Developmental trajectories of adolescent popularity: 
A growth curve modelling analysis

Antonius H.N. Cillessen\textsuperscript{a,}\textsuperscript{*}, Casey Borch\textsuperscript{b}

\textsuperscript{a}Department of Psychology, University of Connecticut, 406 Babbidge Road, Unit 1020, Storrs, CT 06269-1020, USA
\textsuperscript{b}Department of Sociology, University of Connecticut, Storrs, CT 06269-1020, USA

Abstract

Growth curve modelling was used to examine developmental trajectories of sociometric and perceived popularity across eight years in adolescence, and the effects of gender, overt aggression, and relational aggression on these trajectories. Participants were 303 initially popular students (167 girls, 136 boys) for whom sociometric data were available in Grades 5–12. The popularity and aggression constructs were stable but non-overlapping developmental dimensions. Growth curve models were run with SAS MIXED in the framework of the multilevel model for change [Singer, J. D., & Willett, J. B. (2003). \textit{Applied longitudinal data analysis}. Oxford, UK: Oxford University Press]. Sociometric popularity showed a linear change trajectory; perceived popularity showed nonlinear change. Overt aggression predicted low sociometric popularity but an increase in perceived popularity in the second half of the study. Relational aggression predicted a decrease in sociometric popularity, especially for girls, and continued high-perceived popularity for both genders. The effect of relational aggression on perceived popularity was the strongest around the transition from middle to high school. The importance of growth curve models for understanding adolescent social development was discussed, as well as specific issues and challenges of growth curve analyses with sociometric data.

© 2006 The Association for Professionals in Services for Adolescents. Published by Elsevier Ltd. All rights reserved.

Keywords: Peer relations; Popularity; Growth curve modelling

\textsuperscript{*}Corresponding author. Tel.: +1 860 486 3521; fax: +1 860 486 3667.
E-mail address: antonius.cillessen@uconn.edu (A.H.N. Cillessen).
Introduction

An important aspect of adolescent social competence is formed by adolescents’ relationships with peers. The study of peer relations has a long history in the developmental and educational sciences and has yielded in a rich set of findings (see Rubin, Bukowski, & Parker, 1998, for a review). Methodologically, two types of measures have been used to study peer relations: categorical measures of sociometric status types, and continuous measures of the dimensions that underlie these classifications. In the 1980s and 1990s, peer relations researchers focused primarily on understanding the correlates of sociometric categories, such as peer rejection (see, for reviews, Cillessen & Bukowski, 2000; Cillessen & Mayeux, 2004b). Recently, researchers have become more interested in also using continuous measures of peer status, such as social preference. This change is driven at least in part by developments in quantitative methods. Advanced statistical techniques for longitudinal data analysis, such as growth curve modelling, are well suited to answer important developmental questions about peer relations. Because these techniques are more easily applied using continuous measures than categorical measures, continuous measures of peer status have gained more attention.

The surge of interest in peer relations in the 1980s and 1990s was driven to a great extent by researchers’ concerns about peer rejection and the negative correlates and consequences associated with it such as aggression, distorted social cognitions, and emotional problems (see Bierman, 2004, for a review). Rejection continues to be an important focus of research. Recently, however, researchers are also increasingly interested in understanding popularity, at the opposite end of the peer relations spectrum. The interest in popularity is driven by two reasons. On the one hand, popularity is generally associated with high levels of social competence and social skills (Rubin et al., 1998). Thus, an interest in popularity fits with a general interest in positive psychological processes and competence. On the other hand, popularity is not always uniquely associated with positive traits and behaviours. Depending on how it is conceptualized, popularity is sometimes associated with a mixture of prosocial and antisocial behaviours. The second reason why peer relations researchers have been interested in popularity is to understand how potentially negative behaviours such as aggression or health risk behaviours may be linked to high status in the peer group, especially in adolescence (see Cillessen & Rose, 2005, for a review). For both reasons, popularity was also the focus of this study. We examined the peer relations and behaviours over time of students who were initially of high status when they entered the study.

Peer relations researchers interested in popularity have begun to make a distinction between two continuous dimensions of popularity: sociometric popularity (or acceptance) and peer-perceived or reputational popularity (Parkhurst & Hopmeyer, 1998). Sociometric popularity is a measure of liking of a person; perceived popularity reflects their social visibility. Consequently, the operationalization of these concepts is qualitatively distinct. Sociometric popularity is generally measured by asking participants to identify peers they “like most” and “like least.” The number of liked most minus liked least votes received is then used as a measure of sociometric popularity. Perceived popularity is generally measured by asking participants to name the peers in their grade who are the “most popular” and “least popular.” The number of most popular minus least popular votes received by each individual is then used as a measure of their perceived popularity.
While the overlap between these two constructs is relatively high by the end of elementary school, the correlation decreases rapidly throughout middle school and high school. The decrease is also faster for girls than for boys. For example, Cillessen and Mayeux (2004a) found a correlation of about .70 between continuous measures of sociometric and perceived popularity in Grade 5. Over the next 4 years, the correlation decreased systematically to about .30 for boys, and even crossed over to −.20 for girls by Grade 9.

Several studies have examined the unique correlates of sociometric and perceived popularity in cross-sectional analyses. One finding that stands out in this research is the differential association of measures of aggression with both types of popularity. Whereas aggression is generally negatively correlated with sociometric popularity, it correlates positively with perceived popularity. Rose, Swenson, and Waller (2004) found that overt and relational aggression were positively correlated with perceived popularity among third to ninth graders. Farmer and Rodkin (1996) found that physical aggression was positively related to social prominence and perceived popularity in third to sixth graders. Using open-ended descriptions as well as sociometric methods in Grades 4–8, LaFontana and Cillessen (2002) also found that aggression predicted lower sociometric popularity, but higher perceived popularity.

Although the results from these cross-sectional studies were consistent, effect sizes differed, possibly due to sample differences between the studies. Age-related changes may be hidden in cross-sectional studies that use a mixture of age groups. Cillessen and Mayeux (2004a) were able to compare the associations between aggression and popularity for representative samples of the same population in each year from Grade 5 to Grade 9. In addition to the two types of popularity (sociometric and perceived), they examined two types of aggression (physical and relational). In Grade 5, both types of aggression negatively predicted peer acceptance, and positively predicted perceived popularity or prominence in the peer group. With increasing grade levels, however, physical aggression was less predictive of both disliking and prominence. In contrast, relational aggression became more predictive of disliking and prominence over time, especially for girls. This study highlighted three important issues for the study of adolescent popularity. First, the correlates of popularity change over developmental time. Second, if aggression is considered, the type of aggression matters. Third, given the differential results for boys and girls, moderation by gender needs to be examined.

