Journal of Educational Psychology
2005, Vol. 97, No. 3, 391–405

0022-0663/05/$12.00 DOI: 10.1037/0022-0663.97.3.391

Disruptive Behavior and School Grades: Genetic and Environmental Relations in 11-Year-Olds

Wendy Johnson, Matt McGue, and William G. Iacono
University of Minnesota—Twin Cities

Disruptive behavior may be associated with poor academic achievement due to covariance with attention problems and low IQ. Evidence is based on clinical samples and on associations between problem behaviors in young children and later achievement difficulties. The contemporaneous relations and their genetic and environmental influences have not been understood. Using the population-based Minnesota Twin Family Study, the authors observed this pattern of associations in 11-year-olds. About 75% of variance in latent inattention, ability, behavior, and school grades variables was genetic. Genetic influences on inattention and grades and on ability and grades were highly overlapping. Those on disruptive behavior were less closely related.

Keywords: genetic and environmental influences, inattention, school grades, disruptive behavior, twin study

Disruptive behavior has long been associated with poor academic achievement (e.g., see reviews in Farrington et al., 1990; Hinshaw, 1992a, 1992b). Youths displaying such behaviors tend to have higher high school dropout rates (Frick et al., 1991), lower college attendance rates (Hinshaw, 1992b), greater incidence of inadequate reading ability (Rutter, Graham, Chadwick, & Yule, 1976), below average achievement test scores (Nelson, Benner, Lane, & Smith, 2004), lower grade point averages (Fergusson & Horwood, 1995), and greater incidence of underachievement, defined as a gap between IQ and achievement (McCall, Evahn, & Kratzer, 1992). It is well known that disruptive behaviors are often highly persistent over time (Loeber, 1991; Patterson, Baryshe, & Ramsey, 1989), but the origins of the relation between the disruptive behaviors and the achievement difficulties are not clearly understood. There are four basic possibilities, summarized by Hinshaw (1992b): (a) Disruptive behaviors lead to achievement difficulties, (b) achievement difficulties lead to disruptive behaviors, (c) each leads to the other, and (d) the associations result from underlying common causes. In general, researchers in the field seem to have reached two basic conclusions. First, the association has a developmental progression (Hinshaw, 1992b). In the elementary school years, the association is indirect, driven primarily by links between disruptive behaviors and inattention problems. By adolescence, however, the association appears to be much more direct and to involve more explicitly antisocial behaviors. Second, all four possibilities are at least partially correct. That is, the relationships between disruptive behaviors and achievement difficulties are thought to develop through a series of reciprocal processes involving parents, children, and teachers within the contexts of the home, school, and peer group (Conduct Problems Prevention Research Group, 1992). This, of course, is an extremely broad statement, but it serves to emphasize the pervasiveness of the presumed effects.

For example, parents' failure to provide appropriate emotional support may contribute to the child's failure to learn appropriate behavioral control. When coupled with insufficient cognitive stimulation, the child may not be, upon school entry, in a position to learn in most programs. That is, the combination of preexisting patterns of disruptive behavior, such as antisociality and hyperactivity, with lack of academic readiness can impede the child's ability to learn early academic skills, particularly reading (Kazdin, 1993; Moffitt & Silva, 1988). The time spent engaging in antisocial and hyperactive behavior decreases the time the child spends on school-related tasks. Lack of time "on task" interferes with the further development of appropriate academic skills, which leads to greater rejection of the child by peers, teachers, and even parents and alienates the child from positive socializing agents (Patterson, Reid, & Dishion, 1992). Alienation increases the likelihood that the child will affiliate with deviant peers (Patterson, 1986), launching further behavioral problems as the child moves into adolescence (Elliott, Huizinga, & Ageton, 1985) and making educational attainment difficult. Both theory and empirical evidence suggest that this example is apt; that is, though all four possible explanations for the association between disruptive behavior and achievement difficulties are partially accurate, it is common that behavior problems have been manifest before the child even reaches school age, so that a direction of association from early disruptive behaviors to later achievement difficulties (Hinshaw, 1992b; Patterson et al., 1992) is apparent. For purposes of this article, we therefore consider achievement, as measured by school grades, to be the outcome resulting from several contributing factors, including disruptive behavior, thus focusing on Hinshaw's (1992b) first and fourth possibilities.

