)
	Middle	83	-	-	81 (0.98)	79 (0.95)	79 (0.95)	79 (0.95)	77 (0.93)	75 (0.90)	71 (0.86)	72 (0.87)
	Oldest	76	-	-	-	-	73 (0.96)	68 (0.89)	69 (0.91)	67 (0.88)	67 (0.88)	66 (0.87)
Total		427	140	140	268	265	389	391	385	377	366	363

Within each assessment, any students who were enrolled and whose parents did not exempt them from the study were allowed to participate (regardless of whether or not they were in the core sample of 427 students). Over 90% of students present at the school at any given assessment participated.

The amount of data missing further varied across response variables. At each assessment, group members' average peer-nominated aggression was available for all students who both participated at that assessment and who had at least one group member who participated at that assessment. Teacher-report data were occasionally unavailable when a teacher failed to complete a survey for a student. In 7<sup>th</sup> grade, data from one teacher was thrown out in the older two cohorts because that teacher only answered a handful of items for each of his students. The exact sample size for each variable is reported in Table 6 on the next page.

### 4.3 Descriptive Information for Response Variables

Descriptive information was obtained for three of the time-varying variables in the study: teacher-reported aggression, number of group members, and average group members' peer-nominated aggression. The syntax to obtain the descriptive information is provided in Figure 5. The sample size, mean, SD, skew, and kurtosis for each of these variables is given in Table 6 below.

```

/* Obtain descriptive information about the response variable */
/* Results will be obtained separately for each grade */

PROC SORT DATA=&udatafile.;
BY Grade;          * Sort the file by grade;
RUN;

PROC MEANS DATA=&udatafile.
MEAN STD SKEWNESS KURTOSIS N;
WHERE core=1;      * Only include those in the core sample;
BY Grade;          * Obtain descriptive info separately by grade;
VAR tagg;          * Identify response variables;
RUN;

/* Obtaining descriptive information about the predictor variables */

PROC MEANS DATA=&udatafile.
MEAN STD SKEWNESS KURTOSIS N MIN MAX;
WHERE core=1;      * Only include those in the core sample;
BY Grade;          * Obtain descriptive info separately by grade;
VAR sz_GM GMavzppagg; * Identify predictor variables;
RUN;

```

Figure 5. SAS Syntax to Obtain Descriptive Information about the Variables

Table 6. Sample Size, Mean, SD, Skew, and Kurtosis for Time-varying Variables

	3rd		4th		5th		6th		7th	
	Fall	Spring	Fall	Spring	Fall	Spring	Fall	Spring	Fall	Spring
<i>Teacher-reported Aggression</i>										
N	140	140	264	262	380	380	382	376	344	339
Mean	1.42	1.61	1.38	1.37	1.51	1.58	1.43	1.51	1.57	1.58
SD	0.80	0.89	0.59	0.55	0.74	0.76	0.72	0.73	0.65	0.76
Skew	2.79	1.98	1.99	2.26	1.82	1.60	2.01	1.63	1.60	1.60
Kurtosis	8.16	3.50	4.10	6.43	3.28	2.23	3.73	2.18	3.51	2.33
<i>Number of group members</i>										
N	140	140	268	265	389	391	385	377	366	363
Mean	5.68	4.46	5.80	5.09	5.43	5.11	9.15	8.90	9.36	10.03
SD	3.13	2.10	3.24	2.56	2.43	2.48	4.85	4.97	5.32	5.10
Skew	0.41	0.06	0.88	0.78	0.03	0.63	0.90	0.83	0.84	0.76
Kurtosis	-0.58	-0.52	1.71	0.15	-0.39	0.34	0.61	0.68	0.58	1.02
<i>Average Group Members' Peer-nominated Aggression</i>										
N	137	136	264	265	377	384	383	373	361	358
Mean	0.05	0.06	0.05	0.05	0.05	0.06	0.01	0.01	0.01	0.01
SD	0.06	0.07	0.04	0.05	0.05	0.07	0.02	0.02	0.01	0.01
Skew	2.43	2.84	1.04	1.26	1.64	2.23	3.32	3.33	3.71	3.84
Kurtosis	6.65	9.77	1.50	1.06	2.60	6.45	11.52	11.74	15.49	23.42

### 4.3.1 Transforming and Centering Variables

Not surprisingly, teacher-rated aggression was skewed (Range = 1.60 to 2.79 across assessments): most students receive low aggression scores from their teachers and only a handful of students receive very high scores of aggression. To reduce skew, teacher-rated aggression was log transformed.

Centering the predictor variables facilitates interpretation of the results. Grade was centered at 5<sup>th</sup> grade, because that is the earliest grade in which all three cohorts were assessed. The number of group members was centered at 5, because that was the average number of group members that students had in the fall of 5<sup>th</sup> grade. Therefore, when the centered variables are included in the model, the intercept indicates the average aggression of a student in the fall of 5<sup>th</sup> grade (i.e., grade – 5 = 0) who has an average number of group members (i.e., group size – 5 = 0).

The syntax to transform teacher-rated aggression and to center the predictor variables is given in Figure 6 below.

```

/* Centering Predictor Variables */
DATA &udatafile.; SET &udatafile.;
  lntagg=log(tagg);    * Log transform teacher-rated aggression;
  grade5 = (grade-5); * Center grade at 5th grade;
  sz_GM5 =(sz_GM-5);  * Center at 5 group members, Fall of 5th grade mean;
  sexF = sex;         * Rename sex variable to include reference category;

LABEL lntagg = "Natural log, teacher-rated aggression";
      sz_GM5 = "Number of group members, centered around mean = 5"
      sexF   = "Gender, girls = 0, boys = 1"
      grade5 = "Grade, centered at 5th grade"; RUN;

```

Figure 6. SAS Syntax for Transforming and Centering Variables

Group-mean centering, described in section 3.2 above, was used to disentangle between-person and within-person group members' peer-nominated aggression. The syntax to group-mean center group members' aggression is given in Figure 7 on the next page. In addition to group-mean centering, a variable for the within-person lag group members' aggression (i.e., group members aggression at the previous assessment) was created.

Below, the between-person effects are referred to as average group (girls: range = -0.51 to 0.89,  $M = -0.13$ ,  $SD = 0.24$ ; boys: range = -0.42 to 2.13,  $M = 0.14$ ,  $SD = 0.37$ ) and the within-person lag 0 effects as current group (girls: range = -1.34 to +2.68; boys: range = -2.31 to +4.78). A positive value for current group indicates that a student had more aggressive group members at that assessment than usual.



```

/* Group Mean Centering */
PROC SORT DATA=&datafile.; BY ID Grade; RUN;

PROC MEANS NOPRINT DATA=&datafile.; BY ID;
VAR GMavzppagg; OUTPUT OUT=PersonMeansGMpeer2
MEAN (GMavzppagg) = BPGMavzppagg; RUN;
  * Obtain the mean for the variable GMavzppagg;
  * OUTPUT the mean to a new file, PersonMeansGMpeer;
  * Save the MEAN of GMavzppagg as BPGMavzppagg;

DATA &datafile.; MERGE &datafile. PersonMeansGMpeer; BY ID;

WPGMavzppaggLag0 = GMavzppagg-BPGMavzppagg;
/* Create a new variable for the current (Lag0) group members' aggression
   This new variable is equal to the observed group aggression score for
   that assessment minus a person's average group aggression */

IDlag1 = LAG1(id);
  * Create a new variable that shifts the ID variable values down 1;

WPGMavzppaggLag1=LAG1(WPGMavzppaggLag0);
  * Create a new variable that shifts the WP variable values down 1;

IF ID=IDlag1 THEN WPGMavzppaggLag1=WPGMavzppaggLag1; ELSE WPGMavzppaggLag1=.;

/* The result of lagging the ID and WP variables is that the last observation
for each person is shifted down to become the first observation for the next
person. The values that are shifted to the next person are then deleted and
left as missing (we do not know what a person's WP aggression was before they
entered the study). The example below shows what a small sample of the data
would look like. Values of WPLag1 in parentheses are set to missing.

      ID  IDlag1  WPLag0  WPLag1
      1    .      .8      .
      1    1      -.3     .8
      1    1      -.5     -.3
      2    1      .6     (-.5)
      2    2      -.2     .6
      2    2      -.5     -.2
      2    2      .1     -.5
      3    2      -.3     (.1)
      3    3      .3     -.3      */

/* Label the new variables */
LABEL BPGMavzppagg = "Between-person group peer-rated aggression"
      WPGMavzppaggLag0 = "Within-person group peer-rated aggression Lag=0"
      WPGMavzppaggLag1 = "Within-person group peer-rated aggression Lag=1";
RUN;

```

Figure 7. Syntax for Group-Mean Centering of Group Members' Aggression

#### 4.4 Examining Raw Individual Change Trajectories

Examining the raw trajectories of aggression for each student provides a sense of whether and how individuals change in teacher-rated aggression over time, it helps to.

The syntax to obtain these raw individual change trajectories (or “spaghetti plots”) is given in Figure 8 and the individual change trajectories are given in Figure 9.

```

/* Set system options for graphics */
GOPTIONS RESET=ALL NOBORDER FTEXT=TRIPLEX FTITLE=TRIPLEX VSIZE=6.0in
HSIZE=7.0in INTERPOL=JOIN GSFNAME=OUTGRAPH DEV=EMF GSFMODE=REPLACE;

/**** Teacher-rated aggression for Girls *****/
FILENAME outgraph "&filesave.\Spaghetti Plot_Teacher-ratedAgg_Girls.bmp";
* Save the plot as bitmap to location specified by global function;

TITLE JUSTIFY=CENTER HEIGHT=1.5 "Natural Log Teacher-rated Aggression:Girls";
* Center the text in the title & use HEIGHT = 1.5 for the text;

PROC GPLOT DATA=&datafile.;
WHERE sex=0 & core=1; * Only plot girls in the core sample;
  AXIS1 LENGTH=4.5in LABEL=(HEIGHT = 1.4 ANGLE=90
    "Natural Log Teacher-rated Aggression") ORDER=(0 TO 2 BY .5);
  * AXIS1 = Y-axis;
  * Y-axis has a LENGTH = 4.5 inches (out of 6 inches above);
  * HEIGHT = 1.4 sets the size of the label;
  * ANGLE = 90 places the axis label at a 90 degree angle;
  * ORDER = indicates that the scale goes from 0 to 2 in increments of .5;
  AXIS2 LENGTH=5.5in LABEL=(HEIGHT = 1.4 "Grade") ORDER=(3 TO 7 BY .5);
  * AXIS2 = X-axis;
  * X-axis has a LENGTH = 5.5 inches (out of 7 inches above);
  * HEIGHT = 1.4 sets the size of the label;
  * ORDER = indicates that the scale goes from 3 to 7 in increments of .5;
  PLOT lntagg*Grade=ID / NOLEGEND VAXIS=AXIS1 HAXIS=AXIS2;
  * Plot natural log of teacher rated aggression vs. grade for each person;
  * Do not include a legend;
  * Obtain specifications for VAXIS (i.e., Y-axis) from AXIS1 from above;
  * Obtain specifications for HAXIS (i.e., X-axis) from AXIS2 from above;
RUN; QUIT;

```

Figure 8. SAS Syntax to Plot Student-Specific Trajectories of Teacher-rated Aggression

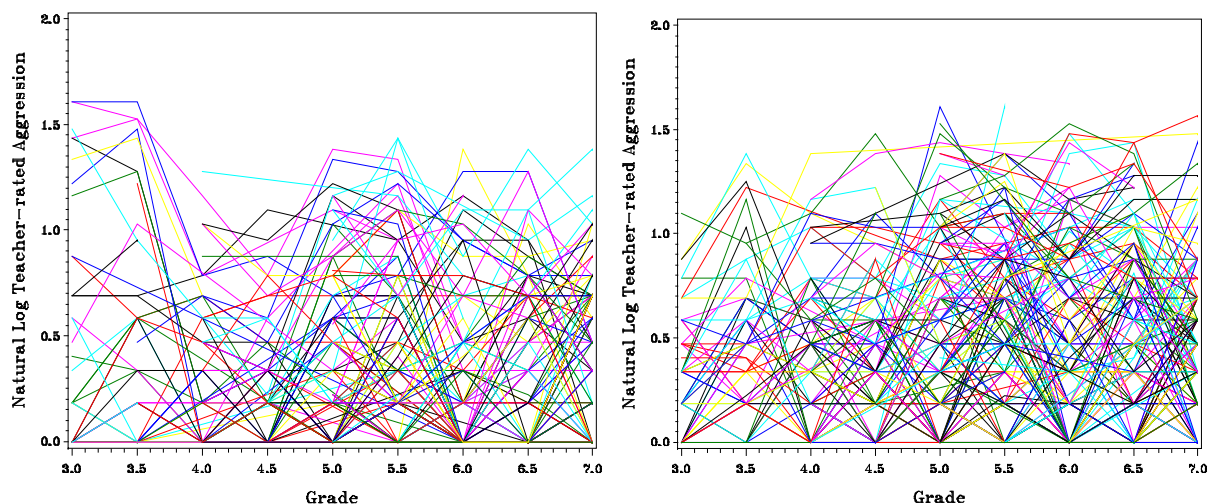


Figure 9. Observed Student-specific Aggression Trajectories for (a) Girls and (b) Boys

In these plots, each line represents a student. Considerable variation in teacher-rated aggression is evident; thus even though aggression is often assumed to be a stable individual characteristic, there were clear changes in aggression over time. In addition, the current analyses will focus on how changes in group members' aggression are associated with teacher-rated aggression. Therefore, student-specific variation in group members' was examined by adapting the syntax above. Figure 10 below displays these individual trajectories for group members' aggression and again, considerable within-person variation is evident.

