Child Development, July/August 1999, Volume 70, Number 4, Pages 930–943

The Structure of Achievement and Behavior across Middle Childhood

Lori Kowaleski-Jones and Greg J. Duncan

This research used data from the National Longitudinal Survey of Youth (NLSY) to describe and model developmental trajectories across middle childhood. Our sample consisted of approximately 1,000 children of NLSY women who were aged 6 to 7 years in either 1986 or 1988. Assessments of PIAT math and reading scores and the mother-reported Behavior Problem Index in 1986, 1988, 1990, and 1992 provided data for middle-child trajectories of children aged 6 to 7 in 1986. Assessments in 1988, 1990, 1992, and 1994 provided data for children aged 6 to 7 in 1988. We used the raw score form of these data to estimate LISREL-based models of their autoregressive structure. As with other samples, average math and reading achievement trajectories were parabolic for NLSY children, with scores increasing at a decreasing rate over this period. Average behavior-problem trajectories were flat. Behind these average shapes was extreme diversity in level (and in some cases, slopes), of individual trajectories and a pronounced tendency for above average changes between two adjacent assessments to be followed by opposite-signed changes in the subsequent period. Estimates from our structural models showed great heterogeneity in the average level of achievement and behavior for all three outcomes and heterogeneous slopes for reading scores as well. Boys but not girls were found to have heterogeneous slopes for math and behavior problems, whereas girls but not boys showed a significantly higher degree of persistence if "shocked" off of their expected trajectories.

INTRODUCTION

Understanding how the abilities, achievements, and behaviors of children evolve across childhood is a key concern of developmentalists. Recent advances in both methodology (e.g., Willett & Sayer, 1994) and data collection (Chase-Lansdale, Mott, & Brooks-Gunn, 1991) now make possible the description and modeling of achievement and behavior trajectories for large national samples of children spanning nearly a decade. This article uses data from the National Longitudinal Survey of Youth (NLSY79; Baker, Keck, Mott, & Quinlan, 1993) to describe and model developmental trajectories across middle childhood.

We begin with a simple description of the behavior and achievement patterns between ages 6 and 7 years, and 12 and 13 years. Consistent with countless studies based on more specialized samples, we found that, on average, reading and mathematics achievement increased at a decreasing rate, whereas averages of mother-reported behavior problems changed relatively little. These averages, however, concealed a great deal of individual variability in test-score and behavior trajectories.

Crucial in this variability was a pervasive tendency for sharp changes in scores from one assessment period to the next (intervals of 2 years in our data) to be followed by pronounced changes in the opposite direction over the following measurement interval. This regression to the mean was consistent with a measurement-error interpretation, although it also was consistent with a process involving genuine change, caused, say, by a family disruption (if negative) or an unusually good teacher or mentor (if positive) that was not sustained.

Unlike previous work on growth-curve dynamics (e.g., Willett & Sayer, 1994), we viewed the speed with which children with sharp changes in scores revert to their "expected" trajectories as a key parameter of interest. This parameter may reflect institutional factors such as school-based tracking or the cumulative nature of learning of a subject like mathematics, both of which might convert an otherwise transitory downward "shock" to achievement into a permanent downward displacement in the trajectory.

That children differ in average levels of achievement and behavior is beyond doubt. The extent of variability in the slopes of developmental trajectories is less certain, especially if one allows for transitory shocks to children's observed scores, coupled with some degree of persistence in the deviation of "shocked" scores from longer-run trajectories. Our approach allowed us to test explicitly for the degree of variability in trajectories, once mean differences and the degree and persistence of shocks were taken into account.

We begin by summarizing some of the key literature on developmental trajectories in achievement and behavior for the middle-childhood period. We then provide a description of our data, followed by a description of various empirical features of the trajectories of these scores across middle childhood. We then de-

© 1999 by the Society for Research in Child Development, Inc.
All rights reserved. 0009-3920/99/7004-0010

scribe and estimate our models of achievement and behavior-score dynamics and discuss our results.

Background

Given its continuous nature, dividing human development into discrete time periods is somewhat arbitrary. Yet, the period of middle childhood has been distinguished from early childhood and adolescence along a number of key dimensions (e.g., Collins, 1984). Around the age of 6 or 7, children show skills and characteristic modes of thought and behavior that differ significantly from previous periods. Children become involved in the process of consolidating, extending, and integrating social and personal knowledge acquired previously. These changes manifest themselves in the increased capacity to solve problems and the beginning capacity to form intimate relationships. Second, and even more important, middle childhood marks the emergence of heterogeneous capacities and typical behaviors. An important research question is to what extent and in what way children's basic competencies and behavioral patterns become "crystallized" during middle childhood into relatively consistent forms that persist into adolescence and early adulthood (Collins, 1984; Caspi, Elder, & Bem, 1987). Moreover, is the extent of crystallization greater for girls than boys? Our research addressed this question and suggested the "crystallization" process during middle childhood might be sharper for girls than boys.

We recognize that middle childhood is a period of change in many dimensions such as emotional identities, physical maturation, and peer attachment. We focused our empirical research and our literature review on two key domains of the middle childhood period, however: achievement and problem behavior.

Achievement. There is a large psychological literature on the nature of children's cognitive development during the middle childhood ages (Honzik, MacFarlane, & Allen, 1948; McCall, Appelbaum, & Hogarty, 1973; Piaget, 1971). Though our primary interest was in children's achievement rather than ability, the literature on cognitive development was of obvious importance for understanding achievement trajectories. Collins (1984) observed that ages 5 to 7 are recognized to be a period of significant cognitive changes, as in the increasing ability to deal systematically with abstract representations of objects and ideas. Between the ages of 6 and 9, children gain skills that allow them to reason about increasingly complex problems and situations in both physical and social worlds. Later in the period, between the ages of 10 and 12, children begin to develop increasing facility in generalizing across discrete instances and are more able to reason by generating and testing hypotheses. More generally, during this period children are increasingly able to organize their thoughts and behavior. The ability to monitor one's own activities and mental processes increases during middle childhood (Brown, Bransford, Ferrara, & Campione, 1983).

Children become more adept during middle childhood at acquiring new information and using this information in reasoning and action. It also marks the entrance of children into the formal education system. Given this, much of the new information they are likely to acquire will be in an academic context. Though school may be the overt catalyst to acquire academic knowledge, schools also function to socialize children in terms of dominant cultural norms and values, including attitudes about gender expectations.

According to Erikson (1963), middle childhood is a period of industry in which children master the basic tools and skills needed for adult life. Children are consumed with the task of becoming producers; they learn skills; they begin to learn what constitutes success in the larger society; to the extent that they feel that their family characteristics, race, or gender limits their potential for success, their competence might suffer. Indeed, one of the potential dangers of middle childhood is that children's sense of personal adequacy is developing and therefore vulnerable. To the extent that they feel their achievement skills are inadequate, they might be discouraged from further developing these skills. Moreover, if events divert them from their expected trajectories, otherwise temporary diversions may become permanent.

Using data from the Fels Longitudinal Study, McCall and colleagues (1973) observed that the median range of standardized IQ scores of middle-class children is 29 IQ points between the ages of 2 and 17. They also found that 1 in 7 children displayed shifts of more than 40 points. Cluster analysis techniques revealed multiple groups of children who exhibited different pattern of IQ change over age. Their work revealed trajectories for some children in which IQ increased quite sharply from the age of 6 to peak at the age of 9, only to then decrease to a level by the age of 17 that was equivalent to the level at the age of 6. This significant variability in children's mental performance trajectories suggested that achievement trajectories also might be characterized by high degree of variability. Our work sought to chart this potential variability in children's math and reading achievement.