Even though these results disentangled the effects of age and sample, the analyses were still cross-sectional as they consisted of comparisons of within-grade regression results between grades. The goal of the current study was to extend these findings by more fully taking advantage of the longitudinal nature of the data. We were interested in examining trajectories of sociometric and perceived popularity, and the effects of physical aggression, relational aggression, and gender on these trajectories. We use the term ‘trajectory’ to refer to the relationship between a dependent variable and developmental time for an individual. This relationship can take many forms; it may be linear, but can also be nonlinear in various ways. In our conceptualization, a developmental trajectory is similar to what Wohlwill (1973) referred to as a ‘developmental dimension.’ Developmental trajectories exists at the level of the individual, but can also be summarized for groups of persons with similar values on other variables (e.g. gender). Graphs of developmental trajectories at the group level are similar to what Singer and Willett (2003) call ‘prototypical plots.’ Developmental trajectories are sometimes referred to as developmental ‘pathways.’ Analyses aimed at describing or understanding developmental trajectories fall under the rubric of
person-centred methods (Bergman, 2002), and have applications in various areas in developmental science (Pulkkinen & Caspi, 2002). The analyses of this paper illustrate the application of growth curve models to examine developmental trajectories in peer relations data collected with sociometric methods.

An important feature of sociometric data is that they are standardized in the reference group within which they are collected. This is necessary to control for differences in classroom or grade size that will otherwise influence the scores. However, this aspect of sociometric data poses a difficulty for longitudinal analyses: standardization of the mean to zero at each time point leads to a constant mean over time. Growth curve modelling is only possible when there is change. Peer nomination scores are unable to capture a growth process when the entire reference group is used. Therefore, for his study we focused on a subgroup of high-status adolescents who were high in popularity at the time they entered a longitudinal study. We selected students who were high on either sociometric or perceived popularity. The average peer nomination derived z-scores for this subgroup of participants is not constrained to zero over time, and allows growth curve modelling analyses. Given our interest in understanding popularity over time in adolescence, we then analysed the longitudinal data of these adolescents. The longitudinal data available for the current study were collected across eight years from Grade 5 to Grade 12.

In summary, we were interested in describing and predicting developmental trajectories of sociometric and perceived popularity of high-status adolescents across 8 years from Grade 5 to Grade 12. Our first goal was to describe the trajectories of change. Our second interest was in predicting the characteristics of the observed trajectories. Because previous cross-sectional research found important predictive effects of physical and relational aggression, our overarching goal was to illustrate growth curve modelling by examining the same predictive effects in a longitudinal context. The role of gender was examined as well. We examined linear change of popularity, but also considered the possibility that change might not be linear.

The multilevel model for change

The multilevel model for change, also known as growth curve modelling, is a flexible and powerful method for the analysis of longitudinal data. A number of important and comprehensive tutorials exist that describe this method (Bryk & Raudenbush, 1987, 1992; Lindenberger & Ghisletta, 2004; Raudenbush & Bryk, 2002; Rogosa & Saner, 1995; Singer & Willett, 2003; Snijders & Bosker, 1999). It is impossible to provide a comprehensive review within the context of this paper. Hence, we will highlight some of the most important points that are directly related to or illustrated by our example.

A multilevel model of change includes four types of variables: the outcome measure, a measure of time, one of more time-varying predictors, and one or more time-invariant predictors. The outcome variable is the dependent variable of interest that changes over time and is measured at each time point. In our example, the continuous measures of sociometric and perceived popularity were the dependent variables. Time is a predictor variable than can be measured in any number of ways (e.g. months, years, measurement wave number, trial number, see Singer & Willett, 2003). In addition to the linear effect of time, higher order effects may be included to test nonlinear effects. In our example, we used grade number as the
indicator of time, and examined both its linear and quadratic effects. Time-varying predictors are independent variables that change over time and are measured at each time point. In our example, the continuous measures of physical and relational aggression, measured at each time point from Grade 5 to Grade 12, were time-varying predictors. Time-invariant predictors are also hypothesized to influence the outcome but do not change over time. In our example, gender was a time-invariant predictor. A score for gender is coded for each person at each time point, but it remains constant over time.

Most readers are familiar with the expression of growth curve models as two-level models. Because the outcome and time-varying predictors change within persons over time, they are referred to as within-person variables and reside at Level 1. Time-invariant predictors are referred to as between-person variables and reside at Level 2.

The Level 1 model

The Level 1 model represents the estimated within-person change over time for the outcome variable in the population, and the effect of time-varying predictors on this change. Following the notation of Singer and Willett (2003), the form of this model with two time-varying predictors (as in our example) is:

\[
Y_{ij} = \pi_{0i} + \pi_{1i} TIME_{ij} + \pi_{2i} X_{2ij} + \pi_{3i} X_{3ij} + \epsilon_{ij}.
\]  

(1)

\(Y_{ij}\) is the predicted outcome for person \(i\) at time \(j\), \(TIME_{ij}\) is the value of time for person \(i\) at time \(j\), \(X_2\) and \(X_3\) are the two time-varying predictors at the within-person level, \(\pi_{0i}\) (initial status) is the value of \(Y\) when time is zero and both time-varying predictors are zero, \(\pi_{1i}\) (rate of change) is the slope of the linear trajectory of person \(i\), \(\pi_{2i}\) is the unique effect of \(X_2\) on \(Y\), \(\pi_{3i}\) is the unique effect of \(X_3\) on \(Y\), and \(\epsilon_{ij}\) is the within-person error term. The variance of this error term is estimated in the model. If this variance is large, it suggests that additional Level 1 predictors may improve the fit of the model.

In our example, \(Y\) is sociometric or perceived popularity, \(X_2\) and \(X_3\) are physical and relational aggression. More time-varying predictors can be added. If there are no time-varying predictors, Eq. (1) is:

\[
Y_{ij} = \pi_{0i} + \pi_{1i} TIME_{ij} + \epsilon_{ij}.
\]  

(2)

This is known as the unconditional growth model (UGM). This model estimates the average within-person initial status and rate of change when there are no other predictors in the model. It is important that the effect of time is statistically significant in this model. Otherwise, there is no growth and growth curve modelling is not meaningful. Thus, the UGM should be tested as a precondition for further analyses. For developmental data, the UGM fails if scores are standardized to 0 at each time point.

The Level 2 model

At Level 2, the parameters estimated at Level 1 are the outcome variables of new equations, in which the time-invariant variables are the predictors. The form of the Level 2 model for our
example with two time-varying predictors and one time-invariant predictor is:

\[ p_{0i} = \gamma_{00} + \gamma_{01} F_i + \xi_{0i}, \]

\[ p_{1i} = \gamma_{10} + \gamma_{11} F_i + \xi_{1i}, \]

\[ p_{2i} = \gamma_{20} + \gamma_{21} F_i + \xi_{2i}, \]

\[ p_{3i} = \gamma_{30} + \gamma_{31} F_i + \xi_{3i}. \] (3)

In this model, \( F \) (Female) is a dummy coded variable for the time-invariant predictor gender (girls, \( F = 1 \); boys, \( F = 0 \)). The Level 2 intercepts \( \gamma_{00}, \gamma_{10}, \gamma_{20}, \text{and} \gamma_{30} \) are the estimates of the four Level 1 parameters \( p_{0i}, p_{1i}, p_{2i}, \text{and} p_{3i} \) when all time-invariant predictors are zero. In this case, with only one time-invariant predictor and \( F = 0 \) representing boys, they are the estimates of the four Level 1 parameters for boys. The Level 2 intercepts \( \gamma_{01}, \gamma_{11}, \gamma_{21}, \text{and} \gamma_{31} \) are the effects of gender. Because of the way gender was coded, they indicate how each of the four Level 1 parameters changes for girls compared to boys. The model can be expanded with additional between-person variables. Their estimated effects will have similar interpretations. The error terms, \( \xi_{0i}, \xi_{1i}, \xi_{2i}, \text{and} \xi_{3i} \), represent individual differences in the Level 1 parameters that are not explained by the Level 2 predictors. The variances and covariances of these error terms are also estimated in the model. A large error variance for a dependent variable indicates that additional between-person variables may improve the prediction of that variable.