Wendy Johnson, Matt McGue, and William G. Iacono, Department of Psychology, University of Minnesota—Twin Cities.

This research was supported by U.S. Public Health Service Grants AA00175, DA 13240, and DA 05147. We thank the twins and their families and the recruiting, interviewing, data management, and lab staffs of the Minnesota Twin Family Study.

Correspondence concerning this article should be addressed to Wendy Johnson, Department of Psychology, University of Minnesota, 75 East River Road, Minneapolis, MN 55455. E-mail: john4350@tc.umn.edu

The conclusions regarding the association between disruptive behavior and school achievement are both stated in a developmental context, and most of the research on which they are based has been conducted longitudinally. This means that little is known about how the association is manifested at any point within the developmental period. The question examined here is the degree to which the association between disruptive behavior and school grades is direct or indirect when both are measured contemporaneously at age 11, between childhood and adolescence. The theory and evidence summarized by the example given above suggest that it would be reasonable to consider the association to be direct.

Distinguishing Processes

There is also evidence that the association between early disruptive behavior and later achievement difficulties is not direct. That is, associations between pre- and elementary school disruptive behavior and adolescent school performance can be explained by attention problems and low IQ (Fergusson & Horwood, 1995; Fergusson, Horwood, & Lynsky, 1993; Frick et al., 1991) that persist throughout childhood and adolescence. Thus, there is evidence that, rather than early pre- and elementary school disruptive behavior directly influencing later adolescent achievement difficulties, there may be two distinct but often co-occurring developmental processes (Fergusson & Horwood, 1995) that each operate separately to present the appearance of a direct association between early disruptive behavior and later achievement difficulties. First, early pre- and elementary school disruptive behavior may simply predict later adolescent delinquent behaviors regardless of school achievement status, the manifestation perhaps of a traitlike tendency toward disruptive behavior that increases in social significance as the child develops. Second, attention problems and/or low IQ that emerge in the pre- and early elementary school years may simply predict later adolescent achievement difficulties regardless of delinquent behavior status, the result perhaps of a poor match between the individual and common school expectations. The apparent link between disruptive behavior and achievement difficulties may thus be spurious: It may be that it is observed because many individuals experience both processes. This co-occurrence of disruptive behavior and educational mismatch processes could explain the association between early disruptive behavior and later achievement difficulties, but it is not clear to what degree the two processes can be differentiated over time, nor is it clear whether they can be distinguished at any particular point in time. Ability to distinguish between them both over time and at particular points in time is important for educational policy. The methods used to address disruptive behavior differ from those used to address attention problems and low IQ. Though schools have the opportunity to follow many children from early childhood into adolescence, allowing for the use of methods that can distinguish over time, there are other children who are presented to schools only in middle childhood and/or for relatively brief periods, requiring the use of methods that can distinguish at particular points in time.

In addition, ascertainment that disruptive behavior and educational mismatch processes can be distinguished contemporaneously would bolster evidence that the processes are in fact distinct, providing additional evidence for the need to ascertain whether there are also attention and/or ability difficulties whenever children present both disruptive behavior and poor school performance. This may be particularly true around age 11. At this age, children stand at the gateway between childhood and adolescence, when two relevant types of change take place. First, the incidence of disruptive behavior increases markedly for many children (Moffitt, 1993). Second, in the elementary years, the school curriculum is not generally highly differentiated, and grading systems tend to reflect individual progress and effort to a greater degree than they do later (Eccles, Midgley, & Adler, 1984), attenuating the association between classroom grades, which provide the child's most direct indication of successful performance, and standardized measures of achievement. As children move into adolescence, however, greater curriculum differentiation takes place, and grades tend to follow objective performance more closely, giving some children a rather different impression of their relative success in the school environment. Should the contemporaneous effects of disruptive behavior, attention problems, and ability on school grades around age 11 be different from the effects that research to date has identified as taking place over time, it would suggest the existence of significant upheavals in those relationships during adolescence, which might have their roots in these two relevant changes.