Research focusing more generally on the dynamic trajectories of reading and math achievement is limited. Pungello, Kupersmidt, Burchinal, and Patterson (1996) used hierarchical linear modeling (HLM) to ex-

amine the role of environmental risk in children's achievement from middle childhood to adolescence. They found that lower family income and minority statuses were significant predictors of children's achievement over time. Theoretical work by Eccles and colleagues (1983) suggested that trajectories of achievement may differ for boys and girls (Eccles, 1984; Eccles, Adler, & Meece, 1984; Eccles, Jacobs, & Harold, 1990). Their model linked achievement choices to expectancies for success to the importance of the task or incentive value of the task, in this case achievement in math or reading. Success in an achievement-related task was shown to be a function of expectations of success and perceived value of the task, as well as factors present within social realm such as gender role and cultural stereotypes (Eccles, 1984). This model assumed that achievement decisions, such as the decision to strive for high attainment on achievement measures, are made in the context of choices. Additionally, it assumed that these choices are guided by one's expectations about various kinds of success, competency needs, and social expectations about gender stereotypes and by the potential cost of investing in one activity rather than another. If, as Erikson (1963) suggests, middle childhood is a time of skill attainment and an increasing awareness that access to success might be limited based on personal characteristics, then girls might be more likely to be diverted from their expected path. Specifically, a transitory shock to an expected trajectory may be more likely to become permanent for girls than boys.

Problem behavior. Studies of children's behavior problems suggest that, for most children, problems observed in early childhood are not enduring, but that behavior problems stabilize at different levels for different children (Achenbach, 1984). Dodge (1993) reviewed the social cognitive mechanisms involved in the development of aggressive and depressive problems and argued that aggressive behavior derives from basic cognitive deficiencies that make it more difficult for some children to find nonaggressive solutions to interpersonal dilemmas. Aggressive children frequently are not able to initiate varied solutions to problems and are more likely to misperceive others' intentions. Dodge (1991) also argued that aggressive children are more likely to perceive peers as having aggressive intent and to act accordingly, and therefore create a situation in which the interaction will become problematic. Children who exhibit depressive behavioral affect are more likely to misperceive daily situations as well. These children are more likely to selectively attend to failure and negative information and tend to blame problems on their own shortcomings.

These reactions to emotional distress may either be temporary or indicative of a long-term self-reinforcing pattern of learned behavior. Established patterns of family interaction patterns may provide repeated reinforcement for problematic behavior and increase their chance of reoccurrence. If persistent, the behavior may lead children to select themselves into environments that further sustain these behaviors (Elder & Caspi, 1988).

Similarly, Crick and Dodge (1994) contended that developmental pathways are based heavily on repeated experiences. They described a process by which the children's paths are initially quite varied but become more deterministic. In essence, they suggested that children are less likely to be "shocked" away from their path, as their path becomes more "worn" by experience.

Wachs (1996) also discussed processes underlying developmental trajectories in childhood. He identified a developmental process by which children choose "niches" that serve to maintain continuity in individual development over time. These niches are locations within the child's social environment and can have a positive or negative valence, depending on whether the niches inhibit or enhance individual development. Some scholars (Scarr & McCartney, 1983) have contended that the choice of niches is the result of rational choice, but still others, such as Wachs, have argued that there are fewer available niches and a higher proportion of negative valence niches should hinder behavioral adjustment over time. Examples include a variety of biological (chronic illnesses), environmental (substandard family interaction), or individual characteristics (innate cognitive abilities and temperament) that may enhance or impair an individual's capacity to discover positive developmental niches. These children's paths may be complex and marked by differential tendencies for children shocked off one trajectory to be propelled into another "niche" that can persist over time.

Much of the available research on childhood behavioral problems has either focused on the determinants of their level at a point in time without (Cooksey, Menaghan, & Jekielek, 1997; Mott, Kowaleski-Jones, & Menaghan, 1997) and with (Menaghan, Kowaleski-Jones, & Mott, 1997) controlling for behavior problems in previous years. Studies of change in childhood aggression typically have been based on data with relatively few data points (Patterson, 1993).

Loeber and his colleagues (1993) found support for three developmental pathways involving severe aggression patterns during childhood and into adolescence. The first consisted of children first engaging in aggressive acts such as annoying and bullying others, then moving to physical fighting, and then to serious

violence in adolescence. The second consisted of escalation in covert and concealing problem behaviors, and the third pathway was an authority conflict pathway that concerned conflict with parents and teachers.

A few studies have used growth curve methods to examine problematic behavioral trajectories throughout the period of middle childhood. One charted the impact of poverty histories on childhood antisocial behavior of children in the NLSY79 (McLeod & Shanahan, 1996). It found that the number of years that children were poor between 1986 and 1990 correlates significantly with changes in children's antisocial behavior during those years and that antisocial behavior increased more rapidly for children with histories of persistent poverty during those years than for transiently poor or nonpoor children. These results demonstrated the accelerating behavioral disadvantages faced by persistently poor children. Growth curve models also have been used to chart the long-term influence of parental divorce on children as they move into adulthood; evidence from a British birth cohort indicated that a parental divorce in childhood continues to affect individual behavioral problems in adulthood (Cherlin, Chase-Lansdale, & McRae, 1998).

As with achievement, we expected that boys and girls may differ in the structure of their behavior trajectories. In a review of the development of aggression, Loeber and Hay (1997) documented gender differences across middle childhood. Boys are persistently more likely to show higher rates of physical aggression than girls. In contrast, girls may be more likely to begin to exhibit more verbal and indirect aggression such as alienation and character defamation.

If girls were to behave in a problematic manner, however, this behavior might receive more attention from key adults and peers who may place more relative importance on female aggressive behavior. Research has indicated that children who develop early behavior problems may elicit reactions from parents, teachers, and others that reinforce those problem behaviors (Belsky, 1984; Patterson, 1986). This process might be even sharper among girls than boys if adults perceive their problem behavior as nonnormative.

McFadyen-Ketchum, Bates, Dodge, and Pettit (1996) used four waves of data to identify early child aggression trajectories. These trajectories were based on assessing whether the child scored high or low on aggression and whether child aggression levels subsequently decreased or increased. The evidence from this research indicated that mother–child interaction patterns were predictive of initial levels of aggression, but boys and girls differed in how coercion and nonaffection predicted changes in aggression across the elementary school years. Boys who lived in coercive and nonaffectionate family environments were more likely to follow a path of high initial aggression and subsequent increasing aggression. In contrast, girls who lived in similar family environments were more likely to follow a behavioral trajectory characterized by high initial aggression and subsequent decrements in aggressiveness. These results suggest that there might be important gender differences in child developmental behavioral trajectories.

As with the literature on achievement, much of the available research has been concerned with identifying and testing predictors of childhood behavior trajectories. The important goal of gaining an accurate dynamic picture of the structure of behavioral pathways remains. These studies point to the value of using longitudinal methods to characterize the potential childhood behavior trajectories, but much work needs to be done to describe the basic shape and variability of behavioral trajectories in middle childhood.

METHOD

Data

We used data from the 1994 and earlier survey rounds of the National Longitudinal Survey of Youth (NLSY79; Baker et al., 1993), a nationally representative study designed by the United States Department of Labor to study variations in labor market behavior and experiences. The mothers of the children we studied were between the ages of 14 and 22 when first interviewed in 1979 and constituted a representative sample of individuals born between 1957 and 1965. Their ages ranged from 29 to 37 in 1994. The NLSY79's initial design oversampled Black, Hispanic, and economically disadvantaged White youth. Design-based weights were used throughout our analysis to produce nationally representative estimates. These youth (and the subset of them who were or became mothers of the children we studied) have been interviewed annually since 1979. The response rate in the 1979 wave was 79%. This survey has had a remarkably high retention rate, with almost 90% of those originally interviewed in 1979 also interviewed in 1994.