**Building and comparing models through measures of model fit**

As shown in our example, the goal of statistical modelling is to find the model that best fit the data. Typically, a sequence of models is tested, beginning with the unconditional means model (UMM) and the UGM. The goal of the UMM is to test whether there is sufficient variability in individuals’ average scores on the dependent variable (averaged over time) for the analyses to proceed. The goal of the UGM is to test whether there is sufficient variability in the data over time. Positive answers to both questions are a precondition for further analysis. If both preconditions are met, further model building takes place with additional predictor variables.

If two models are nested, they can be compared using the deviance statistic. The model with the smaller deviance fits the data better and should be preferred. Two models are nested if one model can be obtained by constraining one of more parameters of the other model. The simpler model (with fewer parameters) will have a smaller deviance than the more complex model (with more parameters). The simpler model is preferred if the difference in deviance between both models is statistically significant. The test of statistical significance is a \( \chi^2 \) test with degrees of freedom equal to the difference in the number of parameters between both models. If two models are not nested, they can be compared with the Akaike information criterion (AIC) and/or the Bayesian information criterion (BIC). The logic here is the same—smaller values are preferred. However, there is no test of statistical significance for the difference of these fit statistics between two models. In our example, we reported all three criteria for model fit for each model and use them to make comparisons between models.

**Centering time**

Centering involves the recoding of a predictor variable to either give the intercept a meaningful interpretation or to address specific issues that are impractical with predictors in their original
Centering the predictor used to represent time involves subtracting a constant from each value for time to facilitate the interpretation of the intercept. In our data, time was initially represented by grade coded as a number from 5 to 12. In an analysis with this variable, the intercept would return estimates for students in Grade 0, which does not exist. In order for the intercept to be meaningful, we must recode time so that 0 will be assigned to one of the existing grade levels. Which grade is selected depends on substantive reasons. If the interest primary interest is in estimating and predicting the dependent variable at the beginning of time (as in a developmental study of subsequent growth), time should be centred at the first time point. In our example, this would be achieved by creating a new variable for time that is grade minus five (range 0–7). This new variable is then used as the predictor in the analyses. If the interest is in estimating and predicting the dependent variable at the end of the longitudinal trajectory (as in an outcome study), time should be centred at the last time point. In our example, this would be achieved by creating a new variable for time that is grade minus 12 (range –7 to 0). It is also possible to centre time in the middle of the longitudinal trajectory, for example, to represent a time point before or after a school transition. In our example, if we were particularly interested in estimating and predicting students’ popularity in their first year after the transition from middle to high school, we could centre time at Grade 9. In our current example, we choose to centre time at the beginning of the study (Grade 5).

Missing data

A longitudinal data set is balanced when all participants have scores for all variables at all time points. Clearly, this is the ideal in longitudinal research. In practice, however, missing data occur frequently in longitudinal research due to a various reasons, such as attrition, absenteeism on data collection days, or the fact that participants may enter an ongoing longitudinal study at different time points. All of these events are common in the practice of longitudinal field research in developmental science. An advantage of the multilevel model for change is that models can still be estimated even when the data set is not perfectly balanced (Singer & Willett, 2003). Even if individuals vary in the number of time points to which they contributed data, or if the spacing of data points differs between study participants, analysis is still possible. Our example also illustrates this feature of the multilevel model for change.

Method

Participants and procedure

Participants were 303 students (167 girls, 136 boys), selected from the larger sample of an ongoing longitudinal study on the social and academic development of children and youth. The larger study took place in the public school system of a medium-sized town in the Northeastern
United States. In this project, data collection took place yearly in 11 consecutive school years over the period 1995–2006 as participants were followed from Grade 4 to Grade 12 and in two follow-up data collections after high school. Data were collected using peer-, self-, and teacher-report instruments. The present study focused on sociometric data collected in the spring of the eight consecutive school years from Grade 5 to Grade 12.

In Grade 5, students were in 10 elementary schools, feeding into two middle schools (Grades 6–8), and then into one high school (Grades 9–12). In each year, all students of the grade were invited to participate. The number of participants in each year was 639, 598, 599, 607, 586, 542, 491, and 481, for Grades 5–12, respectively. The participation rate ranged from 67% to 99%; the lower percentages were in the higher grades. Thus, at least two thirds of the grade cohort participated in sociometric measurement in each year. The proportion of girls varied between 48% (Grade 5) and 54.5% (Grade 12). The ethnic composition of the larger sample was 66% White, 21% African American, 11% Latino, and 2% of other origin.

Not all students participated in all waves of data collection. Changes in the sample composition occurred predominantly at the two school transitions included in the study from elementary to middle school, and from middle to high school. At both transition points, some students did not return from the previous year because they went to a different middle or high school, whereas new students came into the school system.

Selection of subsample

In each year of the study, a continuous measure of sociometric popularity and a continuous measure of perceived popularity were determined (see below). The interest of this paper was on the developmental trajectories of aggression of students who were initially popular. To be included in the sample of this study, students were selected who had a sociometric or perceived popularity score larger than .5 standard deviation above the mean at the first year they entered the study. Based on this criterion, 303 students were identified (167 girls, 136 boys). These students then formed the sample for the analyses of this study. Of the 303 students in the sample, 252 (133 girls, 119 boys) entered the study in Grade 5. The remaining 51 students entered the study at a later time (16, 8, 7, 9, 6, 2, and 3, in Grades 6–12, respectively).

Sociometric measurement

Sociometric data were collected using a group-administered peer nomination instrument in the spring of each school year. As indicated above, the proportion of students in each grade that participated in the sociometric assessment was 67% of larger. A minimum participation rate of 70% is considered acceptable for sociometric assessment when limited nominations are used (Crick & Ladd, 1989). With unlimited nominations, this proportion is lower. Nukulkij (2000) found reliable and stable sociometric scores with unlimited nominations and a minimum participation rate of 60%. Thus, the participation rate in our study was considered large enough to yield reliable sociometric data.

Unlimited nominations were used with the entire grade as the reference group, allowing both same-sex and cross-sex nominations. Four items were used to measure popularity: liked most (“the people in your grade you like the most”), liked least (“the people in your grade you like the
least”), most popular (“the people in your grade who are the most popular”), and least popular (“the people in your grade who are the least popular”). The wording of the items was identical in all study years. Nominations received were counted for each question and standardized within grade. A measure of sociometric popularity was computed by computing the difference between the standardized liked most votes received minus the standardized liked least votes received for each participant. A measure of perceived popularity was computed by computing the difference between the standardized most popular votes received minus the standardized least popular votes received for each participant. Both difference scores were again standardized within each grade to yield the final continuous measures of sociometric and perceived popularity.