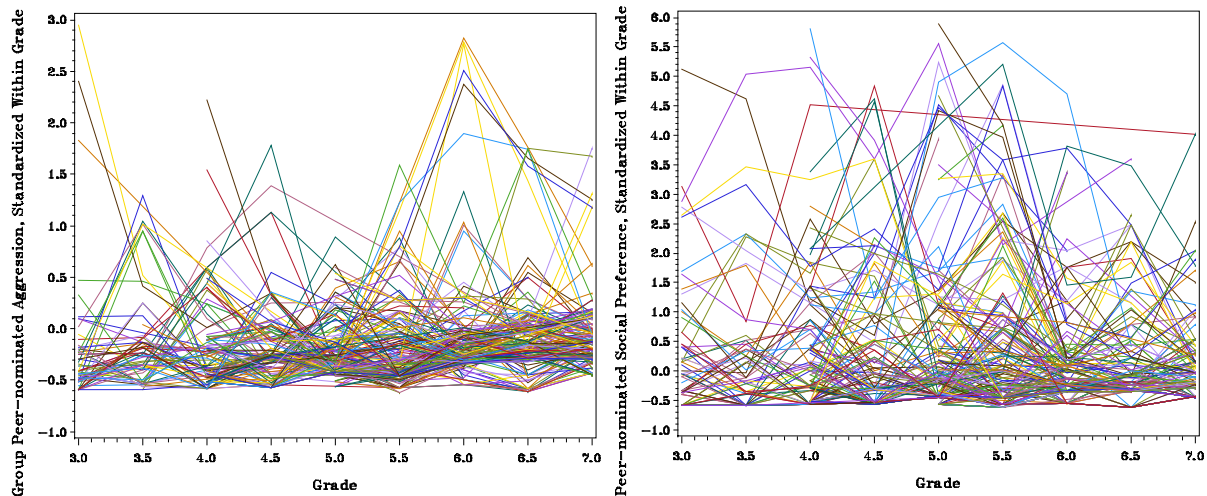


Figure 10. Observed Student-specific Average Group Members' Peer-nominated Aggression Trajectories for (a) Girls and (b) Boys

## 5 Analyses

A multilevel modeling (MLM) strategy was used to explore whether stable individual differences in peer context were associated with a student's average aggression and whether changes in peer context coincided with changes in aggression. For each outcome, a series of two-level models were estimated with SAS PROC MIXED, using the Satterthwaite method to estimate degrees of freedom.

### 5.1 Unconditional Means Model

The unconditional means model (Model 1) estimates the average value of an outcome in the absence of any predictor variables. The Level 1 and Level 2 models are given by:

Level 1 Model:

$$\text{Teacher-rated aggression}_{ti} = \beta_{0i} + e_{ti}$$

Level 2 Model:

$$\beta_{0i} = \gamma_{00} + \zeta_{0i}$$

The Level 1 model indicates that student  $i$ 's teacher-rated aggression at time  $t$  is equal to a student-specific intercept,  $\beta_{0i}$ , plus a residual error,  $e_{it}$ . In the Level 2 model, each student-specific intercept is equal to the overall grand mean across all students,  $\gamma_{00}$ , plus the student's random deviation from that mean.

Figure 11 illustrates the interpretation of the three parameters in the unconditional means model. The light gray dashed line illustrates the grand mean (i.e., the fixed intercept,  $\gamma_{00}$ ) for the entire sample; on average, teacher-rated aggression is low ( $\gamma_{00} = 1.4$ ). However, by including a random intercept, individual students are allowed to vary around this mean. Student  $i$  (darker, dotted line) has a mean teacher-rated aggression score of 2.1. The difference between this student's mean aggression and the grand mean is given by  $\zeta_{0i}$  (the Level 2 residual): here,  $\zeta_{0i}$  is positive because this student's mean aggression is higher than the grand mean for the sample. The difference between this student's mean aggression and his/her observed aggression at any assessment (black squares) is given by  $e_{it}$  (the Level 1 residual). Here, the student has a positive value of  $e_{it}$  in the spring of 7<sup>th</sup> grade (grade = 7.5). Because the residuals are treated as random (rather than fixed), we generally focus on the *variation* of these parameters, rather than their absolute values. The variation around the intercept is given by  $\sigma_{\zeta_{0i}}^2$  and the residual variation is given by  $\sigma_{e_{it}}^2$ .

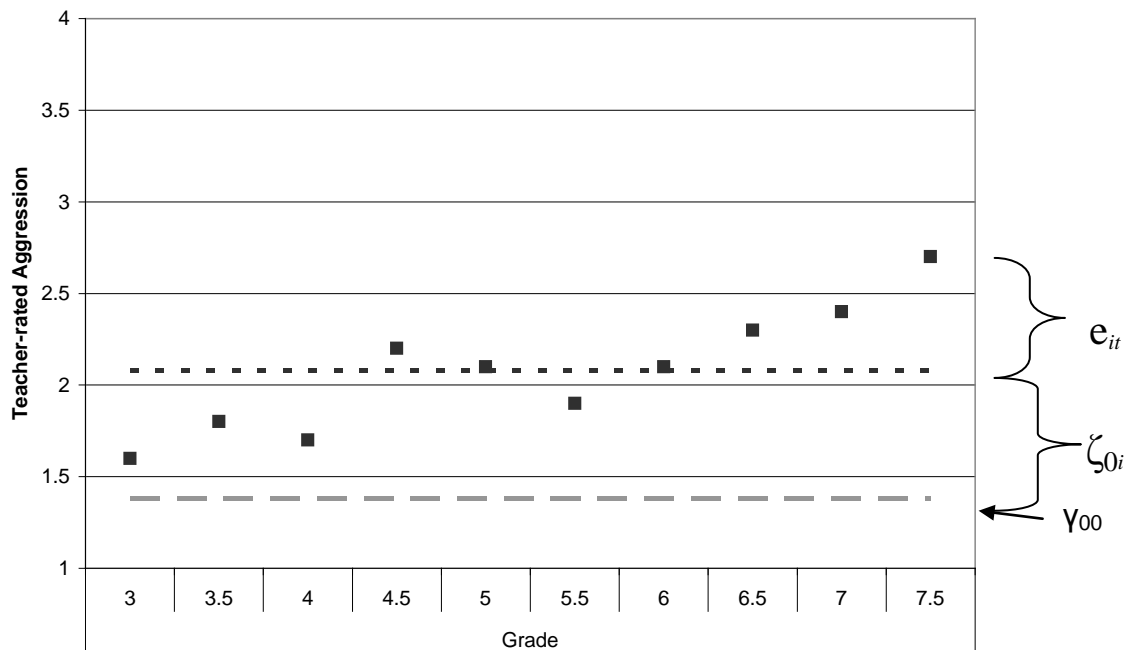


Figure 11. Illustration of the Parameters in the Unconditional Means Model

The SAS syntax to obtain the unconditional means model is provided in Figure 12 and partial output for this model is provided in Figure 13.

Figure 12. SAS Syntax for Estimating the Unconditional Means Model

```

TITLE "Unconditional Means Model";
PROC MIXED DATA = &udatafile. NOCLPRINT NOITPRINT COVTEST METHOD=REML;
  * NOCLPRINT: do not print the class level information;
  * NOITPRINT: do not print the iteration history;
  * COVTEST: print standard error, z-value & p-value for each covariance term;
  * REML: Use restricted maximum likelihood estimation;
WHERE NMISS(core)=0; * Only use observations from the core sample;
CLASS ID;          * Treat ID as a categorical variable;
MODEL lntag = / SOLUTION DDFM=Satterthwaite;
  * The only fixed effect in this model is the intercept;
  * Print the solution;
  * Use the Satterthwaite method to estimate the degrees of freedom;
RANDOM INTERCEPT / TYPE = UN SUBJECT=ID; RUN;
  * RANDOM INTERCEPT: include a random effect for intercept;
  * TYPE=UN: use an unstructured covariance matrix;
  * SUBJECT=ID: observations are nested within child, given in the data by ID;

```

Model Information					
Data Set			STATFP.STATAPR09		
Dependent Variable			lntag		
Covariance Structure			Unstructured		
Subject Effect			id		
Estimation Method			REML		
Residual Variance Method			Profile		
Fixed Effects SE Method			Model-Based		
Degrees of Freedom Method			Satterthwaite		
Dimensions					
Covariance Parameters			2		
Columns in X			1		
Columns in Z Per Subject			1		
Subjects			427		
Max Obs Per Subject			10		
Number of Observations					
Number of Observations Read			3440		
Number of Observations Used			3007		
Number of Observations Not Used			433		
Covariance Parameter Estimates					
Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr >  Z
UN(1,1)	id	0.08405	0.006633	12.67	<.0001
Residual		0.07262	0.002026	35.85	<.0001
Fit Statistics					
-2 Res Log Likelihood			1580.5		
AIC (smaller is better)			1584.5		
AICC (smaller is better)			1584.5		
BIC (smaller is better)			1592.6		
Solution for Fixed Effects					
Effect	Estimate	Standard Error	DF	t Value	Pr >  t
Intercept	0.3285	0.01496	414	21.96	<.0001

Figure 13. SAS Output for Unconditional Means Model

### 5.1.1 Examining the SAS Output

*Model Information.* It is useful to review this section of the output to confirm that the model was actually tested as planned. For example, the output indicates that the data set was StatFP.StatApr09 and that the dependent variable was Intagg.

*Dimensions.* This section of the output indicates that the tested model has two covariance parameters (i.e., the intercept variance and the residual variance). “Columns in X” indicates that the model has one predictor variable (i.e., the intercept) and “Columns in Z per subject” indicates that the model has one dependent variable. There were 427 subjects used in the analysis (i.e., the core sample was used), and the maximum number of observations for any subject was 10.

*Number of Observations.* There were 3440 rows of data corresponding to students in the core sample of 427 students. Of these, 3007 (87.4%) were used to estimate the current model. The remaining rows contained missing values (i.e., the student did not have a value for teacher-rated aggression at a particular assessment).

*Covariance Parameter Estimates.* This section of the output displays the estimates and Z-tests for the random effects. In the unconditional means model, there are two random effects: The random intercept, UN(1,1,), and the residual variance.

*Fit Statistics.* Once terms are added to the model, the fit statistics help to determine which model best describes the data. The REML deviance (-2 Res Log Likelihood), AIC, and BIC for the current model are all provided.

*Solution for Fixed Effects.* This section of the output displays the estimates and t-tests for the fixed parameters. In the unconditional means model there is only one fixed effect: intercept.

### 5.1.2 Interpreting the SAS Output

Across all assessments, the average natural log of teacher-rated aggression, given by the fixed intercept,  $\gamma_{00}$ , is 0.33 ( $t = 21.96$ ,  $p < .001$ ). In other words, teacher-rated aggression was on-average very low: 1.38 on the untransformed 1 to 5 scale. This value is consistent with teachers rating two out of the five items as “somewhat true” of their students.

Although the average teacher-rated aggression is low, there is significant variation around this average: the intercept variance is 0.08 ( $p < .001$ ). To describe the variation across people, this variance can be used to compute the range of scores in which 95% of the population is expected to fall. This range is calculated as the fixed effect  $\pm 1.96$  times the standard deviation of the corresponding random intercept. Here, 95% of the

population is expected to fall within the range of  $0.33 \pm 1.96 * \sqrt{0.08}$ , which translates into a 95% confidence interval of  $[1.00, 2.42]^2$  on the untransformed scale.<sup>3</sup>

In addition, it is also possible to calculate an intraclass correlation, or ICC from the unconditional means model. The ICC is essentially a within-person correlation: it describes the extent to which the Level 1 observations are correlated within the Level 2 units. Here, the ICC indicates the degree to which student's scores on teacher-rated aggression are correlated.

$$ICC = \frac{\textit{BetweenPerson Variation}}{\textit{BetweenPerson Variation} - \textit{WithinPerson Variation}} = \frac{\sigma_{\epsilon_{0i}}^2}{\sigma_{e_{it}}^2}$$

The ICC for teacher-rated aggression is  $0.08 / (0.07 + 0.08) = 0.53$ , which indicates that slightly more than half of the variation in teacher-rated aggression is between people. Typically, aggression is considered to be a highly stable individual characteristic, which would imply that most of the variation should be between people. However, in this sample almost half of the variation was within-people. This within-person variation is consistent with variation that was seen in the plots of observed teacher-rated aggression for individual students in Figure 9.

In sum, average teacher-rated aggression is low within this sample, but there is significant variation in average aggression across students. The considerable between-person and within-person variation in teacher-rated aggression suggests that it will be worthwhile to pursue both student-level (i.e., between-person) and time-varying (i.e., within-person) predictors of aggression. Both of these types of predictors will be explored below.

## 5.2 Unconditional Growth Model

Because the data are longitudinal, the next step after estimating an unconditional means model is to estimate an unconditional growth model. This model accounts for changes in the outcome that are due to time in the absence of any other predictors. There are two indicators of time in the current data: grade in school and semester (fall vs. spring). In addition to a linear effect of grade, it is also important to test whether there are any quadratic effects of grade. For example, it is possible that aggression increases across elementary school and into middle school, but then this increase may slow down across middle school. Theoretically, with up to 10 assessments per person, it would be possible to test higher order effects (e.g., cubic or quartic effect of grade). However, the inclusion of semester is a more interpretable metric of time that will likely capture any patterns beyond linear and quadratic effects.

<sup>2</sup> This range is truncated at 1.0, because values of teacher-rated aggression cannot be less than 1.0. The fact that the confidence interval falls below 1 likely reflects that the transformed variable is still somewhat skewed.

<sup>3</sup> This 95% C.I. given here should not be confused with confidence intervals that indicate precision of the coefficients. The 95% C.I. for the precision of the fixed effect is  $0.08 \pm 1.96$  times the standard error of the estimate, or  $0.08 \pm 1.96 * 0.0066$ , or  $[0.067, 0.092]$ .