Measuring Achievement

Most studies investigating the associations among disruptive behavior, inattention and ability problems, and achievement have used either measures of reading performance (Hinshaw, 1992b) or standardized achievement tests (e.g., Fergusson & Horwood, 1995; Fergusson et al., 1993; Frick et al., 1991). The rationale for using reading tests to measure achievement has been that literacy is critical to success in industrialized societies, and thus, measures of reading skill provide key evidence of academic skill. Yet academic skill is clearly more than merely reading skill, and reading comprehension, or the ability to make functional use of reading skill, may, after some skill acquisition period, be more closely related to general intelligence than it is to reading skill alone (Johnson, Bouchard, Segal, & Samuels, 2005). In addition, there is substantial evidence that though reading is a learned skill, the ability to acquire it and the disabilities that interfere with its acquisition are genetically influenced in ways very similar to intellectual ability (Grigorenko, 2001), thus suggesting that reading ability may be related to school performance primarily because of their shared association with general intelligence. The rationale for using standardized achievement tests is that they provide a uniform yardstick by which to compare the performance of children coming from different schools. But, though we treat academic achievement as the outcome variable in this study, there is also evidence that academic failure reinforces or aggravates existing behavior problems (Hinshaw, 1992b). Thus, we might expect—and there is some evidence to this effect (Hawkins & Lishner, 1987)—that the link between the disruptive behavior and academic achievement would be stronger when the measure of achievement reflects actual performance in the classroom, as do grades.

There would be two reasons for this. First, grades are feedback the child receives directly and so are more likely to affect the child's perception of academic success or failure, and, second, grades are more likely than standardized test performance to reflect any difficulty the teacher has in dealing with the child's disruptive behavior in the classroom (Hinshaw, 1992b). There is thus the potential for perceptual biases to inflate the association by producing common variance between the measures of achievement and disruptive behavior. This means that using school grades as the measure of achievement and reading ability as a measure of ability in conjunction with some more general measure, such as IQ, provides a stronger test of whether inattention and ability problems explain the association between disruptive behavior and achievement, at least as measured by grades. This is not meant to imply that grades are overall a better measure of achievement than standardized tests but merely that grades provide a stronger test of the particular question of interest here involving the existence of a direct association between disruptive behavior and academic achievement.

To the extent that there are no direct associations between disruptive behavior and achievement, there should be some individuals who display disruptive behavior but continue to earn high grades in school. Other researchers have noted this as well (e.g., Clark, Prior, & Kinsella, 2002). Still, description of the characteristics of children displaying this pattern contemporaneously, especially when grades serve as the measure of academic achievement, may provide insight into the ways in which behavior and academic problems can be disentangled. This could help to provide curriculum and guidance ideas for school professionals working with children who appear to be developing disruptive behavior patterns by suggesting possible reasons for their disruptive behavior.

Behavior Genetic Studies

There have been many twin and adoption studies that have investigated the extent of genetic and environmental influences on disruptive behavior (e.g., Jacobson, Prescott, & Kendler, 2000; Rowe, 1983; Taylor, Iacono, & McGue, 2000), attention problems (for a review, see Thapar, 2003), academic achievement (for a review, see Thompson, Detterman, & Plomin, 1991), and intelligence or academic ability separately (e.g., Alarcon, Plomin, Fulker, Corley, & DeFries, 1998; Bouchard & McGue, 1981; McCartney, Harris, & Barnieri, 1990). These studies have consistently revealed substantial genetic contributions to all four of these phenotypes (or traits), despite wide variation in the age groups involved in the samples and the specific measures used to assess the phenotypes. In addition, though the shared or family environment appears to exert considerable influence on disruptive behavior limited to adolescence (Moffitt, 1993), genetic influence appears to be greater in individuals who start young and continue to engage in antisocial behaviors throughout their lives (e.g., Lyons et al., 1995; Moffitt & Silva, 1988; Taylor et al., 2000). Genetic influences on intelligence or academic ability increase throughout the life span, whereas environmental influences that make all members of a family more similar fall to zero by early adulthood (e.g., McGue, Bouchard, Iacono, & Lykken, 1993; Plomin, Fulker, Corley, & DeFries, 1997; Skodak & Skeels, 1949). There is evidence for substantial but far from complete common genetic influence on inattention and disruptive behavior (Burt, Krueger, McGue, & Iacono, 2003; Nadder, Rutter, Silberg, Maes, & Eaves, 2002) and a similar relationship between the genetic influences on academic achievement and intelligence or academic ability (e.g., Petrill & Thompson, 1993; Thompson et al., 1991). However, the relationships among genetic and environmental influences on these four variables together have not been explored in detail, and achievement measures have generally been standardized tests, which may be more closely related to ability than are day-to-day measures, such as grades.