Overt aggression was measured with one item that was identical in all grades (“start fights, say mean things, and/or tease others”). Relational aggression was measured with two items in Grade 5 (“keep others from being in the group during activities or games,” “ignore or stop talking to other kids when they are mad at them”), one in Grade 6 (“ignore others or spread rumours about them when they are mad at them”), and two in Grades 7–12 (“ignore others or spread rumours about others when they are mad at them,” “try to keep others who they don’t like from being in their group”). The measure of overt aggression was the standardized number of overt aggression nominations received in each grade. The measure of relational aggression was the standardized number of relational aggression nominations received (Grade 6) or the average of the standardized nominations of the two relational aggression items (all other grades). (The correlation between the two relational aggression items in each grade ranged from .40 to .60.)

Results

Stability of popularity and aggression constructs

As preliminary analyses, we first examined the stabilities and intercorrelations of the main study variables. To examine stability, correlations were computed for each study variable across all years for the selected sample of this study. These stability coefficients are presented by gender in Table 1 (girls) and Table 2 (boys). As can be seen, almost all stability coefficients were reliably different from zero. Furthermore, the stabilities followed a predictable pattern. Higher correlations were found across shorter intervals, and stabilities across the same interval length were higher when determined later in the developmental trajectory of this study.

Gender differences were examined using Fisher’s r-to-Z tests for independent correlations. Tables 1 and 2 indicate which correlations differed by gender. There were no gender differences in the stability of sociometric popularity. For perceived popularity, 11 correlations were reliably different by gender; all 11 were larger for girls than for boys. For overt aggression, 18 correlations differed by gender—all 18 were larger for boys. For relational aggression, seven stabilities differed by gender: five favoured girls, and two favoured boys.

Associations between popularity and aggression constructs

Table 3 presents the intercorrelations between popularity and aggression constructs by gender and grade for the selected study participants. Consistent with previous findings (Cillessen &
Mayeux, 2004a, b), the association between sociometric and perceived popularity in the total sample of this study declined over time and became negative by Grade 9. When examined by gender, the correlation for boys remained positive, decreasing from .58 in Grade 6 to .03 in Grade 10, before rebounding to .30 by Grade 12. For girls, the correlation reversed in sign over time, from a high of .41 in Grade 6 to .49 in Grade 12 (see Table 3).

The correlation between overt and relational aggression was consistently high for both genders. For girls, the correlation increased from .63 in Grade 5 to .83 in Grade 12. For boys, the correlation decreased from .85 in Grade 5 to .67 in Grade 7, before rebounding to .90 in Grade 10. From Grade 5 to Grade 10, the correlation between both types of aggression was lower for girls than for boys. This trend reversed in Grade 11, and in Grade 12 the correlation was reliably higher for girls than for boys (see Table 3).

**Describing and predicting change in popularity: applying the multilevel model for change**

The multilevel model for change was applied to the selected sample for this study. Five models were tested, labelled A–E. These analyses were run using PROC MIXED in SAS. The five models were tested with sociometric popularity as the dependent variable and again with perceived popularity as the dependent variable. Model A was the UMM in each case; Model B was the UGM in each case. Models C–E were theoretical models in which the effects of the substantive

### Table 1

Stability of main study variables across 8 years for girls (91 < N < 132)

<table>
<thead>
<tr>
<th></th>
<th>Grade 5</th>
<th>Grade 6</th>
<th>Grade 7</th>
<th>Grade 8</th>
<th>Grade 9</th>
<th>Grade 10</th>
<th>Grade 11</th>
<th>Grade 12</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Popularity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 5</td>
<td>.43*</td>
<td>.46*</td>
<td>.22*</td>
<td>.26*</td>
<td>.12</td>
<td>.12</td>
<td>.06</td>
<td></td>
</tr>
<tr>
<td>Grade 6</td>
<td>.66*</td>
<td>.65*</td>
<td>.45*</td>
<td>.36*</td>
<td>.31*</td>
<td>.13</td>
<td>.08</td>
<td></td>
</tr>
<tr>
<td>Grade 7</td>
<td>.63*</td>
<td>.90*</td>
<td>.63*</td>
<td>.50*</td>
<td>.39*</td>
<td>.24*</td>
<td>.23*</td>
<td></td>
</tr>
<tr>
<td>Grade 8</td>
<td>.69*</td>
<td>.89*</td>
<td>.94*</td>
<td>.51*</td>
<td>.62*</td>
<td>.51*</td>
<td>.46*</td>
<td></td>
</tr>
<tr>
<td>Grade 9</td>
<td>.54*</td>
<td>.79*</td>
<td>.85*</td>
<td>.84*</td>
<td>.72*</td>
<td>.54*</td>
<td>.46*</td>
<td></td>
</tr>
<tr>
<td>Grade 10</td>
<td>.43*</td>
<td>.72*</td>
<td>.77*</td>
<td>.76*</td>
<td>.91*</td>
<td>.70*</td>
<td>.62*</td>
<td></td>
</tr>
<tr>
<td>Grade 11</td>
<td>.35*</td>
<td>.65*</td>
<td>.65*</td>
<td>.81*</td>
<td>.93*</td>
<td>.70*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 12</td>
<td>.36*</td>
<td>.63*</td>
<td>.63*</td>
<td>.78*</td>
<td>.89*</td>
<td>.94*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| **Aggression** |         |         |         |         |         |         |         |         |
| Grade 5 | .74*    | .73*    | .53*    | .27*    | .20     | .21     | .21     |         |
| Grade 6 | .49*    | .89*    | .67*    | .30*    | .60*    | .13     | .18     |         |
| Grade 7 | .48*    | .57*    | .61*    | .33*    | .12     | .29*    | .24*    |         |
| Grade 8 | .45*    | .68*    | .82*    | .53*    | .58*    | .43*    | .38*    |         |
| Grade 9 | .35*    | .50*    | .70*    | .78*    | .63*    | .63*    | .50*    |         |
| Grade 10| .28*    | .50*    | .58*    | .68*    | .86*    | .59*    | .63*    |         |
| Grade 11| .23*    | .45*    | .52*    | .61*    | .77*    | .85*    | .61*    |         |
| Grade 12| .26*    | .42*    | .54*    | .70*    | .77*    | .86*    | .80*    |         |

*Note.* *p* < .05. Popularity: sociometric above the diagonal, perceived below the diagonal. Aggression: overt above the diagonal, relational below the diagonal. Correlations that are italicized were significantly different by gender.
predictor variables of interest in this study (gender, overt aggression, relational aggression) were tested. Table 4 summarizes the results of the model tests for sociometric popularity. Table 5 summarizes the results of the model tests for perceived popularity. The unconditional means model