Whether or not there is a population-level pattern of growth or decline of aggression, individual students may become more or less aggressive over time. To test whether there is any variation in trajectories of change across students, random effects for the linear and quadratic effects of grade can be tested. The Level 1 and Level 2 models to test these effects are given by:

Level 1 Model:

$$\text{Teacher-rated aggression}_{ti} = \beta_{0i} + \beta_{1i} (\text{grade}_{ti} - 5) + \beta_{2i} (\text{grade}_{ti} - 5)^2 + \beta_{3i} (\text{semester}) + e_{ti}$$

Level 2 Model:

$$\beta_{0i} = \gamma_{00} + \zeta_{0i}$$

$$\beta_{1i} = \gamma_{10} + \zeta_{1i}$$

$$\beta_{2i} = \gamma_{20} + \zeta_{2i}$$

$$\beta_{3i} = \gamma_{30}$$

The Level 1 model indicates that student  $i$ 's teacher-rated aggression at time  $t$  is equal to a student-specific intercept,  $\beta_{0i}$ , a student-specific effect of grade,  $(\beta_{1i})$ , grade<sup>2</sup> ( $\beta_{2i}$ ), and semester ( $\beta_{3i}$ ), and a residual error,  $e_{ti}$ .

As in the unconditional means model, the Level 2 model indicates that each student-specific intercept is equal to the mean across all students,  $\gamma_{00}$ , plus the student's random deviation from that mean. Because time is now included in the model, this fixed intercept indicates the average teacher-rated aggression in the *fall of 5th grade*, when (grade-5), (grade-5)<sup>2</sup> and semester all equal 0. The Level 2 model also indicates that the student-specific linear slope for grade ( $\beta_{1i}$ ) is equal to the average linear slope for grade ( $\gamma_{10}$ ) and the student's random deviation from this average ( $\zeta_{1i}$ ). Similarly, the student-specific effect of grade<sup>2</sup> is equal to the average quadratic effect of grade ( $\gamma_{20}$ ) and the student's random deviation from this average ( $\zeta_{2i}$ ). Because there are no random effects included for semester, the student-specific effect of semester,  $\beta_{3i}$ , is only a function of the overall fixed effect of semester; in other words, the association between semester and aggression is assumed to be identical across all students.

Essentially, the above equations indicate that the unconditional growth model estimates a population-level regression line and student-specific regression lines. The fixed effects describe the *average* initial value (i.e., intercept) and *average* patterns of change (e.g., linear slope, quadratic slope) within a population. The random effects describe how students' regression lines deviate from the population-level regression line.

Figure 14 illustrates the interpretation of the parameters in an unconditional growth model with a fixed and random linear effect of grade. The light gray dashed line is the average regression line for teacher-rated aggression. Because the variables are centered, the fixed intercept (i.e.,  $\gamma_{00}$ ) for this trajectory corresponds to the average teacher-rated aggression in the *fall of 5th grade*. There is a small positive linear effect of



grade (i.e.,  $\gamma_{10}$ ), which suggests that average teacher-rated aggression increases slightly across grades.

The regression line for student  $i$  is given by the dark, dotted line (intercept = 2.0, slope = 0.20). The difference between this student's estimated aggression and the population-level mean in the *fall of 5<sup>th</sup> grade* is given by  $\zeta_{0i}$ ; here,  $\zeta_{0i}$  is positive because in 5<sup>th</sup> grade this student was more aggressive than the average 5<sup>th</sup> grader. The difference in student  $i$ 's slope for grade and the overall population-level slope is given by  $\zeta_{1i}$ ; here,  $\zeta_{1i}$  is positive because student  $i$  increases in aggression faster than the average student. Once again, the difference between student  $i$ 's estimated aggression and his/her observed aggression at any assessment (black squares) is given by  $e_{it}$ ; Here, the student has a positive value of  $e_{it}$  in the spring of 7<sup>th</sup> grade (grade = 7.5).

When more than one Level 2 residual is in the model, the covariance between the residuals is estimated in addition to the variance of each residual. For example, in an unconditional growth model with a fixed and random linear effect of grade, *four* random parameters are estimated: (1) residual variance, (2) intercept variance, (3) linear slope for grade variance, and (4) the covariance between the intercept and slope. Theoretically, it is possible to interpret the covariance parameter (i.e., is there a significant relationship between students' initial level of aggression and their change in aggression over time?) However, because the magnitude (and even direction!) of parameter can vary depending on where the intercept is placed, any interpretation must note the conditional dependence of the effect (e.g., there is a small positive association between a student's aggression *in 5<sup>th</sup> grade* and their change in aggression over time).

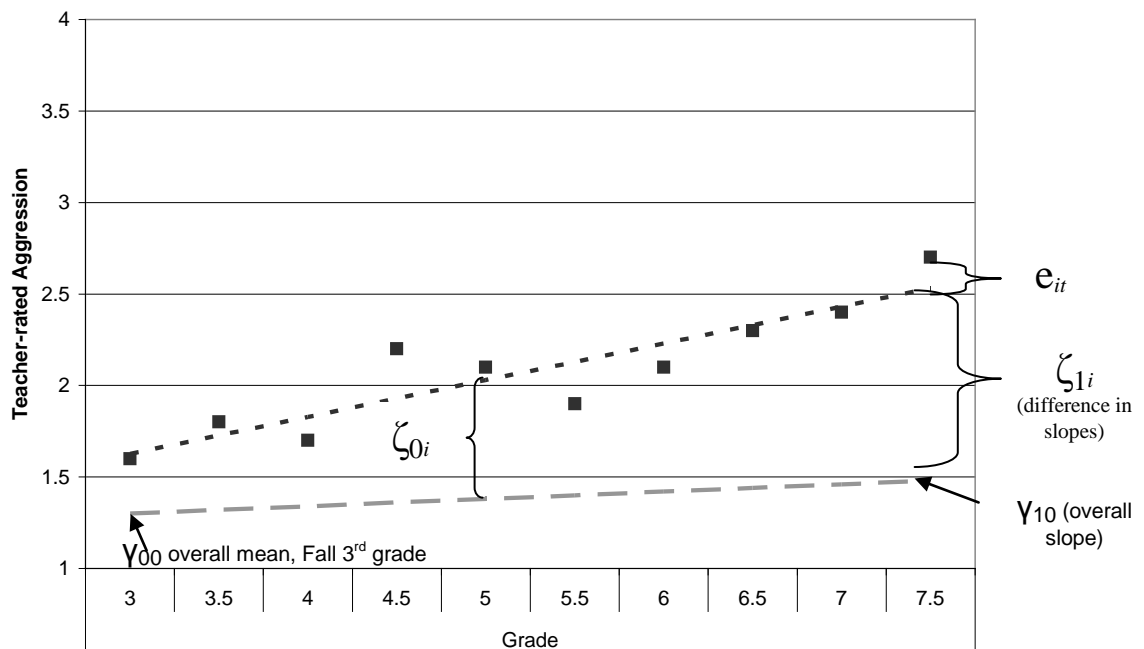


Figure 14. Illustration of Several Parameters in the Unconditional Growth Model

To determine which unconditional growth model fit best, a series of models were sequentially estimated with restricted maximum likelihood (REML). The significance of

the fixed effects was evaluated with *t*-tests. The significance of random effects was evaluated by comparing the difference between the REML deviance (-2 Res Log Likelihood) for the full and restricted models to a Chi-square distribution, with *df* = number of random parameters added to the model.

The syntax to estimate the unconditional growth model with a fixed linear effect of grade (Model 2) is shown in Figure 15 and the output is shown in Figure 16.

```
TITLE "Unconditional Growth Model: Fixed Linear Effect for Grade";
PROC MIXED DATA = &udatafile. NOCLPRINT NOITPRINT COVTEST METHOD=REML;
WHERE NMIS(core)=0;
CLASS ID;
MODEL lntag = grade5 / SOLUTION DDFM=Satterthwaite;
  * A fixed linear effect of grade is added in this model;
RANDOM INTERCEPT / TYPE = UN SUBJECT=ID; RUN;
```

Figure 15. SAS Syntax for Estimating a Model with a Fixed Linear Effect for Grade

Covariance Parameter Estimates					
Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr >  Z
UN(1,1)	id	0.08464	0.006668	12.69	<.0001
Residual		0.07157	0.001997	35.84	<.0001

Fit Statistics	
-2 Res Log Likelihood	1553.0
AIC (smaller is better)	1557.0
AICC (smaller is better)	1557.0
BIC (smaller is better)	1565.1

Solution for Fixed Effects					
Effect	Estimate	Standard Error	DF	t Value	Pr >  t
Intercept	0.3134	0.01520	436	20.62	<.0001
grade5	0.02470	0.004065	2682	6.08	<.0001

Figure 16. Output from a Model with a Fixed Linear Effect for Grade

Adding a fixed effect of grade reduced the intercept from 0.33 (overall mean across all assessments) to 0.31 (mean in the fall of 5<sup>th</sup> grade). There was also a small, positive effect of grade ( $\gamma_{10} = 0.02$ ,  $p < .001$ ). There was very little impact of the fixed effect of grade on the intercept variance or residual variance.

Next, a model that added a random linear effect of grade was estimated (Model 3). The syntax to estimate this model is shown in Figure 17. The random effect of grade is added to the model by including “grade5” as part of the random statement. The covariance between the intercept and the linear effect of grade is automatically included (i.e., it does not need to be specified in the model or random statements).

```

TITLE "Unconditional Growth Model: Fixed & Random Linear Effect for Grade";
PROC MIXED DATA = &udatafile. NOCLPRINT NOITPRINT COVTEST METHOD=REML;
WHERE NMISST(core)=0;
CLASS ID;
MODEL lntag = grade5 / SOLUTION DDFM=Satterthwaite;
  * A fixed linear effect of grade is added in this model;
RANDOM INTERCEPT grade5 / TYPE = UN SUBJECT=ID; RUN;
  * A random linear effect of grade is added in this model;

```

Figure 17. SAS Syntax for Estimating a Fixed and Random Linear Effect for Grade

The output for Model 3 is shown in Figure 18. Although the Z-value for the random linear variance, UN(2,2), is significant ( $p < .001$ ), it is better to use deviance tests to determine whether the addition of a random linear effect of grade improved model fit. The REML deviance difference (2) = 1553.0 – 1446.9 = 106.1,  $p < .001$ : in other words, adding a random linear effect of grade significantly improved model fit. How did the model fit improve? Note that the residual (i.e., Level 1 error) variance for Model 3 ( $\sigma_{e_{it}}^2 = 0.06$ ) was less than the residual variance for Model 2 ( $\sigma_{e_{it}}^2 = 0.07$ ). By allowing students to have their own slopes, it was possible to explain some of the variance that had been attributed to error in Model 2.

Covariance Parameter Estimates					
Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr >  Z
UN(1,1)	id	0.08762	0.007117	12.31	< .0001
UN(2,1)	id	-0.00354	0.001926	-1.84	0.0659
UN(2,2)	id	0.006027	0.000942	6.40	< .0001
Residual		0.06195	0.001883	32.91	< .0001
Fit Statistics					
-2 Res Log Likelihood			1446.9		
AIC (smaller is better)			1454.9		
AICC (smaller is better)			1454.9		
BIC (smaller is better)			1471.2		
Solution for Fixed Effects					
Effect	Estimate	Standard Error	DF	t Value	Pr >  t
Intercept	0.3129	0.01546	406	20.24	< .0001
grade5	0.02580	0.005663	329	4.56	< .0001

Figure 18. Output from a Model with a Fixed & Random Linear Effect for Grade

The covariance between students' intercepts and slopes, UN(2,1) was not significant ( $\sigma_{\zeta_{0i}\zeta_{1i}}^2 = -0.004$ ,  $p = .07$ ), indicating that students who were more aggressive in the fall of 5<sup>th</sup> grade experienced similar levels of growth in aggression over time (i.e., similar slopes) as students who were less aggressive. This covariance is not of interest here, however, as 5<sup>th</sup> grade does not represent any meaningful starting point (other than the earliest point at which all cohorts were assessed); therefore, this value will not be interpreted in other models.

Combining the fixed and random effects of grade also indicated that on average there is a small increase in teacher-rated aggression across grade ( $\gamma_{10} = 0.03$ ,  $p < .001$ ), with significant between-person variation in this trajectory, 95% C.I. = [-0.13, 0.18]. From the confidence interval it is clear that some students became more aggressive across grades, whereas other students became less aggressive across grades.<sup>4</sup>

Next, a model adding a fixed effect of grade<sup>2</sup> (Model 4) was estimated by adding grade5\*grade5 to the model statement in the SAS syntax shown in Figure 17. The fixed effect of grade<sup>2</sup> was not significant ( $\gamma_{20} = -0.0001$ ,  $p = .97$ ). However, before this effect was removed from the model, a random effect of grade<sup>2</sup> (Model 5) was tested. To do so, grade5\*grade5 was added to the random statement in the SAS syntax shown in Figure 17. Adding a random quadratic effect of grade significantly improved model fit, REML deviance difference (3) = 52.4,  $p < .001$ . Therefore, even though on average there was no significant quadratic effect of grade, there was significant variation in this effect across people, 95% C.I. = [-0.04, 0.04]. Therefore, the non-significant fixed effect of grade<sup>2</sup> was retained in the model.

Finally, a fixed effect of semester was added to the model (Model 6). There was a small positive effect of semester ( $\gamma_{30} = 0.03$ ,  $p < .01$ ), indicating that teacher-rated aggression was slightly higher in the spring than in the fall.

Table 7 summarizes the parameter estimates for the six models that were estimated. REML deviance, AIC, and BIC are also provided for each model.