Development of a fuller understanding of these relations would help to articulate ways in which the educational environment can be used both to develop necessary academic skills in all children and to develop additional skills in those who are capable and interested. To the extent that genetic influences explain substantial proportions of individual differences, it will be important to tailor interventions to the specific characteristics of the individuals involved. Thus, for example, when we understand that reading disability runs in a particular child's family, educators might be alert for the emergence of reading difficulties in that child and tailor early interventions to the particular difficulties that are manifest. Such a child might also be expected to be more vulnerable to the emergence of different behavior problems than would a child experiencing reading difficulties due to mental retardation. This is an example of how an intervention might make use of information about a genetic vulnerability, but the point here is not to develop separate interventions that could be applied according to whether the origin of the difficulty is genetic or environmental. Rather, the point is to develop a fuller understanding of how children's behavioral and academic problems emerge in order to develop truly effective interventions, and this will involve recognition of the existence of genetic as well as environmental vulnerabilities. Transaction between genetic and environmental influences is one of the likely realities of the processes involved, and one of the first steps in understanding them is quantifying the magnitude of the genetic and environmental influences in various samples. In addition, to the extent that the same genetic influences affect manifest (phenotypic, resulting from the transaction of genetic and environmental influences) traits that can be distinguished from each other, we might expect similar effects on the traits from environmental interventions. Thus, if, for example, the same genes influence both inattention and disruptive behavior, an environmental intervention that is effective in reducing the manifestations of one might also reduce the manifestations of the other. There should be less reason to have this expectation if the genetic link between the two traits is smaller.

Some might object that this is an aptitude by treatment interaction, considered only at the genetic level, and we know that there are no such effects at the observed or phenotypic level. In fact, there is considerable evidence that such effects do exist at the phenotypic level in education, as well as in parenting, and other therapeutic programs. For example, students susceptible to test anxiety tend to perform better in highly structured instructional programs, but these programs do not similarly benefit nonanxious students (Snow, 1991). Sternberg, Grigorenko, Ferrari, and Clinkenbeard (1999) reported that students performed better when placed in a course that matched their pattern of abilities than when placed in a course that did not. Beutler et al. (2003) reported similar kinds of effects in treating comorbidly depressed and chemically dependent patients. At the temperamental level, there is evidence that behaviorally inhibited children respond better to some forms of parenting than to others (Kagan, 1999) and that impulsive youth are at much greater risk for delinquency in poor neighborhoods than in neighborhoods not considered poor (Lynam et al., 2000). Thus, there would appear to be little reason to reject the plausibility of an aptitude by treatment interaction at the genetic level. In fact, there is increasing evidence that gene–

environment interactions are common (Moffitt, Caspi, & Rutter, 2005).

Gender

There are substantial mean gender differences in academic achievement as well as in disruptive behavior and attention problems. Both disruptive behavior and attention problems are much more common in males than in females (e.g., for disruptive behavior, see Butts et al., 1995; Cohen et al., 1993; for attention problems, see Gomez, Harvey, Quick, Sharer, & Harris, 1999; Rhee, Waldman, Hay, & Levy, 2001), with ratios of males to females on the order of 4–5:1 commonly observed. In general, females receive higher grades than males and score higher on achievement tests, from elementary school through college (e.g., Kimball, 1989; Mau & Lynn, 2001), in spite of the fact that males tend to score higher than females on college and other aptitude tests (Mau & Lynn, 2001). Effect sizes for the sex differences in achievement vary considerably, depending on the measure used and the age level, but can range as high as .5 standard deviation. Effect sizes for the sex differences in aptitude are generally very small. These data suggest that something about the school environment is more effective for students with attributes more common in females than in males.