Model A is the UMM. This model consists of one Level 1 Eq. (4) and one Level 2 Eq. (5). The Level 1 model has an intercept but no slope. It assumes that the change trajectory for all persons is a straight line and that there is no change over time. If this model is not rejected, there is no evidence for longitudinal change, and further model testing is not useful. The general form of the UMM is:

\[ Y_{ij} = \pi_{0i} + \epsilon_{ij}, \]  

(4)

\[ \pi_{0i} = \gamma_{00} + \zeta_{0i}, \]  

(5)

where \( \pi_{0i} \) is the mean of \( Y \) for person \( i \) (within-person mean), \( \gamma_{00} \) is the mean of \( Y \) across all persons in the sample (grand mean), and \( \epsilon_{ij} \) is the difference between person \( i \)'s score at time \( j \) and the within-person mean. The variance of this term is the within-person variance (\( \sigma^2_{\epsilon} \)). If the within-person variance is statistically significant, there is substantial variation of individuals' scores over
time around their mean, and it is useful to add time-varying predictors to the model. Finally, $\zeta_{0i}$ is the difference between a person’s mean and the grand mean. The variance of this term is the between-person variance in initial status ($\sigma_{0i}^2$). If this is statistically significant, there is substantial variation in initial status due to individual differences (e.g. girls may score higher than boys on the dependent variable at the beginning of the study). This is an indication that it may be useful to add time-invariant predictors to the model.

Thus, in Model A, we are looking for: (1) model fit statistics indicating that a model that assumes no change over time is a poorly fitting model, (2) statistically significant within-person variation, indicating that time-varying predictors may be at play, and (3) statistically significant between-person variation, indicating that individual differences in starting values may be due to time-invariant predictors.

Table 4 shows that for sociometric popularity the grand mean (a fixed effect) is reliably different from zero ($\gamma_{00} = .629$, $p < .001$). For perceived popularity (Table 5) the grand mean was also reliably different from zero ($\gamma_{00} = .572$, $p < .001$). This makes sense as the sample consisted of students who were relatively popular when they entered the study.

The more important results of Model A are found in the estimates of the variance components (random effects). For both sociometric and perceived popularity, the estimated within-person variances were reliably different from zero ($\sigma_{2}^2 = .524$ and .333, respectively, both $p$’s < .001). This indicates statistically significant change over time in both measures of popularity. For both measures, the estimated between-person variances also differed reliably from zero ($\sigma_{0}^2 = .333$ and

<table>
<thead>
<tr>
<th>Overall ($N = 303$)</th>
<th>Girls ($N = 167$)</th>
<th>Boys ($N = 136$)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sociometric and perceived popularity</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 5</td>
<td>.18*</td>
<td>.20*</td>
</tr>
<tr>
<td>Grade 6</td>
<td>.47*</td>
<td>.41*</td>
</tr>
<tr>
<td>Grade 7</td>
<td>.29*</td>
<td>.19*</td>
</tr>
<tr>
<td>Grade 8</td>
<td>.22*</td>
<td>.19*</td>
</tr>
<tr>
<td>Grade 9</td>
<td>-.04</td>
<td>-.16</td>
</tr>
<tr>
<td>Grade 10</td>
<td>-.30*</td>
<td>-.47*</td>
</tr>
<tr>
<td>Grade 11</td>
<td>-.15</td>
<td>-.42*</td>
</tr>
<tr>
<td>Grade 12</td>
<td>-.24*</td>
<td>-.49*</td>
</tr>
<tr>
<td><strong>Overt and relational aggression</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 5</td>
<td>.69*</td>
<td>.63*</td>
</tr>
<tr>
<td>Grade 6</td>
<td>.69*</td>
<td>.62*</td>
</tr>
<tr>
<td>Grade 7</td>
<td>.53*</td>
<td>.46*</td>
</tr>
<tr>
<td>Grade 8</td>
<td>.65*</td>
<td>.66*</td>
</tr>
<tr>
<td>Grade 9</td>
<td>.40*</td>
<td>.61*</td>
</tr>
<tr>
<td>Grade 10</td>
<td>.57*</td>
<td>.64*</td>
</tr>
<tr>
<td>Grade 11</td>
<td>.48*</td>
<td>.73*</td>
</tr>
<tr>
<td>Grade 12</td>
<td>.44*</td>
<td>.83*</td>
</tr>
</tbody>
</table>

*Note. $* p < .05. Correlations that are italicized were significantly different by gender.
.497, respectively, both p’s < .001), indicating statistically significant individual differences in average popularity. Because both variance components were not zero, additional predictors at both Level 1 and Level 2 may improve model fit.

Results from the UMM can also be used to gauge the relative magnitude of the $\sigma^2_e$ and $\sigma^2_0$ variance components by computing the intraclass correlation coefficient (ICC), or $\rho$ (Singer & Willett 2003). Without a statistically significant intraclass correlation, multilevel modelling is not an improvement over traditional way of estimating fixed effects (de Leeuw & Kreft, 1995). The ICC is a measure of the proportion of total variation in the dependent variable that is among persons. Formally, $\rho = \sigma^2_0 / (\sigma^2_e + \sigma^2_0)$. For sociometric popularity, the ICC was $\rho = .333 / (.524 + .333) = .389$. This indicates that about 40% of the total variation in sociometric popularity was due to differences between students. For perceived popularity, the ICC was $\rho = .497 / (.333 + .497) = .599$. This indicates that nearly 60% of the variation in perceived popularity was due to individual differences.

The ICC also is a measure of the average autocorrelation of the dependent variable over time (Singer & Willett, 2003). Thus, the estimated average stability of sociometric popularity was .39,
while the estimated stability of perceived popularity was .60. Thus, perceived popularity is more stable than sociometric popularity.

The unconditional growth model

Model B is the UGM. This model adds time as a Level 1 predictor of $Y$, but adds no predictors at Level 2 (the Level 2 model is left “unconditional”):

$$ Y_{ij} = \pi_{0i} + \pi_{1i} \text{TIME}_{ij} + \epsilon_{ij}, $$

$$ \pi_{0i} = \gamma_{00} + \zeta_{0i}, $$

$$ \pi_{1i} = \gamma_{10} + \zeta_{1i}, $$
In this model, $Y$ is a function of an error term ($e_{ij}$), an intercept ($\pi_{0i}$), and an individual growth parameter ($\pi_{1i}$). The latter two are the sum of an intercept (either $\gamma_{00}$ or $\gamma_{10}$) and a residual (either $\zeta_{0i}$ or $\zeta_{1i}$). In the Level 2 model, there are two random effects, $\zeta_{0i}$ and $\zeta_{1i}$, with variance $\sigma_{\zeta_0}^2$ and $\sigma_{\zeta_1}^2$, respectively, and a covariance $\sigma_{\zeta_0\zeta_1}$. In this case, the covariance estimates the association between initial status and rate of change. The variances and covariance can be used to determine the correlation between initial status and rate of change. In a linear model, a negative correlation suggests that a low initial status is associated with a steeper slope (stronger gains over time), and a high initial status is associated with a shallower slope (less change).