*Table 7. Parameter Estimates & Fit Statistics for Unconditional Means & Growth Models*

Fixed Effect	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
	<i>RC</i>	<i>SE</i>	<i>RC</i>	<i>SE</i>	<i>RC</i>	<i>SE</i>	<i>RC</i>	<i>SE</i>	<i>RC</i>	<i>SE</i>	<i>RC</i>	<i>SE</i>
<i>Intercept</i>	0.33***	0.01	0.31***	0.02	0.31***	0.02	0.31***	0.02	0.31***	0.02	0.30***	0.02
<i>Level 1</i>												
Grade - 5			0.02***	0.004	0.03***	0.006	0.03***	0.007	0.03***	0.006	0.02***	0.007
(Grade-5) <sup>2</sup>							-0.0001	0.003	-0.0003	0.004	-0.0006	0.004
Semester (Fall=0)											0.03**	0.009
Variance Components												
Random Effect	<i>RC</i>	<i>SE</i>	<i>RC</i>	<i>SE</i>	<i>RC</i>	<i>SE</i>	<i>RC</i>	<i>SE</i>	<i>RC</i>	<i>SE</i>	<i>RC</i>	<i>SE</i>
<i>Level 1</i>												
	0.07***	0.002	0.07***	0.002	0.06***	0.002	0.06***	0.002	0.06***	0.002	0.06***	0.002
<i>Level 2</i>												
Intercept	0.08***	0.007	0.08***	0.007	0.09***	0.007	0.09***	0.007	0.10***	0.008	0.10***	0.008
Grade - 5					0.006***	0.001	0.006***	0.001	0.006***	0.001	0.006***	0.001
(Grade-5) <sup>2</sup>									0.002***	0.0004	0.002***	0.0004
-2 Resid LL	1580.5		1553		1446.9		1456.7		1404.3		1401.9	
AIC	1584.5		1557		1454.9		1464.7		1418.3		1415.9	
BIC	1592.6		1565.1		1471.2		1480.9		1446.7		1444.3	

\*\*  $p < .01$ ; \*\*\*  $p < .001$

Note. Covariance parameters were estimated but are not displayed here

<sup>4</sup> As noted above, the confidence interval indicates the extent to which the *slope* varies across people, rather than the precision of the slope parameter. Therefore, the result is still significant even though the confidence interval contains 0. Because the confidence interval has both positive and negative values, this suggests that some people increased in aggression over time (positive slope), some people decreased in aggression over time (negative slope) and some people did not have any consistent linear change in aggression over time (zero slope)

### 5.3 Adding Controls

Once an unconditional model of time was established, the next step was to add control variables (i.e., covariates) to the model. Three variables were added:

*Group size.* Because the number of peers that students hang around with may be related to their aggression, it was important to control for the number of group members that a student had. Group size varies from assessment to assessment (i.e., a within-person effect), so number of group members was added as a time-varying covariate to the Level 1 model.

*Gender.* Direct aggression is more common among boys than among girls, so gender (a between-person effect) was added to the Level 2 model.

*Cohort.* As described above, the cohort-sequential study design provides a way to test the generalizability of results within a single study. In order to do so, the results must be consistent across cohorts. In the current model, the cohort indicators were added to the Level 2 model to test whether the intercept (centered at the fall of 5<sup>th</sup> grade) was different across cohorts.<sup>5</sup>

#### Level 1 Model:

$$\text{Teacher-rated aggression}_{ti} = \beta_{0i} + \beta_{1i}(\text{grade}_{ti} - 5) + \beta_{2i}(\text{grade}_{ti} - 5)^2 + \beta_{3i}(\text{semester}_{ti}) + \beta_{4i}(\text{\# of group members}_{ti} - 7) + e_{ti}$$

#### Level 2 Model:

$$\beta_{0i} = \gamma_{00} + \gamma_{01}(\text{gender}_i) + \gamma_{02}(\text{middle cohort}_i) + \gamma_{03}(\text{oldest cohort}_i) + \zeta_{0i}$$

$$\beta_{1i} = \gamma_{10} + \zeta_{1i}$$

$$\beta_{2i} = \gamma_{20} + \zeta_{2i}$$

$$\beta_{3i} = \gamma_{30}$$

$$\beta_{4i} = \gamma_{40}$$

The output for this model is shown in Figure 19 on the next page. The intercept of 0.27 in this model indicates that on average, girls had a very low teacher-rated aggression in the fall of 5<sup>th</sup> grade (1.31 on the 1 to 5 untransformed scale)<sup>6</sup>. There was a significant effect of gender on the intercept: in the fall of 5<sup>th</sup> grade, all else being equal (i.e., cohort and group size), boys had a higher natural log teacher-rated aggression score than girls ( $\gamma_{01} = 0.11$ ,  $p < .001$ ); in other words, on average, boys had a teacher-rated aggression

<sup>5</sup> When all three control variables are added in a single step, then we are actually testing whether the intercept is different for girls (gender = 0) with an average sized group (# of group members = 5) in the fall of 5<sup>th</sup> grade. Because of this, an intermediate model testing adding just the cohort indicators, before adding gender and group size, was estimated. No differences in the intercept were found for either cohort indicator, both  $p > .1$ .

<sup>6</sup> Note that at this point, the model assumptions have not been tested, therefore any interpretations of parameter estimates are preliminary. Interpretation is discussed here solely to facilitate an understanding of the information that the model provides.

of 0.38 (or 1.46 on the untransformed scale). There was no difference in aggression for girls in the oldest (cohort 1) and youngest (cohort 3) cohorts ( $\gamma_{03} = -0.02, p = .66$ ); however there was a trend for girls in middle cohort (cohort 2) to have a lower teacher-rated aggression than girls in the youngest cohort ( $\gamma_{02} = -0.06, p < .10$ ). There was no effect of group size at time  $t$  on a student's aggression at time  $t$  ( $\gamma_{40} = -0.0002, p = .66$ ).

Solution for Fixed Effects						
Effect	cohort, year 4	Estimate	Standard Error	DF	t Value	Pr >  t
Intercept		0.2695	0.03051	462	8.83	<.0001
grade5		0.02382	0.006880	269	3.46	0.0006
grade5*grade5		-0.00090	0.003786	327	-0.24	0.8120
semesterF		0.02817	0.009054	2033	3.11	0.0019
sz_GM5		-0.00019	0.001345	2650	-0.14	0.8858
sexF		0.1097	0.02953	406	3.72	0.0002
cohort	1	-0.01642	0.03740	424	-0.44	0.6608
cohort	2	-0.06332	0.03525	395	-1.80	0.0732
cohort	3	0	.	.	.	.

Figure 19. Fixed Effect from a Model Adding Controls for Group Size, Gender, & Cohort

#### 5.4 Full Model

Once the effect of control variables was established, a full model could be estimated. There are 10 predictor variables that can be included in the model. Testing all possible interactions among the variables would not be practical: not only would higher order interactions be difficult to interpret, but testing a large number of interaction terms would likely inflate the Type I error rate (just by chance, some higher order interaction effects are likely to be significant). Therefore, the statistical questions are used to guide how the statistical models should be set up.

The primary statistical questions focus on whether generally affiliating with aggressive group members (i.e., average group), affiliating with more aggressive group members than usual at the current assessment (i.e., current group), or affiliating with more aggressive group members than usual at the *past* assessment are associated with student's aggression. In addition, the fourth statistical question is whether or not these associations are moderated by gender. Finally, given the cohort-sequential design of the study, it is important to test whether the effects are consistent across cohorts.

Average group is a between-person effect: students who generally affiliate with aggressive group members may have different regression lines than students who generally affiliate with non-aggressive group members. This question does not focus on whether the association between average group and a student's own aggression varies across time; therefore, average group is added as a predictor of the student-specific (Level 2) intercepts, but not as a predictor of the linear or quadratic effects of grade.

To test whether the association between average group and average teacher-rated aggression are moderated by gender, an interaction between gender and average group is added as a predictor of the student-specific intercepts. In addition, to test whether the association between average group and average aggression is consistent across cohorts, two interaction terms (middle cohort by average group and oldest cohort by average group) are added as predictors of the student-specific intercepts.

In contrast, current and past group are within-person, or time-varying effects: students may be more or less aggressive than predicted by their regression line because of the aggressiveness of their group members at the current assessment or the aggressiveness of their group members at the past assessment. Therefore, current and past group are added as Level 1 predictors of teacher-rated aggression and each term is given its own regression coefficient. In turn, these regression coefficients may be predicted by gender, cohort, or average group. Essentially, adding these variables to the Level 2 prediction of the coefficients for current group and past group creates interaction terms between each variable and the within-person effect. For example, adding gender as a predictor of the regression coefficient for current group is the same as testing a gender by current group interaction. It is also possible that the coefficients may vary across students, so a random error term for current group and a random error term for past group should also be tested.

Translating these statistical questions into coefficients for the multilevel model leads to the complete Level 1 and Level 2 models specified below.

#### Level 1 Model:

$$\text{Teacher-rated aggression}_{ti} = \beta_{0i} + \beta_{1i}(\text{grade}_{ti} - 5) + \beta_{2i}(\text{grade}_{ti} - 5)^2 + \beta_{3i}(\text{semester}_{ti}) + \beta_{4i}(\# \text{ of group members}_{ti} - 7) + \beta_{5i}(\text{current group}_{ti}) + \beta_{6i}(\text{past group}_{ti}) + e_{ti}$$

#### Level 2 Model:

$$\beta_{0i} = \gamma_{00} + \gamma_{01}(\text{gender}_i) + \gamma_{02}(\text{middle cohort}_i) + \gamma_{03}(\text{oldest cohort}_i) + \gamma_{04}(\text{average group}_i) + \gamma_{05}(\text{gender}_i) * (\text{average group}_i) + \gamma_{06}(\text{middle cohort}_i) * (\text{average group}_i) + \gamma_{07}(\text{oldest cohort}_i) * (\text{average group}_i) + \zeta_{0i}$$

$$\beta_{1i} = \gamma_{10} + \zeta_{1i}$$

$$\beta_{2i} = \gamma_{20} + \zeta_{2i}$$

$$\beta_{3i} = \gamma_{30}$$

$$\beta_{4i} = \gamma_{40}$$

$$\beta_{5i} = \gamma_{50} + \gamma_{51}(\text{gender}_i) + \gamma_{52}(\text{middle cohort}_i) + \gamma_{53}(\text{oldest cohort}_i) + \gamma_{54}(\text{average group}_i) + \zeta_{3i}$$

$$\beta_{6i} = \gamma_{60} + \gamma_{61}(\text{gender}_i) + \gamma_{62}(\text{middle cohort}_i) + \gamma_{63}(\text{oldest cohort}_i) + \gamma_{64}(\text{average group}_i) + \zeta_{4i}$$

The output for the full model is shown in Figure 20. The number of observations for the full model indicates fewer observations were used here (N = 2451) than in the unconditional means model (N = 3007). This reduction occurred because in the full model, only observations in which a student had a within-person group members' aggression score at *both* the current and past assessments were used. All observations from the initial assessment (i.e., fall of 3<sup>rd</sup>, 4<sup>th</sup>, or 5<sup>th</sup> grade, depending on the cohort) were thus excluded from the model. In addition, seven students were dropped from the analysis because they did not have any consecutive observations. The remaining

observations that were dropped were from students who were included in the analysis but who had some non-consecutive observations.<sup>7</sup>

Number of Observations						
Number of Observations Read		3440				
Number of Observations Used		2451				
Number of Observations Not Used		989				
Covariance Parameter Estimates						
Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr >  Z	
UN(1,1)	id	0.07684	0.007248	10.60	<.0001	
UN(2,1)	id	0.003345	0.002801	1.19	0.2324	
UN(2,2)	id	0.005784	0.001740	3.32	0.0004	
UN(3,1)	id	-0.00507	0.001670	-3.03	0.0024	
UN(3,2)	id	-0.00181	0.000887	-2.04	0.0418	
UN(3,3)	id	0.002483	0.000613	4.05	<.0001	
Residual		0.05730	0.002094	27.37	<.0001	
Fit Statistics						
-2 Res Log Likelihood		1224.5				
AIC (smaller is better)		1238.5				
AICC (smaller is better)		1238.6				
BIC (smaller is better)		1266.8				
Solution for Fixed Effects						
Effect	cohort, year 4	Estimate	Standard Error	DF	t Value	Pr >  t
Intercept		0.3242	0.03249	413	9.98	<.0001
grade5		0.02029	0.009485	200	2.14	0.0336
grade5*grade5		0.003029	0.004977	326	0.61	0.5432
semesterF		0.02477	0.01032	1574	2.40	0.0165
sz_GM5		-0.00108	0.001483	2122	-0.73	0.4652
sexF		0.01086	0.03236	385	0.34	0.7373
cohort	1	-0.02719	0.03658	404	-0.74	0.4576
cohort	2	-0.06549	0.03436	377	-1.91	0.0574
cohort	3	0	0	0	0	0
BPGMavzppagg		0.4147	0.1162	386	3.57	0.0004
WPGMavzppaggLag0		0.1029	0.02826	1863	3.64	0.0003
WPGMavzppaggLag1		0.06694	0.02677	1811	2.50	0.0125
sexF*BPGMavzppagg		-0.1529	0.1132	397	-1.35	0.1774
sexF*WPGMavzppaggLag		-0.07256	0.02903	1958	-2.50	0.0125
sexF*WPGMavzppaggLag		-0.02786	0.02788	1953	-1.00	0.3177
BPGMavzpp*WPGMavzppa		-0.00110	0.02983	1941	-0.04	0.9705
BPGMavzpp*WPGMavzppa		-0.05021	0.02950	1955	-1.70	0.0889
BPGMavzppagg*cohort	1	0.07595	0.1027	393	0.74	0.4598
BPGMavzppagg*cohort	2	0.000111	0.1118	369	0.00	0.9992
BPGMavzppagg*cohort	3	0	0	0	0	0
WPGMavzppaggL*cohort	1	0.08994	0.03525	2009	2.55	0.0108
WPGMavzppaggL*cohort	2	-0.03066	0.02734	1953	-1.12	0.2622
WPGMavzppaggL*cohort	3	0	0	0	0	0
WPGMavzppaggL*cohort	1	-0.00170	0.03558	2005	-0.05	0.9619
WPGMavzppaggL*cohort	2	-0.01882	0.02707	1967	-0.70	0.4869
WPGMavzppaggL*cohort	3	0	0	0	0	0

Figure 20. SAS Output for Full Model

Non-significant effects interactions were removed one at a time, and the model was re-estimated after each interaction was removed. Because all of the main effects were

<sup>7</sup> Note that models which include different subsets of students should not be directly compared. To make any conclusions that the full model is better than the unconditional growth model, the unconditional growth model needs to be re-estimated using only those students who are included in the full model. When the unconditional growth model was re-estimated, the results were essentially the same as those presented for the unconditional growth model in this report. When the final model, described later, was compared to the unconditional growth model with the same number of cases, the final model was found to significantly improve model fit.



either control variables (grade, grade<sup>2</sup>, semester, number of group members, gender, and cohort) or substantively interesting even if they were non-significant (average, current, and past group), all main effects were left in the model. Random effects of current group and past group were tested, but they did not significantly improve model fit, both REML deviance difference (4) < 9.5. A significant gender by current group interaction and a significant average by current group interaction were found. To ensure that these effects did not differ across cohorts, three-way interactions (and all two-way interactions among the effects in the three-way interactions) were tested. Neither three-way interaction with cohort was significant,  $p > .05$ , so they were removed from the model. In addition to these two-way interactions, a significant current group by cohort interaction was found for the oldest cohort.