Aims of Current Study

This synopsis raises several questions that we addressed in the current study. Using the intake assessment of the 11-year-old cohort of the Minnesota Twin Family Study (MTFS), a population-based sample of twins, we asked the following questions: First, considering academic achievement, measured by grades, to be the outcome, can educational mismatch reflected by inattention problems and poor academic ability, measured by both general intellectual ability and reading ability, explain contemporaneous associations between disruptive behavior and grades at the transition between childhood and adolescence? We expected to answer this question positively. Doing so would extend previous research by making the observation contemporaneously at age 11, a point at the gateway to adolescence, and by using a measure of day-to-day academic performance clearly visible to the child. We also expected to be able to identify and describe some children who behaved disruptively yet continued to earn good grades. Second, we asked how the genetic and environmental influences on all four variables (disruptive behavior, inattention, academic ability, and grades) were related. Third, we asked how both the phenotypic and genetic and environmental models for boys and girls compared, seeking possible differences in the associations among the variables that might help to explain the gender differences in achievement levels.

Method

Sample

Participants were drawn from the intake assessment of the 11-year-old cohort of the ongoing MTFS, a longitudinal study of a community-based sample of like-sex twins and their parents. The MTFS sample was compiled by using a population-based method. Starting from state birth records, the current status and location of more than 90% of the like-sex twin pairs born in Minnesota in the targeted years were determined by using various publicly available databases. Located twins who were without any significant physical or mental handicap and living within a day's drive of Minneapolis with at least one biological parent were invited to complete a day-long, in-person assessment at our labs at the University of Minnesota; less than 20% declined. The parents in the participating families were generally born between 1950 and 1965 and were generally between the ages of 30 and 40 at the time of interview. Fathers averaged slightly over 14.5 years of education, and mothers averaged about 1 year less. The average Hollingshead occupational level for the families was about 4, indicating possession of jobs that required some education just beyond the skilled blue-collar level, although the sample included parents working in highly professional occupations as well as parents unemployed or working in semiskilled jobs (the standard deviation was just under 2 Hollingshead levels). In addition to the demographic information provided by the participating families, more than 80% of the families who did not participate completed a brief mail or telephone survey, enabling some comparison of participants and nonparticipants. This comparison revealed that parents in participating families were significantly, though only modestly, better educated than were those in nonparticipating families, with a mean difference of less than 0.3 years of education. The two groups of families did not differ significantly in self-reported mental health. The MTFS intake sample is thus generally representative of families with twins born in Minnesota during the period from the early 1970s to the early 1990s. A complete description of the ascertainment and assessment procedures used in the MTFS as well as an analysis of nonparticipants is given in Iacono, Carlson, Taylor, Elkins, and McGue (1999).

The intake assessment of the 11-year-old cohort was administered when the twins were on average 11 years old, although a few were not quite 11 and some had recently turned 12 at the time of assessment. Data were available for 376 pairs of boys (254 monozygotic [MZ], 122 dizygotic [DZ]) born in 1977–1982 and 424 pairs of girls (259 MZ, 165 DZ) born in 1981–1984 and in 1988. Consistent with the demographics of Minnesota for the birth years sampled, over 98% of the twins are Caucasian. There is a small cohort difference (3–5 years) that might appear to confound our observation of gender differences. The sample contains both girls and boys born in 1981 and 1982, however, and there is no difference in reported grades for the girls and boys born in those years compared with those born in other years, suggesting that the cohort difference does not confound gender comparisons.

Measures

Inattention. Inattention was assessed both by interview and self-report questionnaire. Parents reported on inattention-related behaviors during a formal diagnostic interview,[1] the Diagnostic Interview for Children and Adolescents—Revised (DICA–R; Welner, Reich, Herjanic, Jung, & Amada, 1987). These diagnostic interviews were tabulated and symptom counts were assigned according to *Diagnostic and Statistical Manual*

[1] The twins completed a diagnostic interview on themselves as well, responding to the same questions about their own behavior as did their parents. The twins also completed self-report questionnaires covering many of the same issues about their behavior and attitudes as in the interview completed by their parents. The child's report variables were much less correlated both with each other and with either the parent's or teachers' report variables than were the parent's and teachers' reports among themselves and with each other, however, especially for the inattention variables. Self-reports of attention-related difficulties tend to be less accurate than parents' and teachers' reports (Barkley, Fischer, Smallish, & Fletcher, 2002), and the lower correlations associated with the child's reports made it difficult to form the latent variables that were the foundation of our analysis. We did carry out the analysis both with and without the children's report variables, with very similar results. As the results were more coherent without the children's reports, we present and discuss them in that manner.