The second columns of Tables 4 and 5 provide the results for Model B. TIME was computed as Grade 5 so that TIME = 0 at Grade 5 and TIME = 7 at Grade 12. Thus, time was “centred” at the beginning of the study to facilitate interpretation of the parameters. In this case, Grade 5 is the baseline against which the other grades were compared.

**Sociometric popularity**

The UGM was first tested with linear time only as specified in the previous section. To examine the possibility of nonlinear effects, a second model was tested with quadratic time added, and a third model with cubic time added. Neither the quadratic nor the cubic model improved model fit. Therefore, the model with linear time was retained, and linear change of sociometric popularity was modelled in all subsequent analyses.

In the UGM for sociometric popularity, initial status (intercept) was $\gamma_{00} = .845$ and the rate of change (slope) was $\gamma_{10} = -.081$. Both were reliably different from zero. The difference from zero for initial status was expected because the sample consisted of adolescents who were relatively popular at the beginning of the study. The negative effect of TIME indicates that they became less sociometrically popular over time. The linear change is shown in Fig. 1.

In this model, the Level 1 residual variance $\sigma_e^2$ measures the scatter of students’ data around their linear change trajectory. As a result of adding TIME, $\sigma_e^2$ was reduced from .524 in Model A
to .379 in Model B; a reduction of nearly 30%. However, the within-person variance was still reliably different from zero, meaning that other time-varying predictors should be added to the model at Level 1.

Level 2 residual variances ($\sigma_0^2$ and $\sigma_1^2$) summarize the between-person variability in initial status and rate of change. At Level 2, both $\sigma_0^2 = .263$ and $\sigma_1^2 = .022$ were reliably different from zero, indicating that additional Level 2 predictors will improve model fit.

The covariance, $\sigma_{01} = -.015$, serves two purposes in the UGM. It represents the relationship between the Level 2 residuals, and measures the covariance between the estimated initial status and rate of change. It is easier to interpret this term as a correlation, computed by dividing the covariance by the square root of the product of its associated variance components. Thus, $r = -.015/(.263 \times .022)^{-1} = -.197$. This indicates that the relationship between initial status and rate of change in sociometric popularity was negative although small in size.

To test whether the addition of time as a Level 1 predictor improved the model, the fit of the UMM and UGM were compared. Because the UMM is nested within the UGM, the deviance statistic can be used. Since the UGM estimates three more parameters ($\gamma_{10}$, $\sigma_1^2$, and $\gamma_{01}$), the $\chi^2$ difference test has three degrees of freedom. The reduction in deviance due to the addition of time was statistically significant, $\chi^2(3) = 3928.1 - 3667.7 = 260.4, p < .001$.

**Perceived popularity**

The UGM was again first tested with linear time only, followed by two models with quadratic and cubic time added. The quadratic model reliably improved model fit over the linear model, $\Delta \chi^2(4) = 62.00, p < .001$. The cubic model did not further improve model fit. Therefore, quadratic change of perceived popularity was modelled in all subsequent analyses. (As required, the underlying linear term was also consistently retained.)

For perceived popularity, initial status was $\gamma_{00} = .830$, the linear rate of change was $\gamma_{10} = -.181$, and the quadratic term was $\gamma_{20} = .019$. All three were reliably different from zero. The quadratic trend followed a U-shaped curve shown in Fig. 1. As can be seen, perceived popularity decreased from Grade 5 to Grade 9, but increased again from Grade 10 to Grade 12. Adding the trend terms reduced the within-person variance by 33% from .333 to .222, but it remained different from zero (see Table 5). Thus, additional Level 1 predictors may further improve model fit. At Level 2, both $\sigma_0^2 = .559$ and $\sigma_1^2 = .017$ were reliably different from zero, meaning that additional Level 2 predictors may also improve model fit. Because we used a quadratic trend term in the model, the covariance is not easily interpreted as the association between the initial status and rate of change. Finally, the UGM yielded a statistically significant increase in model fit over the UMM, $\chi^2(4) = 3404.4 - 3112.8 = 291.6, p < .001$.

**Main effects of aggression and gender**

Model C was the first theoretical model in which the effects of gender (dummy coded, girls = 1, boys = 0), overt aggression, and relational aggression were added to the best fitting UGM. In Model C, the added variables can have an effect on the initial status in Grade 5 and linear change of both dependent variables, and on the quadratic change of perceived popularity. The results are shown in the Model C column in Tables 4 and 5.
Sociometric popularity

The intercept ($\gamma_{00} = .903$) is the average estimated sociometric popularity for boys (gender = 0) in Grade 5 when all other predictors are zero. Again and not surprisingly, given the selection of study participants, this intercept was reliably different from zero. Substituting 1 for gender in the equation for Level 2 yields the corresponding value for girls. With other predictors set to zero, the average initial popularity for girls was $\gamma_{01} + \gamma_{00} = .903 + (-.113) = .790$. Because the effect of gender on initial status was not statistically significant, the estimated means of sociometric popularity in Grade 5 did not differ between boys and girls.

Relational and overt aggression had statistically significant negative effects on the initial status of sociometric popularity. Higher levels of relational or overt aggression during the study were associated with less social acceptance by peers at the beginning of the study in Grade 5. The effect of linear time was statistically significant and negative (as in Model B), indicating that peer acceptance decreased linearly over the course of this study for this initially popular sample.

For the random effects, the addition of the two Level 1 predictors reduced the within-person variance ($\sigma_v^2$) from .379 in the UGM to .369. This proportion was still reliably different from zero, suggesting that additional Level 1 predictors may improve model fit. Because the UGM has no Level 2 predictors, it is not surprising that the inclusion of gender at Level 2 reduced the variance in initial status from .263 to .238 and in rate of change from .022 to .020. Because these variances were still reliably different from zero, adding Level 2 predictors may also improve model fit. Finally, Model C fit the data better than the UGM, $\chi^2(3) = 3667.7 - 3585.2 = 82.5, p < .001$.

Perceived popularity

The intercept was again reliably different from zero ($\gamma_{00} = .842$), indicating that the initial status of boys at average levels of both types of aggression (because they were z-scores, 0 is average) was high. The estimated difference for girls ($\gamma_{01} = -.061$) was again not reliably different from zero. Thus, there was no gender difference in initial perceived popularity in Grade 5. The quadratic change term again was evident, indicating that perceived popularity declines at a decreasing rate over time (i.e. the slope is U shaped).

The effects of both Level 1 predictors, overt and relational aggression, were statistically significant. Both types of aggression predicted higher levels of initial perceived popularity (see Table 5). Note, however, that relational aggression had a much stronger effect than overt aggression. Adding the Levels 1 and 2 predictors considerably lowered the within-person variance and the variances in initial status and rate of change. Model C fit the data reliably better than the UGM, $\chi^2(3) = 3112.8 - 2761.7 = 351.1, p < .001$.

Moderation by gender

In Model D, we tested whether the effects that were statistically significant in Model C were moderated by gender. Thus, moderation by gender was examined for the linear and quadratic rates of change, and for the effects of both types of aggression on initial status.