Before concluding that this was the final model, however, one additional model specification was considered. Ideally, in a multilevel model the addition of random effects of time accounts for the patterns of variances and covariances over time. However, it is also possible to add a covariance structure to the residuals errors to account for any remaining patterns (Hedeker & Gibbons, 2006). For example, a first-order autoregressive structure, AR(1), implies that the correlation between two assessments decreases exponentially as the time between the two assessments increases. A Toeplitz structure implies that the relationship between two assessments decreases as the time between the assessments increases, but a different correlation is allowed for each lag (i.e., consecutive assessments have a different correlation than assessments that are separated by additional assessments).

Thus, several alternative covariance structures (AR(1), TOEP, TOEPH) were tested and compared using REML deviance differences. The covariance structure is added to the multilevel model by adding a repeated statement to the SAS syntax. To treat time as both continuous (in the random statement) and categorical (in the repeated statement), the non-centered effect of grade was specified as a categorical factor. The SAS syntax to estimate a multilevel model with autocorrelated errors is given in Figure 21.

```
TITLE "Adding a Covariance Structure for the Level 1 Residuals";
PROC MIXED DATA = &udatafile. NOCLPRINT NOITPRINT COVTEST METHOD=REML;
WHERE NMISSED(core)=0;
CLASS ID cohort grade; * Grade is added treated as a categorical factor;
MODEL lntagg = grade5 grade5*grade5 semesterF sz_gm5 sexF cohort
      BPgmavzppagg WPGmavzppaggLag0 WPGmavzppaggLag1 WPGmavzppaggLag0*sexF
      BPgmavzppagg*WPGmavzppaggLag1 WPGmavzppaggLag0*cohort
/ SOLUTION DDFM=Satterthwaite;
RANDOM INTERCEPT grade5 grade5*grade5 / TYPE = UN SUBJECT=ID;
REPEATED grade/ TYPE=AR(1) SUBJECT = ID; RUN;
/* Adding a REPEATED statement applies a covariance structure to the Level 1
residuals
Grade, rather than grade5, is added to the repeated statement, as grade
is being treated as a categorical factor
The best-fitting structure was of TYPE = AR(1) */
```

Figure 21. SAS Syntax for Adding a Covariance Structure to the Level 1 Residuals

The best fitting model was the AR(1) model. The model with the AR(1) structure improved model fit compared to the model with no covariance structure for the Level 1 errors, REML deviance difference (1) = 25.5,  $p < .001$ . With the addition of the AR(1) error structure, the interaction between average group and past group became non-significant ( $p > .05$ ). Therefore, it was removed from the final model. The final model is shown in Figure 22.

Covariance Parameter Estimates					
Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr >  Z
UN(1,1)	id	0.06453	0.007698	8.38	<.0001
UN(2,1)	id	0.001002	0.002855	0.35	0.7257
UN(2,2)	id	0.003327	0.001915	1.74	0.0411
UN(3,1)	id	-0.00145	0.001810	-0.80	0.4233
UN(3,2)	id	-0.00089	0.000926	-0.96	0.3389
UN(3,3)	id	0.001222	0.000681	1.79	0.0364
AR(1)	id	0.2034	0.04245	4.79	<.0001
Residual		0.06793	0.004059	16.74	<.0001

Fit Statistics	
-2 Res Log Likelihood	1171.6
AIC (smaller is better)	1187.6
AICC (smaller is better)	1187.7
BIC (smaller is better)	1220.0

Solution for Fixed Effects						
Effect	cohort, year 4	Estimate	Standard Error	DF	t Value	Pr >  t
Intercept		0.3043	0.02943	411	10.34	<.0001
grade5		0.02026	0.009749	201	2.08	0.0389
grade5*grade5		0.003102	0.005000	324	0.62	0.5354
semesterF		0.02700	0.009410	926	2.87	0.0042
sz_GM5		-0.00087	0.001474	2121	-0.59	0.5535
sexF		0.02052	0.03095	387	0.66	0.5078
cohort	1	-0.02898	0.03648	408	-0.79	0.4274
cohort	2	-0.05723	0.03378	377	-1.69	0.0911
cohort	3	0	.	.	.	.
BPGMavzppagg		0.3370	0.04679	396	7.20	<.0001
WPGMavzppaggLag0		0.09995	0.02776	1910	3.60	0.0003
WPGMavzppaggLag1		0.02496	0.01227	1962	2.03	0.0421
sexF*WPGMavzppaggLag		-0.06602	0.02846	1976	-2.32	0.0205
WPGMavzppaggL*cohort	1	0.08695	0.03446	1982	2.52	0.0117
WPGMavzppaggL*cohort	2	-0.03140	0.02750	1965	-1.14	0.2537
WPGMavzppaggL*cohort	3	0	.	.	.	.

Figure 22. SAS Output for Final Model

Before interpreting the final model, however, it is important to check whether or not the model assumptions are met. This topic is examined in the next section.

## 5.5 Checking Assumptions

As with other regression models the multilevel model makes several assumptions (Singer & Willet, 2003). First, error is expected to be normally distributed. There are multiple error terms in the final model (Level 1 residuals, and Level 2 errors for intercept, linear effect of grade, and the quadratic effect of grade), and all of them should be approximately normally distributed. Second, the error terms should exhibit

constant variance across levels of the predictor variables. This can be examined by plotting the Level 1 variables against the Level 1 predictor variables and plotting the Level 2 variables against the Level 2 predictor variables.

The SAS syntax for checking model assumptions is provided in Appendix 1. The residual plots are provided in Appendices 3 (checking for normality of residuals), 4 (checking for constant variance of level 1 variables across levels of level 1 predictors), and 5 (checking for constant variance of level 2 variables across levels of level 2 predictors).

Examination of the residual distributions suggests that the normality assumption is probably adequately met. There is some small positive skew of the Level 1 residuals and the intercept error terms, which probably reflects that even after transforming teacher-rated aggression, it is still a little skewed. The skew in the residual distributions is not too extreme, however.

Examination of the Level 1 and Level 2 residuals against the predictor variables suggests that the assumption of homoscedasticity is met for the most part. Variance appears to decrease slightly as the number of group members increases (only a small number of students have the largest number of group members). However, the decrease was not severe enough to warrant testing additional models to try to account for the decrease in variance. It also appears that the youngest cohort exhibited more variance in the Level 2 residuals than the other two cohorts (most likely because they were in the study for the longest period of time). Estimating separate variance terms for each cohort (by adding cohort as a grouping variable on the random statement) did not significantly improve model fit, however. Therefore, it was concluded that although there appeared to be greater variance for the youngest cohort, this variance did not significantly impact model estimation.

## 5.6 Interpreting the Final Model

Because there are many effects in the final model, the intercept must be interpreted carefully. The intercept of 0.30 is the natural log of teacher-rated aggression of girls in the youngest cohort in the Fall of 5<sup>th</sup> grade, who have an average number of group members, who on average have slightly aggressive groups, and whose current and past groups are as aggressive as they usually are. In other words, the intercept is the expected natural log of teacher-rated aggression when all of the predictor variables are zero (i.e., grade-5 = 0, semesterF = 0, sexF = 0, cohort1 = 0, cohort2 = 0, group size – 5 = 0, Average Group = 0, Current Group = 0, Past Group = 0). These girls have an untransformed teacher-rated aggression score of 1.35, with a 95% C.I. = [1.00, 2.22].

Consistent with earlier models, there was a small, positive effect of the fixed effect of grade ( $\gamma_{10} = 0.02$ ,  $p < .05$ ), no effect of the fixed effect of grade<sup>2</sup> ( $\gamma_{20} = 0.003$ ,  $p = .53$ ), a small, positive effect of semester ( $\gamma_{30} = 0.03$ ,  $p < .01$ ) and significant random variation in both the linear and quadratic effects of grade. The number of group members was not associated with teacher-rated aggression.

There was also a positive effect of average group ( $\gamma_{04} = 0.34, p < .001$ ), such that students who typically affiliate with aggressive group members on average have higher teacher-rated aggression than youth who typically affiliate with less aggressive group members. The main effect of gender can no longer be interpreted directly, because there was a significant gender by current group interaction. There was a small, positive association between current group and teacher-rated aggression for girls, but this association was smaller for boys. For both boys and girls, there was a very small (but significant) effect of past group, such that students who had more aggressive group members at the previous assessment had higher aggression at the current assessment.

### 5.6.1 Interpreting Interaction Effects

To interpret the significant current group by cohort interaction, results were plotted for otherwise “typical” youth (i.e., Fall of 5th grade, average number of group members). Results were plotted separately for each cohort. Current group was on the X-axis and predicted values were plotted for youth with low (1 SD below the mean;  $\approx -0.3$ ) and high (1 SD above the mean;  $\approx 0.3$ ) average group members. Results were also plotted separately for boys and girls; however, only the results from girls will be shown here (as suggested by the significant negative gender by current group interaction, the slope for current group was slightly lower for boys than for girls).

To create the interaction plots, the values of each variable for the prototypical youth had to be assigned. A small data set with the values for two prototypical girls was created in Excel (see Figure 23). One of these prototypical girls had an average group aggression of -0.3 and one girl had an average group aggression of 0.3. Then, different values of current group were assigned (from -0.6 to 0.6). Once these prototypical girls were created, it was possible to plot their predicted teacher-rated aggression in SAS. The syntax to plot the interactions for these prototypical youth is given in Appendix 2. The process was repeated with prototypical girls in the middle and oldest cohorts and for boys, simply by altering the relevant variables in the Excel worksheet.

Order	ID	Cohort	Grade	Grade5	BPgmvazppa	WPgmvazppa	WPgmvazppa	core	sz_gm5	sexF	semesterF
1	-102	3	5	0	-0.3	-0.6	0	1	0	0	0
1	-102	3	5	0	-0.3	-0.4	0	1	0	0	0
1	-102	3	5	0	-0.3	-0.2	0	1	0	0	0
1	-102	3	5	0	-0.3	0	0	1	0	0	0
1	-102	3	5	0	-0.3	0.2	0	1	0	0	0
1	-102	3	5	0	-0.3	0.4	0	1	0	0	0
1	-102	3	5	0	-0.3	0.6	0	1	0	0	0
2	-104	3	5	0	0.3	-0.6	0	1	0	0	0
2	-104	3	5	0	0.3	-0.4	0	1	0	0	0
2	-104	3	5	0	0.3	-0.2	0	1	0	0	0
2	-104	3	5	0	0.3	0	0	1	0	0	0
2	-104	3	5	0	0.3	0.2	0	1	0	0	0
2	-104	3	5	0	0.3	0.4	0	1	0	0	0
2	-104	3	5	0	0.3	0.6	0	1	0	0	0

Figure 23. Excel Worksheet with Prototypical Girl in Cohort 3

The interaction plots obtained from SAS are shown in Appendix 6. Before examining the impact of cohort, it is clear that the strongest effect in all three cohorts is the difference between students who typically affiliate with aggressive groups (high average group)

and students who typically affiliate with non-aggressive groups (low average group). The difference in the light grey and black lines illustrates that students who typically have aggressive group members are on average more aggressive than students who typically have non-aggressive group members. These results are consistent with the significant positive effect of average group ( $\gamma_{04} = 0.34, p < .001$ ).

Affiliating with more aggressive group members than usual at a particular assessment is associated with slightly higher teacher-rated aggression, although this association is strongest in the oldest cohort. There is little difference between the middle and youngest cohort. These results are consistent with the significant current group by oldest cohort interaction ( $\gamma_{07} = 0.09, p < .05$ ) and the non-significant current group by middle cohort interaction ( $\gamma_{06} = -0.03, p = .25$ ).

## 5.7 Summary

In sum, the results highlight some of the potential benefits of using an MLM strategy to address questions about peer influence. Using MLM, rather than repeated measures ANOVA, allows all observations to be used, rather than requiring complete case analysis. This is particularly important when longitudinal cohort-sequential designs are used. In addition, MLM allows for variation across students in how they change over time (i.e., it allows for subject-specific slopes in addition to subject-specific intercepts).