(*DSM*) criteria during a consensus process conducted by teams of two qualified graduate students. In this consensus process, teams of two reached complete agreement as to the appropriateness of the symptom assignment. We made use of the symptom counts for attention deficit disorder (not including hyperactivity symptoms as we intended to focus specifically on inattention) as reported by parents (generally mothers) from the revised third edition of the *DSM* (*DSM–III–R;* American Psychiatric Association, 1987), which was the diagnostic system current at the time of assessment. The symptoms in *DSM–III–R* overlap substantially with those in the fourth edition of the *DSM* (American Psychiatric Association, 1994), the current diagnostic standard. For these symptoms, we assigned a value of 2 for a rating of *definitely present* and 1 for a rating of *possibly present*. Possible scores ranged from 0 to 12.

Parents also reported on behaviors related to school as part of the self-report questionnaire. Among many other questions, they rated the statements "Has difficulty concentrating or paying attention in class" and "Easily distracted in class" on a 4-point scale ranging from 1 (*definitely true*) to 4 (*definitely false*). Scores for these two items were summed. Possible scores thus ranged from 2 to 8. Alpha coefficients for the two items were .80 and .87 for girls and boys, respectively.

As many as four different teachers nominated by each twin completed an extensive report of student behavior and achievement. Minnesota state guidelines stipulate that twins should be in different classrooms whenever possible; thus, co-twins usually did not nominate the same set of teachers. Teachers represent an important source of information because they assigned some of the grades on which they reported; they saw the twins in a normative, structured setting on a frequent basis away from their immediate families; and they had no a priori emotional ties to them. They are thus able to put the twins' behavior in a broader and less biased context than are parents. At the same time, however, the twins' behavior may be somewhat more constrained at school than at home as the setting at school is public. Return rates on the Teacher's Rating Forms exceeded 70%, but because multiple teachers were nominated by each child, at least one teacher rating was available for 81% of the sample. Conduct, oppositional and defiant behaviors, and inattention were rated by teachers with items we developed that modeled items from the Conners Teacher Rating Scale (Connors, 1969), the Rutter Child Scale B (Rutter, 1967) and *DSM* criteria. The 67-item inventory asked teachers to assess the applicability of the specified behavior to the student compared with the average student in the classroom on a 4-point scale ranging from *not at all* (1) to *very much* (4). The specific items used to rate inattention, conduct, and oppositional and defiant behaviors were determined by factor analysis. Average internal consistency and interteacher agreement reliabilities were .96 and .75, respectively, for the inattention-related items. We computed average teacher scores for inattention on the basis of the number of teacher reports obtained for each twin for this study. Scores ranged from 8 to 32.

Ability. The twins were assessed at age 11 with an abbreviated version of the Wechsler Intelligence Scale for Children—Revised (WISC–R). The abbreviated versions of these tests consist of two verbal (Vocabulary and Information) and two performance (Block Design and Picture Arrangement) subscales. These subscales were selected for their high correlation (.90) with total WISC–R IQ based on all subscales. The twins also receive the reading test from the Wide Range Achievement Test (WRAT). We made use of the grade-equivalent reading score as a second measure of ability because reading ability is so fundamental to academic performance. WISC–R and WRAT scores correlated at .43 for girls and .48 for boys.

Disruptive behavior. As for inattention, we made use of the *DSM–III–R* symptom counts for oppositional and defiant disorder (ODD) and conduct disorder (CD) as reported by parents from the DICA. For these symptoms, we assigned a value of 2 for a rating of *definitely present* and 1 for a rating of *possibly present*. Possible scores thus ranged from 0 to 18 for ODD and 0 to 26 for CD. We also made use of the items loading on the conduct and oppositional and defiant factors from the teachers' rating of CD (T-CD) and ODD. Average internal consistencies were .93 for the items loading on the conduct factor and .80 for the items loading on the oppositional and defiant factor. Interteacher agreement reliabilities averaged at approximately .75. As for inattention, we computed average teacher scores on the basis of the number of teacher reports obtained for each twin for this study. Possible scores ranged from 8 to 32 for ODD and 14 to 56 for CD.