To test for moderation of the rates of change by gender, the interactions of gender with linear time and quadratic time (for perceived popularity) were included in the model. Moderation by gender would exist if the linear trend in sociometric popularity would differ for boys and girls, or the curvilinear trend in perceived popularity would vary by gender. However, neither of
these interactions was statistically significant, indicating that the linear decrease of sociometric popularity and the curvilinear decrease of perceived popularity were the same for boys and girls. Because adding these moderator effects did not improve model fit for either sociometric popularity or perceived popularity, these effects were not included in Tables 4 and 5.

To test whether the effects of aggression on initial status were moderated by gender, interaction terms between gender and each type of aggression were added to the model. This illustrates that interactions are possible between predictors across levels, in this case, between gender (a Level 2 variable) and each of the aggression constructs (Level 1 variables). Addition of these effects did improve model fit as described below.

**Sociometric popularity**

Adding the interactions between gender and the two types of aggression reduced the within-person, initial status, and rate of change variance. Model D fit the data better than Model C, \( \chi^2(2) = 3585.2 - 3554.0 = 31.2, p < .001 \). The interaction between gender and relational aggression was not statistically significant, but the interaction between gender and overt aggression was (Table 4). Post-hoc testing showed that, consistent with the literature, overt aggression had a stronger negative effect on initial acceptance for girls than for boys. Details of the post-hoc probing are shown in the prototypical plots section below.

**Perceived popularity**

Model D reduced the within-person variance, but the initial status and rate of change variance remained the same. Model D fit the data better than Model C, \( \chi^2(2) = 2761.7 - 2751.5 = 10.2, p < .05 \). The interaction between gender and overt aggression was statistically significant. Again, post-hoc probing showed that overt aggression had a stronger negative effect on initial perceived popularity for girls than for boys. Details of the post-hoc probing procedure are presented below.

**Effects of aggression on rate of change**

In Model E, two further effects were added. We tested whether overt and relational aggression also influenced the rate of change of sociometric and perceived popularity over time. Model E for both outcome variables is shown in the last column of Tables 4 and 5.

**Sociometric popularity**

As can be seen in Table 4, adding the effects of overt and relational aggression on the linear rate of change improved model fit for sociometric popularity, \( \chi^2(2) = 3554.0 - 3528.2 = 25.8, p < .001 \). Furthermore, both effects were statistically significant. To further understand the effects, prototypical plots were created and are shown in the next section. To complete the model tests, we examined whether the effects of both types of aggression on the linear change of sociometric popularity were further moderated by gender. The addition of these moderator terms did not improve model fit, \( \chi^2(2) = 3528.2 - 3527.3 = .9, p = .638 \), and both effects were not statistically significant. Therefore, these effects are not further shown in Table 4.
Perceived popularity

Because perceived popularity changed nonlinearly, the effects of overt and relational aggression on both the linear and the quadratic rate of change were added in Model E. As can be seen in Table 5, adding these four effects improved model fit, $\chi^2(4) = 2751.5 - 2713.8 = 37.7$, $p < .001$. Furthermore, all four effects were statistically significant. Prototypical plots of these effects are shown in the next section. Finally, we examined whether the effects of the two types of aggression on both linear and quadratic change of sociometric popularity were further moderated by gender. The addition of these four terms did not improve model fit, $\chi^2(4) = 2713.8 - 2709.5 = 4.3$, $p = .367$, and none of the effects were statistically significant. Therefore, they are not further shown in Table 5.

Prototypical plots

As our example may illustrate, effects are often difficult to interpret from parameter estimates alone. An important aid towards understanding effects in the multilevel model for change is the creation of prototypical plots (Singer & Willett, 2003). A prototypical plot is a graph of the trajectory of a dependent variable for selected values of the predictors. If a predictor is categorical, such as gender, separate lines are plotted for each group. If a predictor is continuous, lines are drawn for representative values. Often, the mean, mean plus 1 standard deviation, and mean minus 1 standard deviation are chosen to represent medium, high, and low values of a predictor. Other values may be chosen based on substantive considerations. The procedure for creating prototypical plots is thus identical to the procedure of plotting trends or interactions in regression (see Aiken & West, 1991). The full equation that results from an estimated model is written out, and values of the predictor(s) are systematically substituted to obtain predicted scores for each combination of predictor values. There is one difference between multiple regression and the multilevel model for change. Because the latter includes more than one equation, the Level 2 equations first need to be substituted into the Level 1 equation to create a composite specification (Singer & Willett, 2003). Once the composite equation is known, it can be used to create prototypical plots in the same way as for regression.

Fig. 2 shows the prototypical plots for the effects of gender and overt aggression (Fig. 2a) and gender and relational aggression (Fig. 2b) on sociometric popularity. These prototypical plots were created using the equation for Model E in Table 4. Gender was dummy coded as before ($1 =$ girls, $0 =$ boys). For overt aggression, $M \pm 1$ s.d. ($0.029 \pm 0.941$) were chosen to plot the prototypical trajectory of adolescents high and low in overt aggression. For relational aggression, $M \pm 1$ s.d. ($0.188 \pm 1.053$) were chosen to plot the prototypical trajectory of adolescents high and low in relational aggression.

Fig. 2a shows the main effect of overt aggression on initial sociometric popularity, further qualified by gender. Overt aggression predicted lower peer acceptance, especially for girls. Fig. 2a shows the effect of overt aggression on the rate of change. The overall decrease in peer acceptance documented earlier (see Fig. 1) was moderated by overt aggression. More overt aggression predicted a slower decrease in acceptance. While this may seem counterintuitive at first, it makes sense given that overt aggression also predicted lower initial values of peer acceptance (see Fig. 2a).
Fig. 3 shows the main effect of relational aggression on initial sociometric popularity, which was not moderated by gender. Interestingly, relational aggression predicted higher starting values of peer acceptance at Time 1 (Grade 5). Importantly, relational aggression also influenced the overall decrease in peer acceptance over time. For adolescents who were not relationally aggressive, peer acceptance stayed constant over time. However, more relational aggression predicted a stronger decrease in peer acceptance. The decrease was the same for boys and girls, but because girls had lower starting values, the lowest peer acceptance scores were obtained for relationally aggressive girls from Grade 7 on.

Fig. 2 shows the prototypical plots for the effects of gender and overt aggression (Fig. 3a) and gender and relational aggression (Fig. 3b) on perceived popularity, created using the equation for Model E in Table 5. The same values of gender, high and low overt aggression, and high and low relational aggression were used as for Fig. 2.
Fig. 3a shows the main effect of overt aggression on initial perceived popularity, further qualified by gender. Overtly aggressive boys started with the highest perceived popularity scores in Grade 5, whereas the estimates of initial status for overtly aggressive girls and children who were not overtly aggressive were about the same. Importantly, overt aggression moderated the change trajectory of perceived popularity. At low levels of overt aggression, perceived popularity decreased linearly over time. At higher levels of overt aggression, however, there is a steeper decrease in the first half of the time span covered by this study, but scores rebounded in the second half. Starting in Grade 9, overt aggression actually increased adolescents’ perceived popularity.