MLM also improves upon the tendency in the field of peer relationships of inferring peer influence from ordinary least squares (OLS) regression models in which peer characteristics predict change over time in a child's outcome (i.e., Time1 [T1]-Peer predicts Time2-Student Aggression after controlling for T1-Student Aggression). This OLS approach assumes that the effect of the T1-Peer variable represents the unique effects of the T1 Peer context. Given that peer contexts have some continuity over time, however, the T1-Peer variable may be better conceptualized as reflecting two distinct components: a child's typical (or average) peer context and his or her deviation from that average at T1. OLS regression "bundles" these conceptually distinct components into a single parameter estimate, whereas MLM allows them to be separated into two parameters (Singer & Willett, 2003; Snijders & Bosker, 1999).

Separating within- and between-person effects in the analyses here allowed for the possibility that within-person effects would drive the association between affiliating with aggressive groups and student's aggression. Instead, the results indicated that past group members only had a very small influence on current aggression, after controlling for the effect of current group and average group. Instead, typically affiliating with an aggressive group (i.e., average group) had a much stronger effect. The interpretation of this between-person effect, however, is ambiguous: it could reflect both selection and on-going influence (e.g., maintenance or cumulative effects). The results suggest that studies relying on OLS regression may overestimate the strength of peer influence, by interpreting a parameter that combines between- and within-person effects as only reflecting socialization.

## 5.8 References

### 5.8.1 Background References

- Cairns, R. B., Cairns, B. D., Neckerman, H. J., Gest, S. D., & Garipey, J. L. (1988). Social networks and aggressive behavior: Peer support or peer rejection? *Developmental Psychology, 24*(6), 815-823.
- Coie, J. D., & Dodge, K. A. (1998). Aggression and antisocial behavior. In W. Damon & N. Eisenberg (Eds.), *Handbook of Child Psychology: Vol. 3 Social, Emotional, and Personality Development* (Vol. 3, pp. 55-99). New York: Wiley.
- Gest, S. D., Farmer, T. W., Cairns, B. D., & Xie, H. (2003). Identifying children's peer social networks in school classrooms: Links between peer reports and observed interactions. *Social Development, 12*, 513-529.
- Kandel, D. B. (1978). Homophily, selection, and socialization in adolescent friendships. *American Journal of Sociology, 84*(2), 427-436.
- Neckerman, H. J. (1996). The stability of social groups in childhood and adolescence: The role of the classroom social environment. *Social Development, 5*, 131-145.

### 5.8.2 Statistical References

- Hedeker, D., & Gibbons, R.D. (2006). *Longitudinal Data Analysis*. Hoboken, N.J.: John Wiley & Sons, Inc.
- Kim, J.Y., McHale, S. M., Osgood, D. W., & Crouter, A. C. (2006). Longitudinal course and family correlates of sibling relationships from childhood through adolescence. *Child Development, 77*(6), 1746-1761.
- Singer, J. D., & Willett, J. B. (2003). *Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence*. New York: Oxford University Press.
- Snijders, T.A.B., & Bosker, R. (1999). *Multilevel Analysis*. Thousand Oaks, CA: Sage.

In addition, much of my knowledge of multilevel modeling comes from my class notes for HDFS 597A: Applied Multilevel Analysis that I took several semesters ago from Lesa Hoffman. I continue to adapt the SAS syntax from this class for my research and some of that adapted syntax was included in this report.

## Appendix 1. SAS Syntax for Checking Model Assumptions

```

/**** Checking Model Assumptions ****/

DATA work.association;
    SET STATFP.StatApr09;

/*Define global variable for datafile name to be replaced in code below */

%GLOBAL udatafile varname varlabel filesave;
%LET udatafile=STATFP.lntagg;
%LET varname=lntagg;
%LET varlabel=LN Teacher-rated Aggression;
%LET filesave = F:\Courses\Stat 581\Final Project\Resid Plots;
RUN;

/* Rerun the model, this time saving predicted values and residuals */

TITLE1 "Model 4 - Final model";
PROC MIXED DATA = work.association NOCLPRINT NOITPRINT COVTEST METHOD=REML;
    WHERE NMISSED(core)=0;
    CLASS ID cohort grade sexF;
    MODEL lntagg = grade5 grade5*grade5 semesterF sz_gm5 sexF cohort
BPgmavzppagg WPGmavzppaggLag0 WPGmavzppaggLag1 WPGmavzppaggLag0*sexF
WPGmavzppaggLag0*cohort
    / SOLUTION DDFM=Satterthwaite OUTP=L1Resids;
    RANDOM INTERCEPT grade5 grade5*grade5 / SOLUTION TYPE = UN SUBJECT=ID;
    REPEATED grade/ TYPE=AR(1) SUBJECT = ID;
    ODS OUTPUT SolutionR=RandEffects; RUN;

/* OUTP creates a new data set with the predicted values and residuals from
the full model added. This new data set is a subset of the original,
with only students from the core sample in it
SOLUTION on the RANDOM statement creates the random effects
ODS OUTPUT SolutionR=RandEffects saves estimated random effects to a SAS
dataset called RandEffects */

/* Rename the residuals from the saved dataset
Overwrite the file, keeping student's id, grade, and Level 1 residuals */

DATA L1Resids; SET L1Resids;
    L1Resid=Resid;
    KEEP ID grade5 L1Resid; RUN;

/* The original RandEffects data lists the random effects on separate lines
for each student
First, remove any estimates that are not 0;
Then, use PROC SORT to sort the data by id. This gets the random effects
for each student in adjacent rows
Then, use PROC TRANSPOSE to create a new data set that treats random
effects as variables
Output a new data set, RandEffectsT, that has one line for each student */

DATA RandEffects; SET RandEffects;
    IF Estimate ~=0; RUN;
PROC SORT DATA=RandEffects; BY ID; RUN;

```

```

PROC TRANSPOSE DATA=RandEffects OUT=RandEffectsT;
  BY ID; * Specifies that transposing will occur within each ID;
  ID Effect; * Specifies which var in the input has the name for the new var;
  VAR Estimate; * Specifies which var has the value for the new var;
  RUN;

/* Rename the random effects to prevent them from overwriting the original
   variables by the same name */

DATA RandEffectsT; SET RandEffectsT;
  RandInt = Intercept;
  RandGrade = Grade5;
  RandGradeSq = grade5_grade5;
  KEEP ID RandInt RandGrade RandGradeSq; RUN;

/* Merge the Level 1 residuals with the original data by student and time */

PROC SORT DATA = L1Resids;
BY ID grade5; RUN; * Sort L1Resids by ID and grade;

PROC SORT DATA=work.association;
BY ID grade5; RUN; * Sort work.association by ID and grade;

DATA merged;
  MERGE work.association L1Resids; * Identify which two files to merge;
  BY ID grade5; RUN; * Indicate which variables to match the files on;

/* Merge random effects with data from above (original data + residuals) */
DATA Aggslope;
  MERGE merged RandEffectsT;
  BY ID; RUN; * Random effects are at Level 2 so match based on student;

/* Label the new variables */
DATA Aggslope; SET Aggslope;
LABEL L1Resid="Level 1 Residual"
      RandInt="Random Intercept"
      RandGrade="Random Linear Slope for Grade"
      RandGradeSq="Random Quadratic Slope for Grade-Squared"; RUN;

/* Save the new data set to the MSTP library
   To keep the file small, only save the variables that will be needed to
   check assumptions */

DATA &udatafile.; SET Work.aggslope;
KEEP ID Grade SemesterF WPGMavzppaggLag0 WPGMavzppaggLag1 L1Resid sz_GM
      SexF BPGMavzppagg RandInt RandGrade RandGradeSq Cohort; RUN;

/***** Make plots to check assumptions *****/

/* Set generic plotting options for text, size, and saving plots */
GOPTIONS RESET=ALL NOBORDER FTEXT=Triplex FTITLE=Triplex VSIZE=4in HSIZE=5in
GSFNAME=outgraph DEV=bmp GSFMODE=REPLACE; RUN;

```



```

/**** Plot a histogram of Level 1 residuals to check for normality *****/

FILENAME outgraph "&filesave.\&varname. Residual Distribution.bmp";
  TITLE1 "Residual Distribution:" ;
  TITLE2 "&varlabel.";
PROC UNIVARIATE DATA=&udatafile.;
  VAR L1Resid;
  HISTOGRAM L1Resid; RUN;

/**** Plot histogram of Level-2 residuals to check for normality *****/

/* Random Intercept Residual, Level 2 Distribution */
FILENAME outgraph "&filesave.\&varname. Random Intercept Distribution.bmp";
  TITLE1 "Random Intercept Distribution:";
  TITLE2 "&varlabel.";
PROC UNIVARIATE DATA=&udatafile.;
  VAR RandInt;
  HISTOGRAM RandInt; RUN;

/* Random Linear Effect of Grade Residual, Level 2 Distribution */
FILENAME outgraph "&filesave.\&varname. Random Grade Distribution.bmp";
  TITLE1 "Random Effect of Grade Distribution:";
  TITLE2 "&varlabel.";
PROC UNIVARIATE DATA=&udatafile.;
  VAR RandGrade;
  HISTOGRAM RandGrade; RUN;

/* Random Quadratic Effect of Grade, Level 2 Residual Distribution */
FILENAME outgraph "&filesave.\&varname. Random GradeSq Distribution.bmp";
  TITLE1 "Random Effect of Grade-Square Distribution:";
  TITLE2 "&varlabel.";
PROC UNIVARIATE DATA=&udatafile.;
  VAR RandGradeSq;
  HISTOGRAM RandGradeSq; RUN;

/**** Plot L1 residuals by L1 variables to check for constant variance ****/

FILENAME outgraph "&filesave.\&varname. Residuals by Grade.bmp";
  TITLE1 "Residuals as a Function of Grade";
  TITLE2 "&varlabel.";
PROC GPLOT DATA=&udatafile.;
  AXIS1 LABEL = (ANGLE=90 "Level 1 Residuals");
  AXIS2 LABEL = ("Grade");
  PLOT L1Resid*grade/ VAXIS=AXIS1 HAXIS=AXIS2; RUN; QUIT;

FILENAME outgraph "&filesave.\&varname. Residuals by WPGMavzppaggLag0.bmp";
  TITLE1 "Residuals as a Function of Within-Person Current Group Members";
  TITLE2 "&varlabel.";
PROC GPLOT DATA=&udatafile.;
  AXIS1 LABEL=(ANGLE=90 "Level 1 Residuals");
  AXIS2 LABEL=("Within-person Current Group Members") ORDER=(-3 TO 3 BY 1) ;
  PLOT L1Resid*WPGMavzppaggLag0/ VAXIS=AXIS1 HAXIS=AXIS2; RUN; QUIT;

FILENAME outgraph "&filesave.\&varname. Residuals by WPGMavzppaggLag1.bmp";
  TITLE1 "Residuals as a Function of Within-Person Past Group Members";
  TITLE2 "&varlabel.";

```

```

PROC GPLOT DATA=&udatafile.;
  AXIS1 LABEL=(ANGLE=90 "Level 1 Residuals");
  AXIS2 LABEL=("Within-person Past Group Members") ORDER=(-3 TO 3 BY 1) ;
  PLOT L1Resid*WPGMavzppaggLag1/ VAXIS=AXIS1 HAXIS=AXIS2; RUN; QUIT;

FILENAME outgraph "&filesave.\&varname. Residuals by Semester.bmp";
  TITLE1 "Residuals as a Function of Semester";
  TITLE2 "&varlabel.";
PROC GPLOT DATA=&udatafile.;
  AXIS1 LABEL=(ANGLE=90 "Level 1 Residuals");
  AXIS2 LABEL=("Semester") ORDER=(-1 TO 2 BY 1) VALUE=(" " "Fall" "Spring" "");
  PLOT L1Resid*SemesterF/ VAXIS=AXIS1 HAXIS=AXIS2; RUN; QUIT;

FILENAME outgraph "&filesave.\&varname. Residuals by Group Size.bmp";
  TITLE1 "Residuals as a Function of Number of Group Members";
  TITLE2 "&varlabel.";
PROC GPLOT DATA=&udatafile.;
  AXIS1 LABEL=(ANGLE=90 "Level 1 Residuals");
  AXIS2 LABEL=("Number of Group Members") ORDER=(0 TO 20 BY 2);
  PLOT L1Resid*sz_GM/ VAXIS=AXIS1 HAXIS=AXIS2; RUN; QUIT;

/* Plot L2 residuals by L2 variables to check for constant variance */

/* Average Group Members' Aggression */
FILENAME outgraph "&filesave.\&varname. Random Intercept by BPgrp.bmp";
  TITLE1 "Random Intercept Variance by Average Group Aggression";
  TITLE2 "&varlabel.";
PROC GPLOT DATA=&udatafile.;
  AXIS1 LABEL=(ANGLE=90 "Random Intercept");
  AXIS2 LABEL=("Average Group Aggression") ORDER=(-1.5 TO 1.5 BY .5);
  PLOT RandInt*BPGMavzppagg/ VAXIS=AXIS1 HAXIS=AXIS2; RUN; QUIT;

FILENAME outgraph "&filesave.\&varname. Random Grade by BPgrp.bmp";
  TITLE1 "Random Grade Variance by Average Group Aggression";
  TITLE2 "&varlabel.";
PROC GPLOT DATA=&udatafile.;
  AXIS1 LABEL=(ANGLE=90 "Random Slope for Grade");
  AXIS2 LABEL=("Average Group Aggression") ORDER=(-1.5 TO 1.5 BY .5);
  PLOT RandGrade*BPGMavzppagg/ VAXIS=AXIS1 HAXIS=AXIS2; RUN; QUIT;

FILENAME outgraph "&filesave.\&varname. Random Grade Squared by BPgrp.bmp";
  TITLE1 "Random Grade Squared Variance by Average Group Aggression";
  TITLE2 "&varlabel.";
PROC GPLOT DATA=&udatafile.;
  AXIS1 LABEL=(ANGLE=90 "Random Slope for Grade Squared");
  AXIS2 LABEL=("Average Group Aggression") ORDER=(-1.5 TO 1.5 BY .5);
  PLOT RandGradeSq*BPGMavzppagg/ VAXIS=AXIS1 HAXIS=AXIS2; RUN; QUIT;

/* Gender */
FILENAME outgraph "&filesave.\&varname. Random Intercept by Gender.bmp";
  TITLE1 "Random Intercept Variance by Gender";
  TITLE2 "&varlabel.";

PROC GPLOT DATA=&udatafile.;
  AXIS1 LABEL=(ANGLE=90 "Random Intercept");
  AXIS2 LABEL=("Gender") ORDER=(-1 TO 2 BY 1) VALUE=(" " "Female" "Male" "");
  PLOT RandInt*SexF/ VAXIS=AXIS1 HAXIS=AXIS2; RUN; QUIT;