Beginning in 1986, interviewers administered an extensive set of assessment instruments to the children of all the female respondents. These assessments included information about cognitive, socioemotional, and psychological aspects of the child's development as well as about the quality of the home environment (Baker et al., 1993). In 1986, the assessments were completed by 4,971, or about 95%, of the coresident children of mothers interviewed in 1986. These same children, plus newly age-eligible

children, were interviewed again in 1988 (6,266 valid interviews were completed), in 1990 (5,803 valid interviews), in 1992 (6,509 valid interviews), and in 1994 (7,089 valid interviews).

We studied two cohorts of children. The first consisted of children aged 6 to 7 in 1986, 8 to 9 in 1988, 10 to 11 in 1990, and 12 to 13 in 1992. There were 633 children in this cohort from whom interview data was gathered in the three follow-ups. These children were among the oldest children of the NLSY79 mothers. Some 51% of them were born to mothers who were under the age of 20 at the birth of the child, and none were born to mothers who were older than the age of 23 at the time of the birth. The second cohort of children we selected, of which there were 903, were aged 6 to 7 in 1988, 8 to 9 in 1990, 10 to 11 in 1992, and 12 to 13 in 1994. Though more representative than children in the first cohort, almost one third of the children in this cohort were born to mothers under the age of 20 at the time of the birth, and none were born to mothers older than the age of 25 at the time of the birth. When interpreting the results of our research, it is important to keep this sample limitation in mind. Additionally, we excluded children whose mothers were part of the initial oversampling of economically disadvantaged poor White sampling in 1979. For financial reasons, these mothers and their children were dropped from the NLSY79 between 1988 and 1990 surveys. We also limited our multivariate analyses to those children who had valid scores for all four time periods for the given assessment.

Measures

We used two achievement measures and one behavior problem assessment from these data (Baker et al., 1993). Both achievement measures are taken from the Peabody Individual Achievement Tests (PIAT): mathematics and reading recognition (Dunn and Markwardt, 1970). Children aged 5 and older were administered these tests at each of the biennial assessments.

The mathematics assessment consists of 84 multiple-choice items of increasing difficulty. It begins with such early skills as recognizing numerals and progresses to advanced concepts in geometry and trigonometry. A child looks at each problem and then chooses an answer by pointing to or naming one of four options. Children enter the assessment at an age-appropriate item and establish a basal by attaining five consecutive correct responses. Testing stops when the child responds incorrectly to 5 out of 7 items. Our analysis is of the raw score responses to these questions.

The PIAT reading recognition assessment measures word recognition and pronunciation ability. This assessment contains 84 items, each with four options, that increase in difficulty from preschool to high school levels. Skills assessed include matching letters, naming names, and reading single words aloud. The scoring and scaling procedure is similar to that of the PIAT math assessment. Both the PIAT mathematics and reading recognition assessments are generally considered to be highly reliable and valid, and their widespread use is evident in the psychological and sociological literature (Murphy, Conoley, & Impara, 1994).

We measured the children's behavior problems with items that were drawn largely from the Achenbach and Edlebrock Behavior Problems Inventory (BPI; Achenbach & Edlebrock, 1981). We combined the externalizing and internalizing subscales into a single score. Examples of internalizing behavior were whether the child was: unhappy, sad, or depressed; complained no one loves him/her; and demanded a lot of attention. Examples of externalizing behavior were whether the child was: disobedient at home; was stubborn, sullen, or irritable; and had a strong and easily lost temper. Scores on these items were 1 if the mother reports "often true" or "sometimes true," and 0 otherwise.

Researchers who have demonstrated the considerable construct validity of the measure have extensively used this index of children's behavioral problems. For example, Parcel and Menaghan (1988), among others, demonstrated strong and plausible relations between the BPI and a variety of social and demographic variables. This measure of behavior problems had consistently high test–retest reliability. On average, Cronbach's α for the measure across the assessment years is .85.

RESULTS

Patterns of Achievement and Behavior Trajectories

Figures 1 to 3 present the means and standard deviations of the achievement and behavior scores at the four age points provided by our data. Although these data were calculated from a longitudinal sample, the means and standard deviations plotted in the figures were cross-sectional in nature and could have come from independent samples taken at the four age points. As expected, there was curvilinear growth in mean reading and, especially, mathematics scores, but little net change across middle childhood in mean behavior-problem scores. There was some tendency for the dispersion in reading scores to increase over the 8-year interval for both boys and girls and for the dispersion in behavior problems to increase for girls.

To gauge the degree of dispersion in linear trajec-

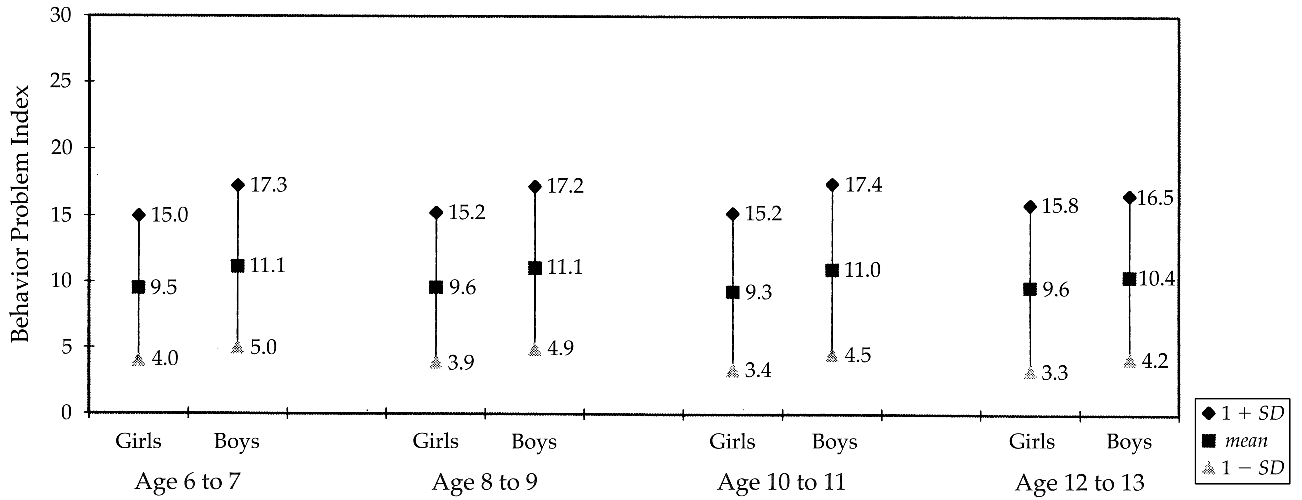

Figure 1 Means and standard deviation ranges of behavior problems for boys and girls aged 6 to 7 in 1986 or 1988.

tories, we calculated for each child and each assessment an average change in scores. We did this by fitting a time-centered OLS regression line to the four observed scores. For a child with scores of y_1, y_2, y_3, and y_4, this amounts to calculating $(-1.5y_1 - .5y_2 + .5y_3 + 1.5y_4)/5$. The slope value can be interpreted as the average amount of test-score change over each 2-year interval. Ranking children by average change revealed considerable dispersion. Consistent with Figure 1, the median change in behavior problems was essentially zero. But the 10% of the sample with the most negative change in BPI scores enjoyed declines that totaled at least 7 points on the BPI scale over the 8 years. The 10% of the sample with the most problematic positive BPI trajectories experienced total increases of at least 7 points on the BPI scale over the 8 years.