School grades. In general, the MTFS does not collect data on actual grades. Rather, parents and teachers report separately on student grades in language arts, math, social studies, science classes, and overall by indicating that the grades are much better than average (As = 4), better than average (Bs = 3), average (Cs = 2), below average (Ds = 1), or much below average (i.e., failing = 0). Possible scores thus ranged from 0 to 4. This approach was taken because of the disparity in grading formats, procedures, and standards taken in the various school systems from which the MTFS families are drawn. We made use of the parent's report of overall grades and computed average teacher scores on the basis of the number of teacher reports obtained for each twin as our measures of school grades. The estimated internal consistency reliability for the teachers' grade reports was .92, and estimated interteacher agreement reliability was .87. We believed the use of both parents' and teachers' reports in this situation to be important. Although the reporting teachers did provide the participating twins with some of their grades, they did not provide them with all of them and, unlike the parents, would not have been in the position to view all of the participant's grades. At the same time, the parents would be less likely to have a clear basis on which to compare their own offspring's grades with those of others. Thus, as is often the case, each reporter provided information with different strengths and weaknesses. In addition, as noted above, although grades may reflect classroom behavior as well as actual achievement, particularly in the elementary school years, any contamination of this sort should serve to increase the observed relationship we have hypothesized to be insignificant in the presence of appropriate controls. Thus, the use of grades provides a stronger test of this hypothesis than would more standardized measures of achievement.

Academic engagement. As part of the self-report questionnaire on school behaviors, twins reported on the degree to which they were interested in schoolwork, studied without being reminded, turned in their homework, and wanted good grades. These items were rated on a 4-point scale ranging from 1 (*definitely true*) to 4 (*definitely false*). Although not directly a part of our investigation, we used this measure to help characterize children who behaved disruptively yet earned high grades. We reverse scored the items and summed the reverse scores to produce a scale we labeled Engagement. Possible scores ranged from 4 to 16. Estimated internal consistency reliabilities for the scale were .74 for girls and .85 for boys.

Analytical Approach

Phenotypic analyses. We made use of maximum likelihood estimation as implemented in LISREL Version 8.53 (Jöreskog & Sörbom, 2003) to fit the phenotypic model of the manifest traits designed to address our first question: Is disruptive behavior independently associated contemporaneously with school grades at the transition between childhood and adolescence? In doing so, we treated all of our variables as continuous, even though, strictly speaking, most of them were ordinal. This is sometimes done when the ordinal variables have numerous categories and a clear quantitative ordering, as was the case with our variables (Tabachnick & Fidell, 2001). We also treated girls and boys as separate groups and used our multiple reports of each construct to form latent inattention, ability, disruptive behavior, and grade variables. We thus refer below to latent phenotypes, by which we mean latent variables resulting from the transaction of latent genetic and environmental influences. To make sure that the parameter estimates from the groups of girls and boys were comparable, we focused on the common metric solution produced by LISREL. In this solution, the standardized solution is computed by using a weighted average of the covariance matrices for the two groups to form the underlying correlation matrix. This preserves at the standardized level any equality constraints that have been placed across groups at the level of the raw data. The resulting solution matrices, however, must be interpreted as covariances relative to the common metric rather than as correlations. For example, when the variances in one group are greater than those in another, standardized covariances greater than 1.00 can result.

We made no adjustment for the lack of independence between twin pairs within the sample. This should have the effect of inflating the model fit statistics so that the model appears to fit more closely but should have little effect on parameter estimates themselves, though it will also tend to inflate the significance of particular paths (McGue, Wette, & Rao, 1984; M. C. Neale, 2003, personal communication, January 15, 2003). To minimize the potential effect of this, we evaluated path significance by only using the overall chi-square statistics rather than the *t* statistics associated with particular paths. We note, however, that our primary goal was to test the significance of disruptive behavior as an independent predictor of school grades. Therefore, our test was especially stringent. The phenotypic model we used is diagrammed in Figure 1. Though not shown in Figure 1 because they were not the focus of our attention, we allowed for rater effects by including parent and teacher residual factors. It would, of course, be possible to specify different models of these basic relationships, with different causal paths. It was not our intent, however, to test the direction of the causal paths in this study but rather to test the significance of the path from disruptive behavior to school grades by assuming, on the basis of prior theory and empirical data, that that was the appropriate causal direction. Thus, we did not develop and compare alternative models.