```

```

FILENAME outgraph "&filesave.\&varname. Random Grade by Gender.bmp";
  TITLE1 "Random Grade Variance by Gender";
  TITLE2 "&varlabel.";
PROC GPLOT DATA=&udatafile.;
  AXIS1 LABEL=(ANGLE=90 "Random Slope for Grade");
  AXIS2 LABEL=("Gender") ORDER=(-1 TO 2 BY 1) VALUE=(" " "Female" "Male" "");
  PLOT RandGrade*SexF/ VAXIS=AXIS1 HAXIS=AXIS2; RUN; QUIT;

FILENAME outgraph "&filesave.\&varname. Random Grade Squared by Gender.bmp";
  TITLE1 "Random Grade Squared Variance by of Gender";
  TITLE2 "&varlabel.";
PROC GPLOT DATA=&udatafile.;
  AXIS1 LABEL=(ANGLE=90 "Random Slope for Grade Squared");
  AXIS2 LABEL=("Gender") ORDER=(-1 TO 2 BY 1) VALUE=(" " "Female" "Male" "");
  PLOT RandGradeSq*SexF/ VAXIS=AXIS1 HAXIS=AXIS2; RUN; QUIT;

/* Cohort */

FILENAME outgraph "&filesave.\&varname. Random Intercept by Cohort.bmp";
  TITLE1 "Random Intercept Variance as a Function of Cohort";
  TITLE2 "&varlabel.";
PROC GPLOT DATA=&udatafile.;
  AXIS1 LABEL=(ANGLE=90 "Random Intercept");
  AXIS2 LABEL=("Cohort") ORDER=(0 TO 4 BY 1) VALUE=(" " "Oldest" "Middle"
  "Youngest" "");
  PLOT RandInt*Cohort/ VAXIS=AXIS1 HAXIS=AXIS2; RUN; QUIT;

FILENAME outgraph "&filesave.\&varname. Random Grade by Cohort.bmp";
  TITLE1 "Random Grade Variance as a Function of Cohort";
  TITLE2 "&varlabel.";
PROC GPLOT DATA=&udatafile.;
  AXIS1 LABEL=(ANGLE=90 "Random Slope for Grade");
  AXIS2 LABEL=("Cohort") ORDER=(0 TO 4 BY 1) VALUE=(" " "Oldest" "Middle"
  "Youngest" "");
  PLOT RandGrade*Cohort/ VAXIS=AXIS1 HAXIS=AXIS2; RUN; QUIT;

FILENAME outgraph "&filesave.\&varname. Random Grade Squared by Cohort.bmp";
  TITLE1 "Random Grade Squared Variance as a Function of Cohort";
  TITLE2 "&varlabel.";
PROC GPLOT DATA=&udatafile.;
  AXIS1 LABEL=(ANGLE=90 "Random Slope for Grade Squared");
  AXIS2 LABEL=("Cohort") ORDER=(0 TO 4 BY 1) VALUE=(" " "Oldest" "Middle"
  "Youngest" "");
  PLOT RandGradeSq*Cohort/ VAXIS=AXIS1 HAXIS=AXIS2; RUN; QUIT;

```

## Appendix 2. SAS Syntax for Plotting Interaction Terms

```

/**** Interpreting Interaction Effects ****/

/* Plot Final Model for Teacher-rated Aggression */

PROC SORT DATA=&udatafile.; BY ID; RUN;

/* Import prototypical people data from ResponsePlotData file in Excel */

PROC IMPORT DATAFILE = "&filesave.\ResponsePlotData.xls"
  OUT=tagPlot DBMS=EXCEL REPLACE; * Output this file as tagPlot;
  SHEET="girl3"; * Read data from the "Girl3" sheet (girls in cohort 3);
  GETNAMES=YES; * Read in the labels;
RUN;

PROC SORT DATA=tagPlot; BY ID; RUN;
* Sort prototypical people by ID variables;

DATA Merged; MERGE &udatafile. tagPlot; BY ID; RUN;
* Merge original data set and file with prototypical people;
* Match based on ID;

/* Run final model to generate predicted values for prototypical people */

TITLE1 "Final Model";
PROC MIXED DATA = Merged NOCLPRINT NOITPRINT COVTEST METHOD=REML;
  WHERE NMISSED(core)=0;
  CLASS ID cohort grade sexF;
  MODEL lntag = grade5 grade5*grade5 semesterF sz_gm5 sexF cohort
    BPgmavzppagg WPgmvavzppaggLag0 WPgmvavzppaggLag1 WPgmvavzppaggLag0*sexF
    WPgmvavzppaggLag0*cohort
  / SOLUTION DDFM=Satterthwaite OUTPM=tagPred;
  RANDOM INTERCEPT grade5 / TYPE = UN SUBJECT=ID;
  REPEATED grade/ TYPE=AR(1) SUBJECT = ID; RUN;

/* Keep only the prototypical people */
DATA tagPred; SET tagPred;
  WHERE Order IS NOT MISSING;
  PredR=exp(Pred); * Untransform the predicted values;
RUN;

PROC SORT DATA=tagPred; BY Order; RUN;
* Sorting by order will facilitate labeling in the legend below;

/* Plot Final Model for Teacher-rated Aggression */

/* Set generic plotting options for text, size, and joining lines*/
GOPTIONS NOBORDER FTEXT=Triplex FTITLE=Triplex VSIZE=7in HSIZE=7in
  INTERPOL=JOIN GSFNAME=outgraph DEV=BMP GSFMODE=replace; run;

FILENAME outgraph "&filesave.\Predicted LnTagg Youngest Cohort Girls.bmp";
  * Save the file to the location specified globally by filesave;
TITLE1 JUSTIFY=LEFT HEIGHT=1.5 "Predicted Teacher-rated Agg as a Function of
Current Group";
TITLE2 JUSTIFY=LEFT HEIGHT=1.3 "Youngest Cohort, Girls";

```

```

PROC Gplot DATA=tagPred;
AXIS1 LENGTH=4.5in LABEL=(HEIGHT = 1.4 ANGLE=90 "Teacher-rated Aggression")
ORDER=(1 TO 3 BY .5);
* AXIS1 = Y-axis;
* Y-axis has a LENGTH = 4.5 inches (out of 6 inches above);
* HEIGHT = 1.4 sets the size of the label;
* ANGLE = 90 places the axis label at a 90 degree angle;
* ORDER = indicates that the scale goes from 1 to 3 in increments of .5;

AXIS2 LENGTH=5.5in LABEL=(HEIGHT = 1.4 "Current Group Members")
ORDER=(-.8 TO .8 BY .2);
* AXIS2 = X-axis;
* X-axis has a LENGTH = 5.5 inches (out of 7 inches above);
* HEIGHT = 1.4 sets the size of the label;
* ORDER = indicates that the scale goes from 3 to 7 in increments of .5;

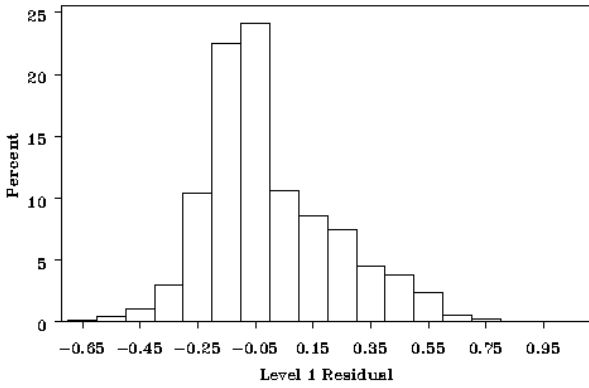
LEGEND1 FRAME POSITION = (CENTER TOP) ACROSS=1 LABEL=NONE VALUE=(HEIGHT=1.1
COLOR=Black JUSTIFY=Left "Low Average Group" "High Average Group");
SYMBOL1 COLOR=Black VALUE=Dot HEIGHT=1.3 LINE=1 WIDTH=2;
SYMBOL2 COLOR=Grey VALUE=Star HEIGHT=1.3 LINE=2 WIDTH=2;
* FRAME the legend
* CENTER the legend at the TOP;
* Only use one value across;
* The first VALUE is Low Average Group;
* The second VALUE is High Average Group;
* For the first values, use black dots with a solid line (LINE = 1);
* For the second values, use grey stars, with a dotted line (LINE = 2);
PLOT predr*WPgmavzppaggLag0=Order / VAXIS=AXIS1 HAXIS=AXIS2 LEGEND=LEGEND1;
* Obtain specifications for LEGEND from LEGEND 1;
* Obtain specifications for VAXIS (i.e., Y-axis) from AXIS1 from above;
* Obtain specifications for HAXIS (i.e., X-axis) from AXIS2 from above;
RUN; QUIT;

```

Appendix 3.  
Plots to Check Normality of Level 1 and Level 2 Residuals

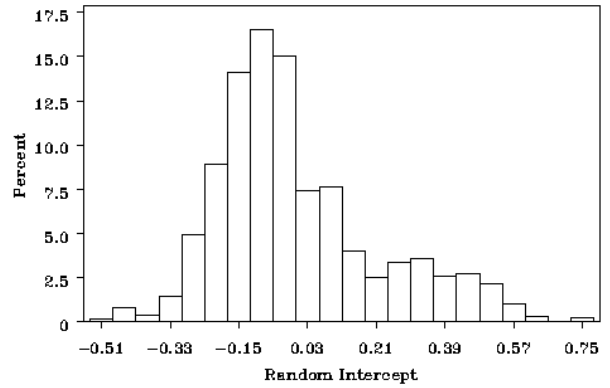
**Residual Distribution:**

LN Teacher-rated Aggression



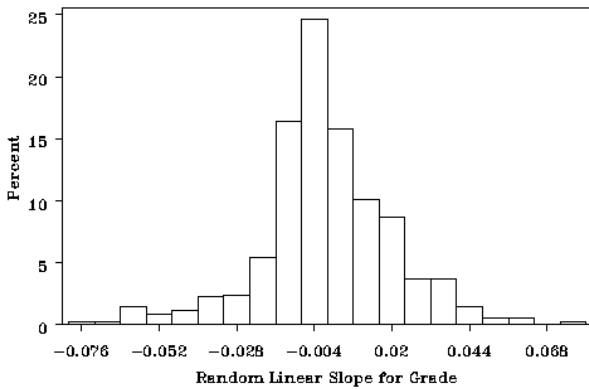
**Random Intercept Distribution:**

LN Teacher-rated Aggression



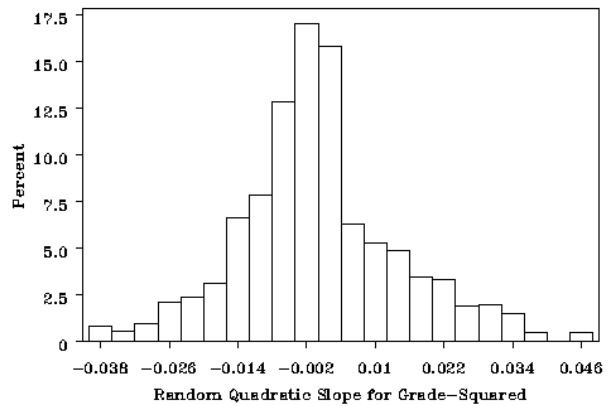
**Random Effect of Grade Distribution:**

LN Teacher-rated Aggression



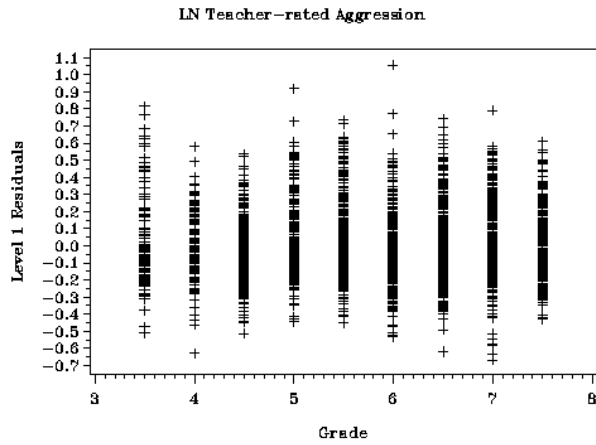
**Random Effect of Grade-Square Distribution:**

LN Teacher-rated Aggression

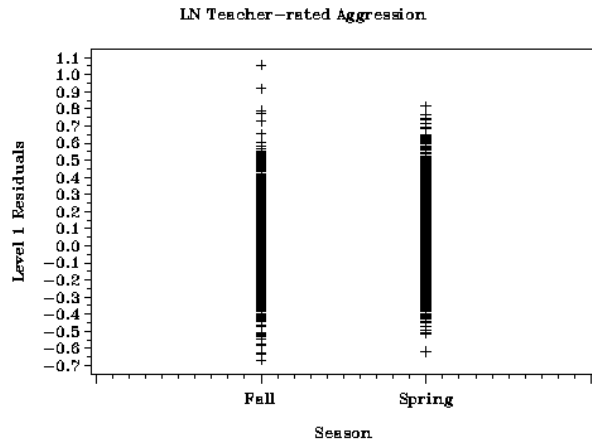


Appendix 4:  
Plots to Check Constant Variance of Level 1 across Level 1 Predictors