In the case of the two achievement scores, the 10% of the children with the worst (i.e., least positive) trajectories still had positive changes: 6.0 to 6.4 points every 2 years in mathematics scores, and 5.4 to 6.3 points every 2 years in the case of reading. The changes for the 10% of children with the most positive trajectories were, of course, much larger: 13.5 to 13.7 points every 2 years in mathematics scores, and 15.5 to 16.2 points every 2 years in the case of reading.

Linear trajectories such as these describe an average rate of change across the measurement interval and can lead to the impression that most children display monotonic patterns of increasing or decreasing scores. In fact, the observed score trajectories were rarely monotonic, with increases much more likely to be followed by decreases than further increases. By taking the simple difference between time 1 and time 2 scores and correlating them with the simple differ-

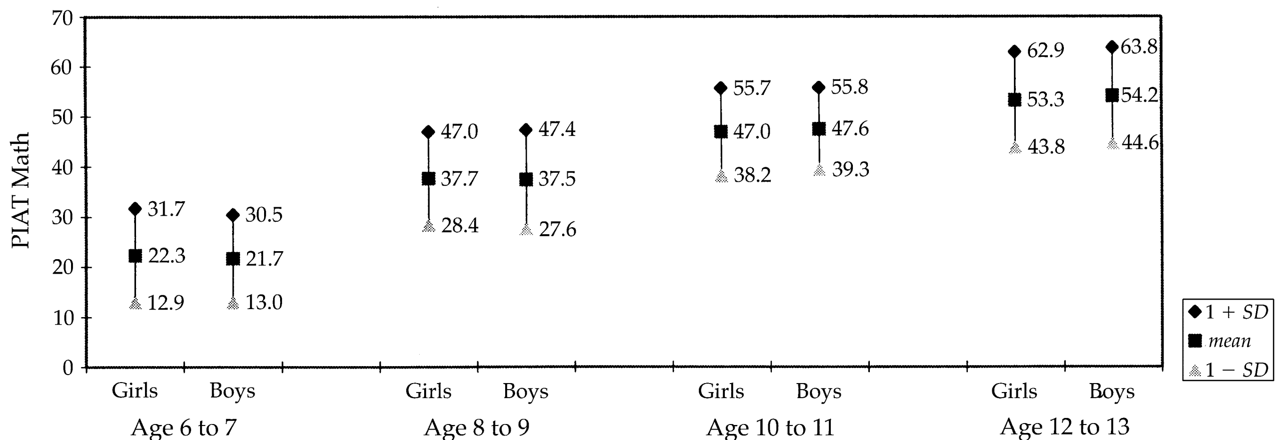

Figure 2 Means and standard deviations of math scores for boys and girls aged 6 to 7 in 1986 or 1988.

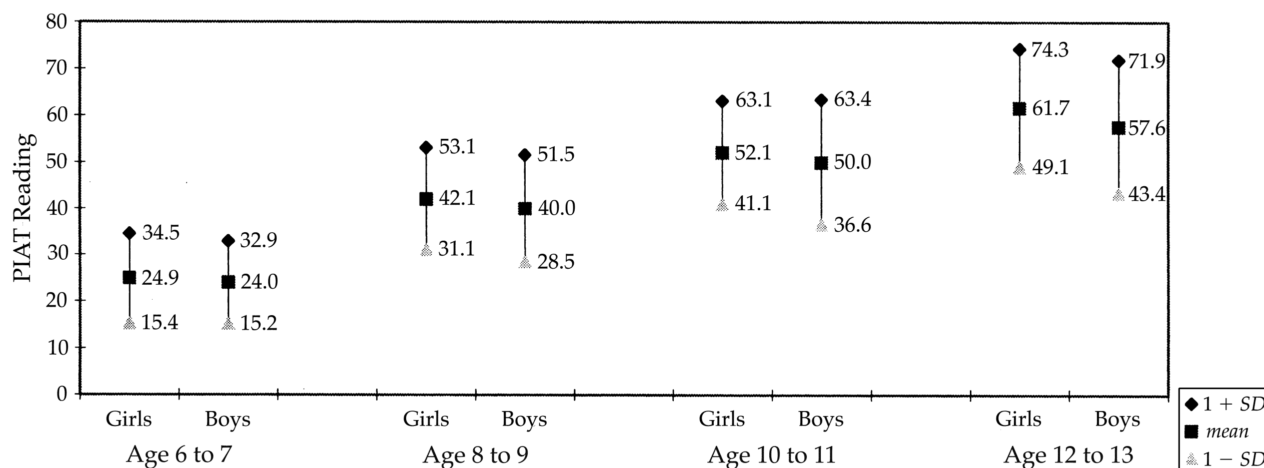

Figure 3 **Means and standard deviation ranges of reading scores for boys and girls aged 6 to 7 in 1986 or 1988.**

ence between scores at times 2 and 3, we found that the correlations were negative and substantial in all cases—ranging from −.24 in the case of change in reading scores, to −.31 for behavior problems, and −.37 for math scores. Analogous correlations for changes between times 2 and 3 and times 3 and 4 were similarly negative, ranging from −.34 for reading change, to −.35 for math-score change, and −.39 for behavior-problem change.

Between one fifth and one third of the children in our sample had achievement or behavior scores that were consistently above or below the average of their age mates (Table 1). But what about the persistence of trajectories? We calculated the proportion of children who were consistently above and below average on a simple difference change score between times 1 and 2, times 2 and 3, and times 3 and 4. These calculations are shown in Table 1. Looking across all three assessments, we saw that slightly higher proportions of children were above the mean change on the behavior problems score, .11, than for the math and reading scores, .09 and .06, respectively. Very few of the children were consistently below the average on the behavior problem change scores, .02, whereas a slightly higher proportion of children were persistently below the mean for changes in math and reading scores, .06 and .10, respectively. All in all, these proportions indicated that fewer than 1 in 6 children had achievement or behavior trajectories that were either persistently more negative or positive than their age mates.

Negative correlations were consistent with a measurement error interpretation, because an erroneously high score in, say, period 2 will produce an erroneously positive change between time 1 and 2, followed by an erroneously negative change between times 2 and 3. But they also were consistent with a more substantial picture of real change caused by some positive or negative event, followed by a resilient recovery in the case of an initially negative change or disappointing reversion to a longer run, less positive, trajectory in the case of an initially positive change.

Table 1 **Weighted Statistics for Behavior Problems, PIAT Math, and PIAT Reading Scores for All NLSY Children aged 6 to 7 in 1986 or 1988**

	BPI	Math	Reading
Level scores			
Above average scores at all four assessments	.22	.22	.26
Below average scores at all four assessments	.28	.26	.34
Change scores			
Above average on change scores 1, 2, and 3	.11	.09	.06
Above average change 1 and 2, below average change 3	.14	.11	.11
Above average change 1, below average change 2 and 3	.11	.14	.11
Below average change 1, above average change 2 and 3	.13	.11	.10
Below average change 1 and 2, above average change 3	.12	.11	.09
Below average on change scores 1, 2, and 3	.02	.06	.10
Above average change 1, below average change 2, and above average change 3	.19	.19	.16
Below average change 1, above average change 2, and below average change 3	.18	.19	.17
N =	966	1,016	1,020

Our approach to describing these test and behavior-score dynamics fit models to the variance–covariance matrices of the scores. These matrices, along with the associated correlations, are presented in Table 2. Based as they are on individual-level variability across time, they revealed several features of the data not readily apparent in Figures 1 to 3. For behavior problems, the variance of girls' scores at time 1 (aged 6 to 7) was considerably smaller than that of boys. The variance of boys' and girls' math scores fluctuated considerably over time. Noteworthy in the reading scores was the increasing variance for girls and, especially, for boys.