Fig. 3b shows the main effect of relational aggression in initial status. Students who were relationally aggressive had higher perceived popularity scores in Grade 5 than students who were not relationally aggressive. Relational aggression also influenced the change trajectory. At high levels of relational aggression, perceived popularity tended to decline although the scores remained high. At low levels of relational aggression, relational aggression declined steeply in the
first half of the study, but rebounded in the second half. Notice that the effects of relational aggression on initial status and rate of change were not moderated by gender; the plotted trajectories for boys and girls were virtually identical. Notice also that the effect of relational aggression on perceived popularity peaks in Grades 8 and 9, when the difference between the top two and bottom lines in Fig. 3b is the largest.

Discussion

Longitudinal research is the “royal road” to understanding developmental processes. Wohlwill (1973) already specified the ideas behind causal and multilevel modelling approaches to longitudinal data analysis when the statistical methods for these analyses did not yet exist. Today, researchers have available both the conceptual advances made by scholars such as Wohlwill and the technology to analyse the questions he envisioned.

For example, growth curve modelling enables researchers to examine the shape of developmental trajectories, to test the effect of predictors on the starting or ending values of developmental trajectories, as well as the effect of predictor variables on acceleration, deceleration, or other variations of growth. Growth curve models can assess the impact of developmental transitions and can model growth as a continuous or discontinuous process. It is possible to examine whether the effect of an event is developmentally invariant or depends on the moment in developmental time when it occurs. These questions cannot be address with traditional multivariate methods, but are important substantive questions in many domains of developmental science.

Adolescent social development is one of these domains for which longitudinal data and growth curve analyses are of great importance. There are important early developmental predictor variables (e.g. dimensions of parenting) that determine the course of trajectories of social adjustment in adolescence. The progression of important developmental dimensions, such as antisocial behaviour or concerns for physical appearance or popularity is often assumed to be nonlinear and moderated by gender or other between-person variables. The status of academic achievement by the end of adolescence is an outcome many scholars and practitioners alike are eager to predict. Further knowledge about each of these phenomena can be achieved through proper analysis of longitudinal data.

The same applies to researchers interested in peer relations, as one other important domain of adolescent social development. The current study shows how growth curve modelling can be used to describe and predict trajectories of popularity across adolescence.

This study focused on two specific types of popularity, sociometric and peer-perceived, and their trajectories across Grades 5 to 12. The first goal of this study was to illustrate the application of growth curve modelling to an important aspect of adolescent development. The second goal was to extend previous work on the predictors of adolescent popularity, by focusing on the role of overt aggression, relational aggression, and gender.

Preliminary analyses indicated that the constructs of this study were relatively stable over time, with measures closer in time more highly correlated. The intercorrelations demonstrated that perceived popularity was more stable than sociometric popularity or both boys and girls. This was confirmed by the ICCs computed from the UGMs, with average stability estimated of .40 for
sociometric popularity and .60 for perceived popularity. The differences makes sense, because perceived popularity is a reputation that is likely to be observed by all members of the peer group, whereas sociometric popularity is a reflection of individual peers’ liking judgments, which are more likely to vary from person to person.

The stabilities of overt and relational aggression differed by gender. For boys, overt aggression was more stable than relational aggression, although the stabilities of relational aggression were still high. For girls, the pattern was more mixed, with some stabilities higher for relational aggression and some higher for overt aggression. Whenever there were reliable gender differences in stabilities, overt aggression was more stable for boys, whereas relational aggression was more stable for girls. In a number of cases, however, both aggression constructs were equally stable by gender. Although correlations are about the stability of individual differences and not about the means, these mixed findings suggest that we should be careful not to portray overt aggression as uniquely male or relational aggression as uniquely female.

Sociometric and perceived popularity were initially positively correlated, but their association declined across the adolescent years. For boys, the correlation between both constructs remained positive throughout the course of the study. For girls, the correlation crossed over from positive in the early years to negative in the high school grades. It should be recognized, however, that there correlations were computed within a sample of students who were initially relatively popular. Yet, the pattern of these changing correlations is consistent with findings for the larger normative samples in each grade (Cillessen & Mayeux, 2004a). Thus, as before, these findings indicate that boys are able to maintain some compatibility between both popularity roles throughout adolescence, whereas for girls, being liked and being seen as popular are increasingly incompatible roles.

The results from our sequence of model tests are best summarized by the final models (Models E) and the prototypical plots that present their results. Sociometric popularity showed a linear change trajectory; perceived popularity showed nonlinear change. Overt aggression predicted low sociometric popularity but an increase in perceived popularity in the second half of the study. Relational aggression predicted a decrease in sociometric popularity, especially for girls, and continued high-perceived popularity for both genders. The effect of relational aggression on perceived popularity was the strongest in the two school years before and after the transition from middle school to high school.

The final models (Models E) confirm the distinction between sociometric and perceived popularity made in other research (see e.g. Cillessen & Rose, 2005). Our final models show differences in the trajectories and predictors of both constructs within a sample of initially popular students. For sociometric popularity, the rate of change was linear, overt aggression predicted lower acceptance (especially for girls), and the effects of relational aggression decreased over time. For perceived popularity, the rate of change was quadratic, overt aggression predicted higher perceived popularity (although less strongly for girls), and the effects of overt and relational aggression had opposite effects on the nonlinear trajectory.

All variance components (random effects) in our final models of both types of popularity were still different from zero. The within- and between-person predictors (both types of aggression and gender) did substantially reduce unexplained variance and improve model fit in the preceding models. At the same time, the results from the final models suggested that additional predictors at both Level 1 and Level 2 may be needed to further exhaust the residual variances. Decisions
regarding which variables to add cannot be derived from the statistical models, but need to come from theory and substantive considerations. Examples of possible additional between-person predictors are ethnicity and socioeconomic status, especially in school systems that are diverse in these respects. Physical attractiveness and athletic ability may prove to be important additional within-person predictors. Addition of these variables is an important substantive direction for future research.

An additional important direction for future research is methodological rather than substantive. As explained in the introduction, sociometric data are typically standardized by school or grade to control for differences in grade size and the fact that these data are nested within the reference group in which they are collected. There are two important issues here. First, sociometric data are typically analysed as if they are independent from the context in which they are collected. Standardization achieves this independence partly, but incompletely, as there may still be differences in group norms or the distribution of sociometric scores within classrooms or grades. Second, when sociometric data are standardized to a constant mean over time, growth curve analyses for the entire reference group are typically not possible. Solutions to both issues for sociometric data need to be developed and implemented in future research: how to properly deal with the nested data structure, and how to develop new scores that do not lead to constant means over time. Both issues form an important agenda for future sociometric research.

The study of adolescent peer relations in general, and of sociometric and perceived popularity specifically, is an active area of current research in developmental psychology. Future studies are expected to shed further light on the questions that were not answered in this study. The current paper demonstrates how the multilevel model for change and other approaches to the modelling of longitudinal data more generally can be instrumental in this research.

Acknowledgements

This research was supported by a Grant from the University of Connecticut Research Foundation to the first author. The authors are grateful to the students, teachers, and school administrators who participated in this study. Portions of this research were presented at the annual meeting of the American Sociological Association, Philadelphia, PA, August 2005.

References