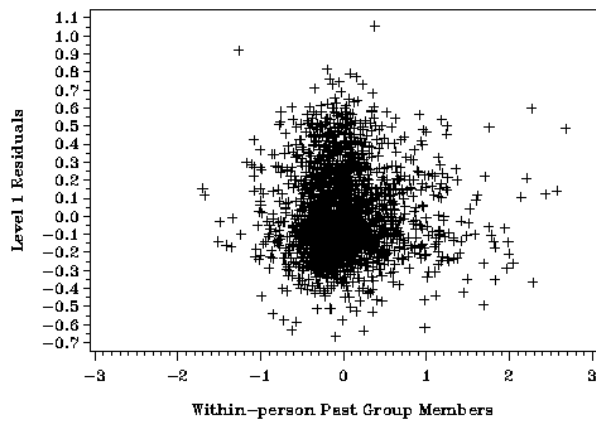
Residuals as a Function of Grade



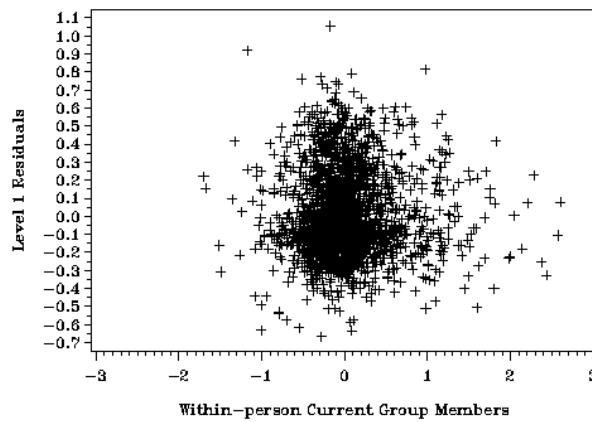
Residuals as a Function of Season



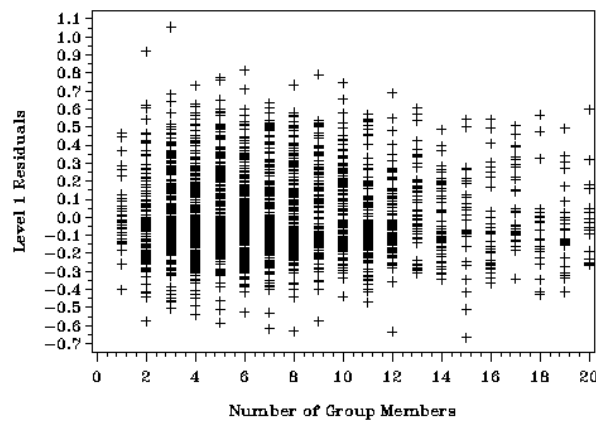
Residuals as a Function of Within-Person Past Group Members



Residuals as a Function of Within-Person Current Group Members



Residuals as a Function of Number of Group Members



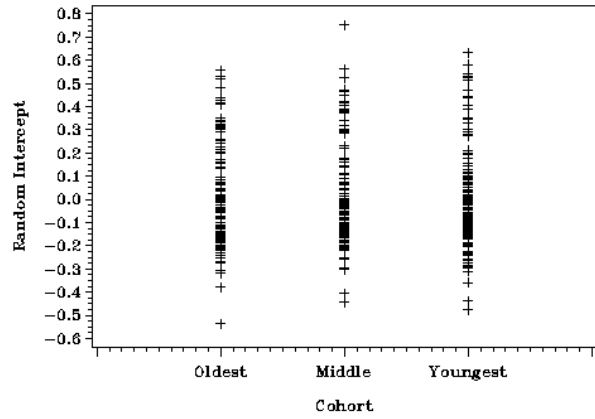
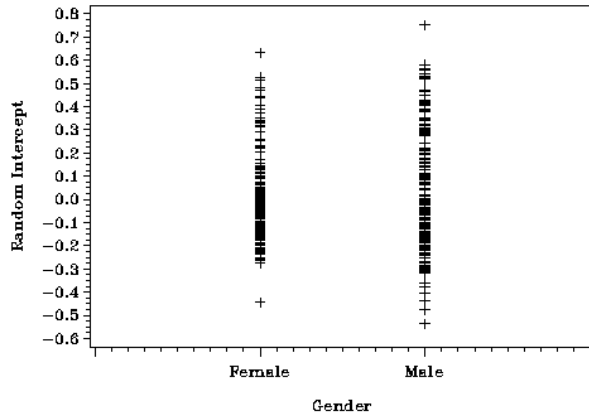
Appendix 5:  
Plots to Check Constant Variance of Level 2 Residuals across Level 2 Predictors

Gender:

Cohort:

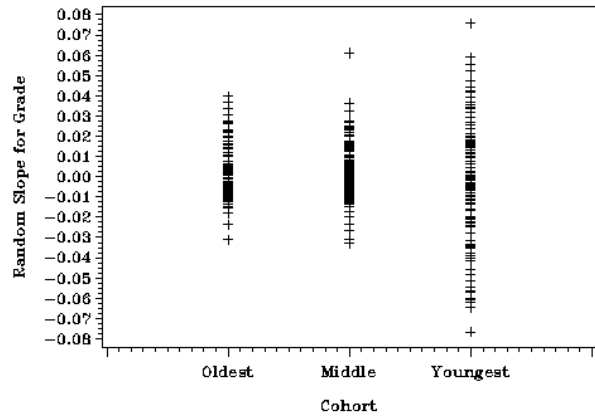
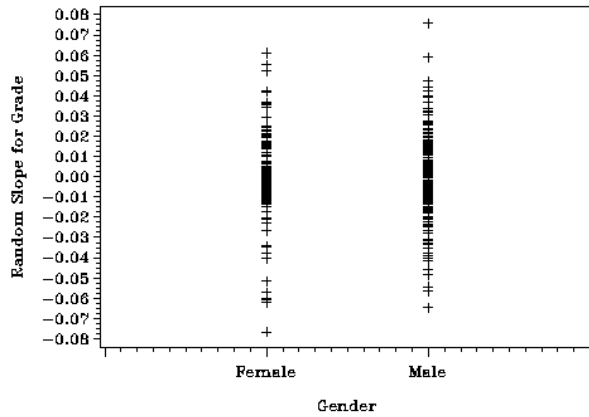
Random Intercept Variance as a Function of Gender  
LN Teacher-rated Aggression

Random Intercept Variance as a Function of Cohort  
LN Teacher-rated Aggression



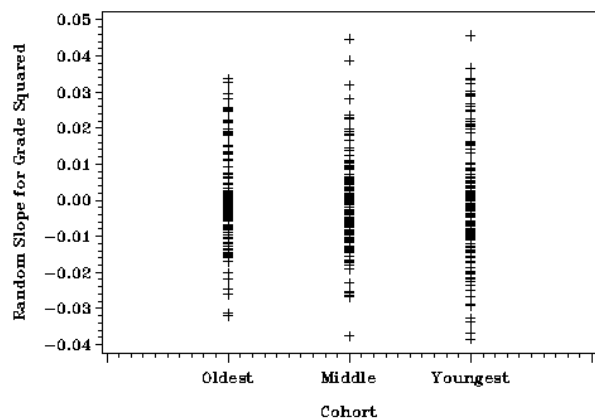
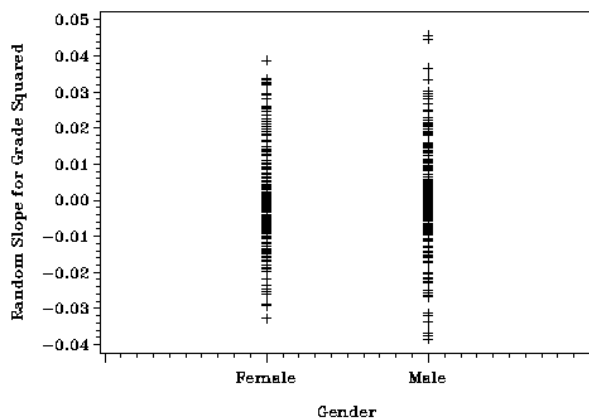
Random Grade Variance as a Function of Gender  
LN Teacher-rated Aggression

Random Grade Variance as a Function of Cohort  
LN Teacher-rated Aggression



Random Grade Squared Variance as a Function of Gender  
LN Teacher-rated Aggression

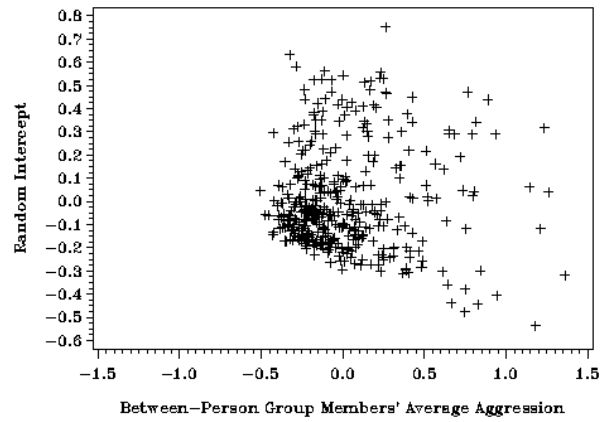
Random Grade Squared Variance as a Function of Cohort  
LN Teacher-rated Aggression



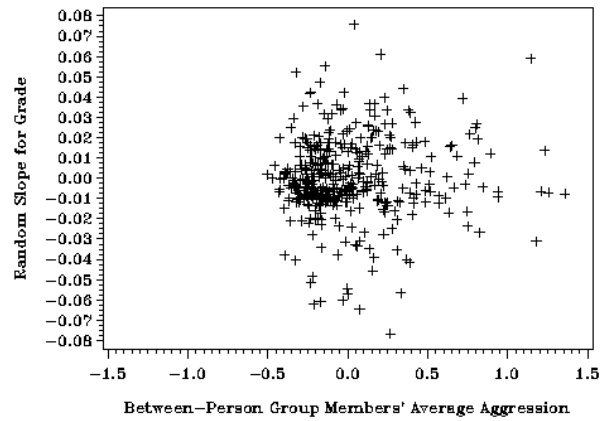


## Between-person (Average) Group Members' Aggression:

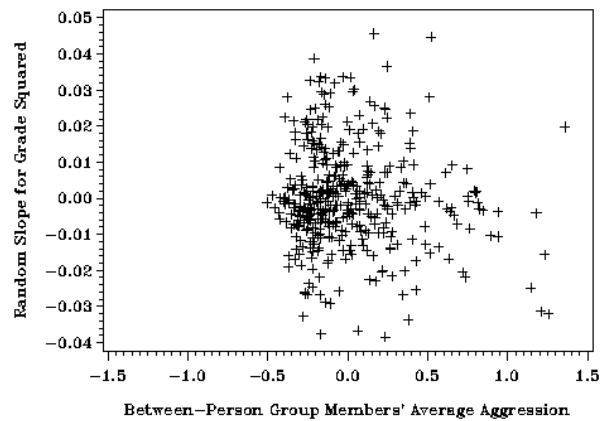
Random Intercept Variance as a Function of Group Members' Aggression  
LN Teacher-rated Aggression



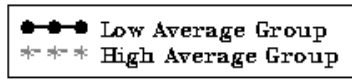
Random Grade Variance as a Function of Group Members' Aggression  
LN Teacher-rated Aggression



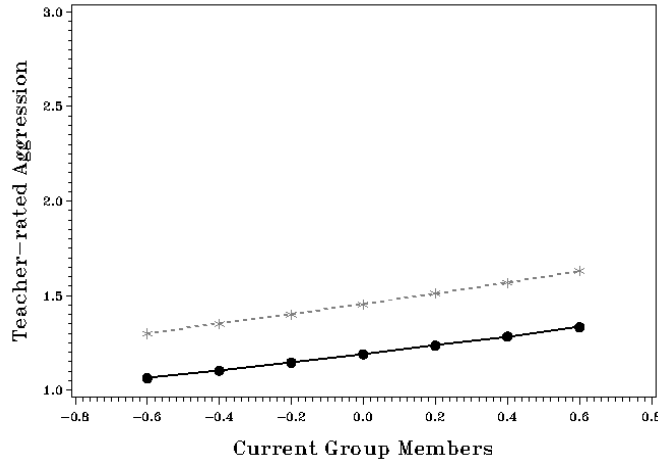
Random Grade Squared Variance as a Function of Group Members' Aggression  
LN Teacher-rated Aggression



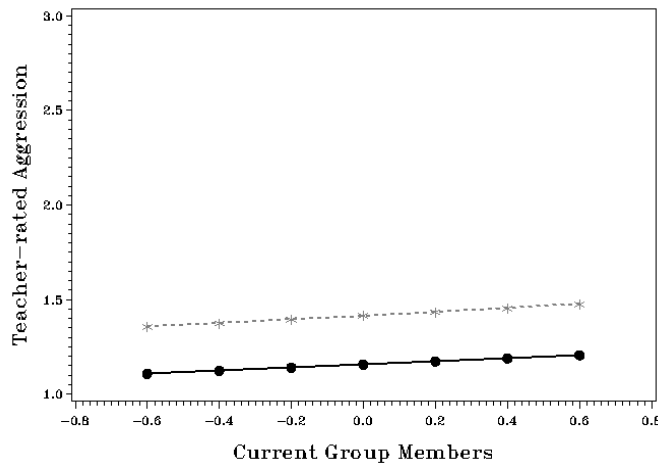
Appendix 6: Predicted Teacher-rated Aggression as a Function of Average Group Members, Current Group Members, and Cohort



*Oldest Cohort*



*Middle Cohort*



*Youngest Cohort:*