Researchers have argued that middle childhood is a period in which behavior problems are likely to stabilize (Achenbach, 1984). Table 2 presents the correlations of behavior scores from time 1 to time 4 that have bearing on this issue. The correlation between boys' time 1 and time 2 behavior problems was .62 (.60 for girls), which would suggest a fair amount of stability between the two time periods. The correlations between the adjacent time periods (2 versus 3, and 3 versus 4) over the rest of the study period, however, only increased slightly and actually decreased between time 3 and 4 among girls. This does not suggest a clear case for the increasing stability of behavior problems over middle childhood. In fact, these descriptive statistics point to the importance of using more sophisticated multivariate methods for the study of developmental trajectory variability over time.

A Statistical Model of Achievement and Behavior Dynamics

Drawing from the literature on income dynamics (Hause, 1980; Lillard & Willis, 1978), we estimated a model of achievement and behavior-score dynamics that allowed for persistent differences in the level and slope of individual trajectories. This model also allowed for "shocks" that throw children off their expected trajectories and for a recovery process from the shocks that brings people back to these trajectories. The model can be written as:

$$y_{it} = \mu_i + \gamma_i t + e_{it} \quad (1)$$
$$\text{where} \quad e_{it} = \rho e_{it-1} + v_{it}.$$

In (1), y_{it} is the achievement or behavior-problem score of child i in year t, expressed as a deviation from the time t sample mean; μ_i is the ith child's unobserved time-invariant achievement or behavior score; γ_i is the linear slope of the ith child's scores across time; e_{it} is a first-order autoregressive error term; ρ is a serial correlation coefficient common to all children; and v_{it} is a purely random error term. In our data, t increased by one unit between the every 2-year assess-

Table 2 Covariance-Correlation Matrices

Behavior Problem Index									
Boys	Time 1	Time 2	Time 3	Time 4	Girls	Time 1	Time 2	Time 3	Time 4
Time 1	37.62	.62	.53	.38	Time 1	29.88	.60	.51	.46
Time 2	23.44	37.65	.63	.55	Time 2	18.56	32.04	.74	.61
Time 3	21.02	25.16	41.90	.65	Time 3	16.52	24.72	38.89	.63
Time 4	14.22	20.86	25.76	38.01	Time 4	15.52	21.38	23.33	38.89
PIAT Math									
Boys	Time 1	Time 2	Time 3	Time 4	Girls	Time 1	Time 2	Time 3	Time 4
Time 1	76.59	.64	.48	.43	Time 1	88.67	.66	.55	.53
Time 2	55.62	98.50	.61	.58	Time 2	57.47	86.74	.70	.64
Time 3	34.19	48.63	67.61	.69	Time 3	45.51	57.05	76.22	.70
Time 4	36.28	55.57	54.40	92.44	Time 4	47.98	56.64	58.67	91.34
PIAT Reading									
Boys	Time 1	Time 2	Time 3	Time 4	Girls	Time 1	Time 2	Time 3	Time 4
Time 1	78.86	.69	.71	.54	Time 1	91.72	.68	.60	.51
Time 2	70.20	133.02	.86	.84	Time 2	71.53	120.72	.84	.74
Time 3	84.98	133.43	179.56	.84	Time 3	63.62	101.15	121.13	.80
Time 4	68.61	137.23	160.54	202.79	Time 4	61.51	102.90	110.65	158.94

Note: Covariances are on and below the diagonal. Correlations are above the diagonal. All are weighted.

ments and ranged from −1.5 to +1.5 to center it within our four assessment points.

In effect, the model presumed that each child's score at a given point differed from his or her age-mates in terms of both average level (μ_i) and slope (γ_i). Economists who have developed these models term the μ_i a "fixed effect," because, in our context, it is an unchanging characteristic of each child, relative to his or her age-mates, across time. The e_{it} term is a residual from this average level and slope that reflects both purely random variation (v_{it}, termed "shock variance" in the income models) as well as some tendency (captured by ρ) to remain above or below the expected trajectory once pushed ("shocked") away from it.

Suppose, for example, that a divorce disrupts a child's life and increases his or her behavior-problem symptoms. Whether the child persists with more behavior problems or quickly returns to his or her "normal" behavior problem trajectory is reflected in a high or near-zero value of ρ.

The model produces a number of parameters of interest. In large part, the observed sample variance of a score is the sum of the fixed effect (σ^2_μ) and transitory (σ^2_v) variances. (Heterogeneous slopes [σ^2_γ] also will contribute to the observed variance, but we ignore that fact for the moment.) Thus, the relative sizes of fixed effect (σ^2_μ) and transitory (σ^2_v) variances are a useful indicator of how much of the observed variability in a score results from permanent as opposed to transitory interindividual differences. Whether once transitory shocks (v_{it}) and recovery (ρ) are taken into account, there are significant differences in the slopes of children's developmental trajectories is revealed by the size and significance of the slope variance (σ^2_γ). The recovery process itself is summarized by ρ, which can vary between −1 and +1 but is expected to be either zero (indicating complete recovery over the 2-year interval between assessments) or positive (indicating partial recovery).

Although transitory variances and the ρ recovery parameters could, in theory, vary across the four assessment points, the four assessments did not provide enough degrees of freedom to estimate such a complete model. After considerable experimentation, we adopted the assumption that the three ρ parameters had a common value, as did the second through fourth shock variances (σ^2_v). Our experimentation included constraining various combinations of period-specific transitory (σ^2_v) and ρ parameters and assessing whether the model fit improved relative to the model outlined above. The rationale for allowing for a different shock variance in the first period is that first-period variance presumably reflects dynamic elements that preceded the point of first measurement. Given the increase in the observed variance of reading scores, we anticipated that the transitory variance might be higher in the fourth and possibly third period than in earlier periods. Given the Crick and Dodge (1994) hypothesis of increasingly deterministic behavior problems trajectories, we anticipated small "shock" variances in the third and fourth periods. In all cases, none of this experimentation significantly improved the fit of the model.

Insight into the structure of the model can be gleaned from the four hypothetical expected and observed test scores for three children drawn in Figure 4. The expected trajectories of the first and second children are parallel to one another, reflecting an assumption of fixed effect (μ_i), but not slope (γ_i) differences between the first and second child. The third child's trajectory is drawn to have both a lower level and more negative slope than those of the first two children.

Measurement error and unexpected events will cause observed scores to deviate from these expected trajectories. Observed scores are drawn in Figure 4 to be fairly close to the expected trajectories, reflecting an assumption that transitory variance (σ^2_v) is small relative to the variance of expected interindividual (fixed effect, σ^2_μ) differences. The models we estimate provided a quantitative assessment of the relative sizes of these two components of variance.

Observed scores in Figure 4 have also been drawn to reflect varying assumptions about the speed of adjustment to the transitory "shocks." In the case of the first child, the score at age 6 was somewhat higher than his or her expected trajectory, reflecting some combination of transitory positive shocks around age 6 and the lingering effects of positive shocks at prior unobserved ages. The first child experienced a substantial negative shock between ages 6 and 8, causing the observed score at age 8 to lie below his or her expected trajectory. (Note that even though the shock was negative, the observed score does increase between ages 6 and 8. The change is negative in the sense that it is less than what would be predicted by the expected trajectory.) Assuming no further shocks, the observed scores at age 10 and 12 showed a pattern of only gradual return to the expected trajectory. In the language of the model, ρ was positive and close to unity for the first child, reflecting initially no recovery from "shocks." Child 2's observed scores were drawn to reflect a more rapid return to the expected trajectory, that is, a less positive ρ. Observed scores for the third child followed an autoregressive pattern similar to that of the first child: a negative shock between the ages of 6 and 8, followed by very slow recovery. Our model provided estimates of ρ that were averaged across all children.

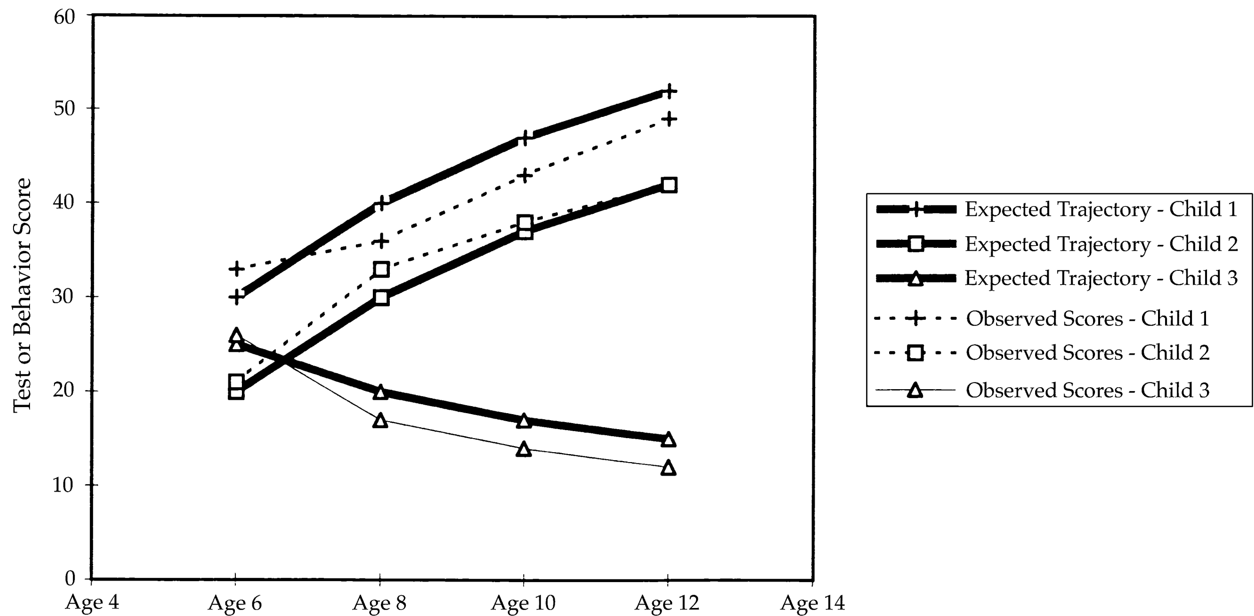

Figure 4 Expected and observed test or behavior score trajectories.

Multivariate Results

Parameter estimates and standard errors from a variety of models are shown in Table 3 for girls and Table 4 for boys. Despite the many similarities between the test and behavior score patterns for boys and girls revealed in our descriptive analysis, tests of equality of parameters in our test score dynamics models always indicated highly significant differences between the two groups. As a result, we have presented all of our results separately by sex and then noted which of the individual parameters differed significantly between the sexes.

We show results from two nested models, the first of which constrained the variance of the slope trajectories (γ_i) to zero and the second relaxing that constraint. The fit of these models was summarized with the model χ^2, which is dependent on sample size, the adjusted goodness of fit index, and the Bayesian Information Criterion (BIC) statistic (Raftery, 1995). Better fitting models were indicated by increasingly negative values of the BIC statistic.

Taking behavior problem scores as an example, we can see that the first model, which allows for heterogeneity in the levels but not slopes of score trajectories, produced only a moderately good fit to the data for both boys and girls; an allocation of observed variance into roughly equal amounts of fixed-effect and transitory variance (comparing σ^2_μ with the second, third, and fourth period values of σ^2_v); and an estimate of substantial and highly significant persistence to deviations from long-run trajectories (ρ).

Allowing for heterogeneous slopes produced a highly significant estimate of slope variance for boys but not girls. The covariance between the level and slope of behavior-problem trajectories was significant for girls but not boys. Allowing for heterogeneous slopes also reduced the estimates of both the transitory variance and ρ. An intuitive explanation of the interrelation between the slope variance, transitory variance, and ρ was that dispersed scores over time can be accounted for in two ways: (1) substantial and continual shocks to essential parallel expected trajectories; and (2) smaller and less persistent shocks to differently sloped expected trajectories. In the case of boys, the second explanation clearly fit the data better than the first. In the case of girls, the difference between these two alternatives was not as clear cut.

All told, the estimates of the level-plus-slope BPI models indicated the following. First, the higher observed variability in behavior problem scores for boys relative to girls appeared to be a function of greater dispersion in the permanent level of boys' behavior; transitory variations in behavior problems were quite similar for boys and girls. Second, boys had heterogeneous behavior-problem slopes but, once "shocked" away from their expected trajectories, quickly returned to them. For girls, trajectories differed in level and not slope but, once shocked away from these ex-

Table 3 Weighted LISREL Maximum Likelihood Estimates for Behavior Problems, PIAT Math, and PIAT Reading Recognition Raw Scores, For all Girls Aged 6 to 7 in 1986 or 1988

	No. of Obs.	Observed Variance (Time 1 to Time 4 Average)	Individual Level Variance (σ^2_μ)	Level-Slope Covariance	Individual Slope Variance (σ^2_γ)	First Period σ^2_v	Second to Fouth Period σ^2_v	ρ_1, ρ_2, and ρ_3	χ^2	*df*	AGFI	BIC
BPI	469	33.92	14.39	—	—	17.16	16.92	.42	43.80	6	.932	6.38
			(1.86)			(1.81)	(.85)	(.05)				
			17.11	1.77	.33	18.23	14.38	.28	31.43	4	.925	6.83
			(2.37)	(.55)	(.65)	(4.14)	(1.39)	(.09)				
Math	503	85.74	48.69	—	—	46.31	32.28	.26	14.64	6	.978	−22.68
			(4.28)			(3.95)	(1.72)	(.05)				
			50.65	1.89	1.19	42.66	29.48	.20	9.08	4	.978	−15.80
			(4.91)	(1.21)	(1.31)	(7.86)	(2.81)	(.08)				
Reading	509	123.13	25.64	—	—	66.33	52.55	.80	84.09	6	.874	46.70
			(14.30)			(14.69)	(1.99)	(.04)				
			82.26	11.91	5.39	40.64	29.93	.24	54.76	4	.880	29.84
			(6.94)	(1.77)	(1.54)	(8.78)	(3.02)	(.09)				

Note: Standard errors are in parentheses. First to fourth period σ^2_v are estimates of transitory variance. ρ_1, ρ_2, and ρ_3 are autoregressive parameter estimates.

Table 4 Weighted LISREL Maximum Likelihood Estimates for Behavior Problems, PIAT Math, and PIAT Reading Recognition Raw Scores For All Boys Aged 6 to 7 in 1986 or 1988

	No. of Obs.	Observed Variance (Time 1 to Time 4 Average)	Individual Level Variance (σ^2_μ)	Level-Slope Covariance	Individual Slope Variance (σ^2_γ)	First Period σ^2_v	Second to Fourth Period σ^2_v	ρ_1, ρ_2, and ρ_3	χ^2	*df*	AGFI	BIC
BPI	497	38.79	16.55	—	—	22.79	19.15	.34	28.41	6	.955	−8.84
			(1.98)			(2.09)	(.98)	(.05)				
			21.84	.23	2.13	13.08	14.41	.08	16.52	4	.960	−8.32
			(1.81)	(.54)	(.48)	(2.18)	(1.21)	(.07)				
Math	510	83.78	36.93			43.84	41.89	.36	65.85	6	.898	28.44
			(4.20)			(4.14)	(2.08)	(.05)				
			47.41	1.22	4.45	26.26	31.61	.10	51.11	4	.878	26.17
			(3.90)	(1.16)	(1.02)	(4.54)	(2.58)	(.07)				
Reading	511	148.56	47.57	—	—	31.32	58.48	.89	213.92	6	.705	176.50
			(17.05)			(16.88)	(2.13)	(.03)				
			117.77	22.12	9.93	21.29	18.99	−.22	124.72	4	.736	99.76
			(7.61)	(2.04)	(.92)	(2.37)	(1.56)	(.07)				

Note: Standard errors are in parentheses. First to fourth period σ^2_v are estimates of transitory variance. ρ_1, ρ_2, and ρ_3 are autoregressive parameter estimates.

pected trajectories, the recovery process was slower for girls than it was for boys.

The picture for mathematics scores was generally similar to that of behavior-problem scores. The level-and-slope models fit the data better than the level-only models; boys had much greater variability in slopes than girls; however, shocks had more persistent effects for girls than boys. Relative to behavior problems, the transitory component to math scores was much less important than the permanent component for both girls and boys.

The model did not fit the reading data as well as the mathematics or behavior-problem data. The general pattern of results was again similar, however: the level-and-slope models fit the data better than the level-only model; boys had much greater variability in slopes than girls; however, shocks were more persistent for girls than boys. In fact, for boys the estimate of ρ was negative and significant, suggesting, somewhat implausibly, that the recovery process overshot the long-run trajectory within the 2-year period.

DISCUSSION

We have attempted to model achievement and behavior dynamics using important national data and growth models that highlighted the characteristics of the transitory component to scores. We found a much larger role for transitory variance in the case of problem-behavior reports than either mathematics or reading achievement, as well as a tendency for "shocks" to math and reading scores to be more persistent for girls than boys.

That the structure of the models differs between boys and girls raises interesting questions. Explanations for gender differences in achievement trajectories can be drawn from the developmental literature on achievement motivation (Eccles, 1984; Eccles et al., 1984; Eccles et al., 1990; Meece, Wigfield, & Eccles, 1990; Wigfield, Eccles, & Blumenfield, 1997). That both boys and girls have differing levels of achievement is not surprising. More surprising is the fact that boys and girls have differing amounts of variability in their trajectories. Perceptions about the value of achieving in math and reading domains might be more strongly held by girls than by boys. This might influence girls to be more systematic in their evaluation of the value of their achievement goals. Prior research has shown that girls weigh the subjective value of an achievement activity more heavily than do boys (Eccles, 1984). This may cause girls to be more rigid in their perception of achievement success and of the successful paths toward achievement goals. Boys, on the other hand, are less affected by the subjective value of achievement tasks and may be more disposed to follow a variety of achievement paths. Consequently, boys may exhibit greater variability in the slopes of their achievement trajectories.

The other striking pattern that has emerged from these analyses is that "shocks" to their expected achievement trajectories appear more persistent for girls than boys. Girls' apparent lack of resilience might be due, in part, to greater value placed on prevailing gender stereotypes pertaining to achievement measures. For example, achievement in math is usually higher among boys (Eccles et al., 1984; Meece, Parsons, Kaczala, & Goff, 1982). It might be that math learning produces anxiety among girls that converts transitory shocks into more persistent deviations from expected trajectories. Results from our analyses fit with this explanation; girls appeared to have a more difficult time than boys recovering from shocks to their expected trajectories of achievement in math. Girls, however, also had a difficult time recovering from shocks to their trajectory of achievement in reading, which was puzzling because reading is generally thought to be an area in which girls traditionally do well.

A partial answer to this puzzle might be that girls value competency more than boys in all areas of achievement. Eccles' (1984) model suggested that girls' achievement decisions are more influenced by their self-perceptions. This may help explain why girls might have a more difficult time "getting back on track" after they have experienced an otherwise transitory shock to their expected trajectories. Girls might be more likely than boys to allow events in their lives to affect how they evaluate their potential for achieving in both math and reading. Given this, girls would be more affected by shocks to their achievement trajectories.

We also found that both girls and boys have considerable variability in the level of behavior problems, but only boys have considerable variability in the slopes of their trajectories. As with the achievement measures, girls are considerably less resilient to shocks away from their natural trajectories. If a girl is shocked away from her original pattern of behavior such that she exhibits higher levels of behavior problems, she might attach more meaning to this trend than would her male counterpart. This also is consistent with concepts from the deviance literature that indicate that there is a process by which children and adolescents come to see themselves as trouble makers (Becker, 1963; Lemert, 1967). This involves an initial act of primary deviance that may or may not lead to more deviant activities, or "secondary deviance," that would cause the child to view herself as a deviant person. If a girl places more subjective value on her

achievement and behavior decisions, then a shock away from an otherwise low behavior problem trajectory may be more likely to propel her down a path characterized by higher behavior problems. In this way, girls might be more likely to make the transition from primary to secondary deviant activities, or in this case increased behavior problems, because girls are more likely to give more meaning to initial evidence of problem behavior. In contrast, a boy might not give as much thought to the meaning of an initial shock away from low behavior problem profile. As a result, he might be more likely to quickly return to his original trajectory.

These, of course, are only possible explanations to the pattern of results we observed in our analyses. Our data included measures of neither perceived competence in math and reading achievement nor the value that boys and girls placed on high attainment in achievement domains. Similarly, we were unable to assess how exhibiting higher behavior problems affect the children's self-perceptions.

An important extension to these analyses is a separate consideration of negative and positive shocks to children's developmental trajectories. Our statistical model treated the two kinds of shocks symmetrically. In our discussion, we have emphasized processes associated with negative shocks. Processes associated with the persistence of positive shocks could be quite different. Further exploring this possibility is an important avenue for our future research.

Another extension is to identify what events or changes might be producing shocks to children's trajectories. For example, it is useful to consider the role of income events such as mothers beginning to receive welfare or experiencing a job loss, as well as family events such as a divorce or separation. Before introducing such covariates, however, it is important to first establish a model of the general structure of achievement and behavior in middle childhood. In this regard, these analyses represent an important first step.

ACKNOWLEDGMENTS

A previous version of this paper was presented at the 1998 Annual Meeting of the American Sociological Association, San Francisco, California, August, 1998. We are grateful to the MacArthur Network on Successful Pathways through Middle Childhood for supporting this research; to Rachel Dunifon for capable research assistance; and to Joe Altonji, Tom Cook, Jacque Eccles, Helena Kraemer, Bruce Meyer, John Modell, Steve Raudenbush, and Jean-Claude Ray for helpful comments.

ADDRESSES AND AFFILIATIONS

Corresponding author: Lori Kowaleski-Jones, Institute for Policy Research, Northwestern University, 2040 Sheridan Road, Evanston, IL 60208; e-mail: l-kowaleski@nwu.edu. Greg J. Duncan is also at Northwestern University.

REFERENCES

Achenbach, T. M. (1984). The status of research related to psychopathology. In A. Collins (Ed.), *Development during middle childhood: The years from six to twelve* (pp. 370–397). Washington, DC: National Academy Press.

Achenbach, T. M., & Edlebrock, C. S. (1981). Behavioral problems and competencies reported by parents of normal and disturbed children aged four through sixteen. *Monograph of the Society for Research in Child Development, 46*(1, Serial No. 188).

Baker, P., Keck, C., Mott, F. L., & Quinlan, S.V. (1993). *NLSY child handbook.* Columbus, OH: Center for Human Resource Research.

Becker, H. S. (1963). *Outsiders: Studies in the sociology of deviance.* New York: Free Press.

Belsky, J. (1984). The determinants of parenting: A process model. *Child Development, 55*, 883–896.

Brown, A., Bransford, J., Ferrara, R., & Campione, J. (1983). Learning, remembering, and understanding. In J. H. Flavell & E. M. Markman (Eds.), P. H. Mussen (Series Ed.), *Handbook of child psychology: Vol. 3. Cognitive development*, New York: Wiley.

Caspi, A., Elder, G. H., & Bem, D. J. (1987). Moving against the world: Life-course patterns of explosive children. *Developmental Psychology, 23*, 308–313.

Chase-Lansdale, P. L., Mott, F. L., & Brooks-Gunn, J. (1991). Children of the National Longitudinal Survey of Youth: A unique research opportunity. *Developmental Psychology, 27*, 918–931.

Cherlin, A. J., Chase-Lansdale, P. L., & McRae, C. (1998). Effects of parental divorce on mental health throughout the life course. *American Sociological Review, 63*, 239–249.

Collins, A. (1984). The status of basic research on middle childhood. In A. Collins (Ed.), *Development during middle childhood: The years from six to twelve* (pp. 398–421). Washington, DC: National Academy Press.

Cooksey, E. C., Menaghan, E. G., & Jekielek, S. M. (1997). Life course effects of work and family circumstances on children. *Social Forces, 76*, 637–665.

Crick, N. R., & Dodge, K. A. (1994). A review and reformulation of social information-processing mechanisms in children's social adjustment. *Psychological Bulletin, 115*(1), 74–101.

Dodge, K. (1993). Social-cognitive mechanism in the development of conduct disorder and depression. *Annual Review of Psychology, 44*, 559–584.

Dodge K. A. (1991). The structure and function of reactive and protective aggression. In D. J. Pepler & K. H. Rubin

(Eds.), *The development and treatment of childhood aggression* (pp. 201–218). Hillsdale, NJ: Erlbaum.

Dunn, L. M., & Markwardt, F. C. (1970). *Peabody Individual Achievement Test manual*. Circle Pines, MN: American Guidance Service.

Eccles, J. S. (1984). Sex differences in achievement patterns. In T. Sonderegger (Ed.), *Nebraska Symposium on Motivation* (Vol. 32, pp. 97–132). Lincoln: University of Nebraska Press.

Eccles, J. S., Adler, T., & Meece, J. L. (1984). Sex differences in achievement: A test of alternate theories. *Journal of Personality and Social Psychology, 46*, 26–43.

Eccles, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kazala, C. M., Meece, J., & Midgley, C. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motives* (pp. 75–146). San Francisco: W. H. Freeman.

Eccles, J. S., Jacobs, J. E., & Harold, R. D. (1990). Gender-role stereotypes, expectancy effects, and parents' role in the socialization of gender differences in self-perceptions and skill acquisition. *Journal of Social Issues, 46*, 182–201.

Erikson, E. H. (1963). *Childhood and society* (2nd ed.). New York: Norton.

Hause, J. C. (1980). The fine structure of earnings and the on-the-job training hypothesis. *Econometrica, 48*, 1013–1029.

Honzik, M. P., MacFarlane, J. W., & Allen, L. (1948). The stability of mental test performance between two and eighteen years. *Journal of Experimental Education, 18*, 309–324.

Lemert, E. M. (1967). *Human deviance, social problems, and social control.* Englewood Cliffs, NJ: Prentice-Hall.

Lillard, L. A., & Willis, R. (1978). Dynamic aspects of earning mobility. *Econometrica, 46*, 985–1012.

Loeber, R., & Hay, D. F. (1994). Developmental approaches to aggression and conduct problems. In M. Rutter, & D. F. Hay (Eds.), *Development through life: a handbook for clinicians.* Oxford: Blackwell.

Loeber, R., & Hay, D. F. (1997). Key issues in the development of aggression and violence from childhood to early adulthood. *Annual Review of Psychology, 48*, 371–410.

Loeber, R., Wung, P., Keenan, K., Giroux, B., Stouthamer-Loeber, M., Van Kammen, W. B. (1993). Developmental pathways in disruptive child behavior. *Development and Psychopathology, 5*, 101–132.

McCall, R. B., Appelbaum, M. I., & Hogarty, P. S. (1973). Developmental changes in mental performance. *Monographs of the Society for Research in Child Development, 38*, (3, Serial No. 150).

McFadyen-Ketchum, S. A., Bates, J. E., Dodge, K. A., & Pettit, G. S. (1996). Patterns of change in early childhood aggressive-disruptive behavior: gender differences in predictions from early coercive and affectionate mother-child interactions. *Child Development, 67*, 2417–2433.

McLeod, J. D., & Shanahan, M. J. (1996). Trajectories of poverty and children's mental health. *Journal of Health and Social Behavior, 37*(3), 207–220.

Meece, J. L., Parsons, J. E., Kaczala, C. M., and Goff, S. B. (1982). Sex differences in math achievement: Toward a model of academic choice. *Psychological Bulletin, 91(2)*, 324–348.

Meece, J. L., Wigfield, A., & Eccles, J. S. (1990). Predictors of math anxiety and its consequences for young adolescents' course enrollment intentions and performances in mathematics. *Journal of Educational Psychology, 82*, 60–70.

Menaghan, E. G., Kowaleski-Jones, L., & Mott, F. L. (1997). The intergenerational costs of parental social stressors: Academic and social difficulties in early adolescence for children of young mothers. *The Journal of Health and Social Behavior, 38*, 72–86.

Mott, F. L., Kowaleski-Jones, L., & Menaghan, E. G. (1997). Paternal absence and child behavior, does child gender and recency of paternal absence make a difference? *Journal of Marriage and the Family, 59*, 103–118.

Murphy, L. L., Conoley, J. C., & Impara, J. C. (1994). *Tests in print iv: an index to tests, test reviews, and the literature on specific tests.* Lincoln: University of Nebraska Press.

Parcel, T. L., & Menaghan, E. G. (1988). Measuring behavior problems in a large cross-sectional survey: Reliability and validity for children of the NLS youth. Unpublished manuscript, Columbus: Ohio State University.

Patterson, G. R. (1986). Maternal rejection: determinant or product of deviant child behavior? In W. W. Hartup & Z. Rubin (Eds.), *Relationships and development* (pp. 73–94). Hillsdale, NJ: Erlbaum.

Patterson, G. R. (1993). Orderly change in a stable world: the antisocial trait as chimera. *Journal of Consulting and Counseling Psychology, 61*, 911–919.

Piaget, J. (1971). The theory of stages in cognitive development. In D. R. Green, M. P. Ford, & G. B. Flamer (Eds.), *Measurement and Piaget.* New York: McGraw-Hill.

Pungello, E. P., Kupersmidt, J. B., Burchinal, M. R., & Patterson, C. J. (1996). Environmental risk factors and children's achievement from middle childhood to early adolescence. *Developmental Psychology, 32*, 755–767.

Rafferty, A. (1995). Bayesian model selection in social research. In P. Marsden (Ed.), *Sociological Methodology* (pp. 111–163). Washington, DC: American Sociological Association.

Scarr, S., & McCartney, K. (1983). How people make their own environments. *Child Development, 54*, 424–435.

Wachs, T. D. (1996). Known and potential processes underlying developmental trajectories in childhood and adolescence. *Developmental Psychology, 32*, 796–801.

Wigfield, A., Eccles, J. S., & Blumenfield, P. C. (1997). Change in children's competence beliefs and subjective task values across the elementary school years: A three-year study. *Journal of Educational Psychology, 89*, 451–470.

Willet, J. B., & Sayer, A. G. (1994). Using covariance structure analysis to detect correlates and predictors of individual change over time. *Psychological Bulletin, 116*, 363–381.