We assessed model fit by using the Bayesian Information Criterion (BIC; Raftery, 1995; $\chi^2 - [df \times \text{natural log of } n]$, with n set at the geometric mean of the number of twin pairs in each group for conservatism) and the root-mean-square error of approximation (RMSEA; Browne & Cudeck, 1992). For BIC, smaller or more negative values are preferred. For RMSEA, values less than .05 indicate a close fit; values between .05 and .08 indicate a reasonable fit. To reduce the effects of outlying observations, transformed values greater than three standard deviations from the mean were trimmed. This affected WISC–R scores for 2 boys, one scoring 3.8 standard deviations above the mean (trimmed to 3.2) and one scoring 3.9 standard deviations below the mean (trimmed to −3.25).

Analyses of genetic and environmental influences. We explored the associations among the genetic and environmental influences on all four latent variables (disruptive behavior, inattention, academic ability, and school grades) by using standard quantitative genetic structural equation models estimated with maximum-likelihood analysis as implemented by Neale (1997). Some data were missing for some participants. There were varied reasons for this. Although there may have been some instances when the reason for the absence of data related directly to the measure in question (e.g., the participant refused to complete a specific measure on the basis of the presentation of the instructions and sample items), the most common reason for missing data was lack of time during the assessment session or failure to return a mail-back questionnaire. Thus, in general, the assumption that the data were missing at random (Little & Rubin, 1987) appeared to be appropriate. We therefore used maximum likelihood estimation to generate the covariance matrices for analysis in order to make use of all of the available data.

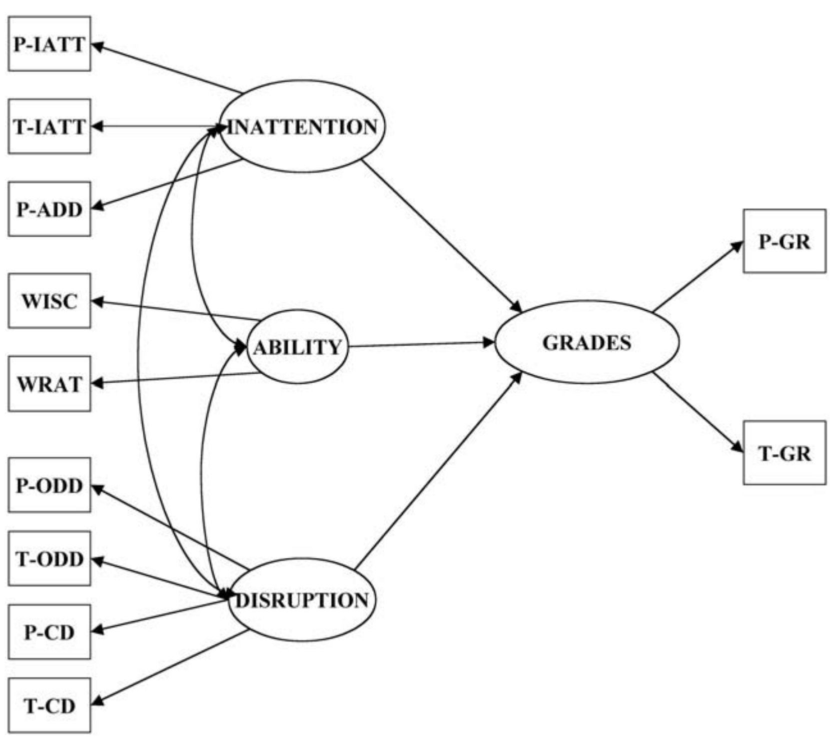

Figure 1. Basic phenotypic model. P-IATT = parent's report of attention problems; T-IATT = teachers' report of attention problems; P-ADD = parent's report of attention deficit disorder symptom count, based on the Diagnostic Interview for Children and Adolescents—Revised (DICA–R); WISC = composite IQ from the Wechsler Intelligence Scale for Children—Revised subtests used; WRAT = grade-equivalent reading level, based on the Wide Range Achievement Test; P-ODD = parent's report of oppositional and defiant disorder behavior, based on DICA–R symptom counts; T-ODD = teachers' report of oppositional and defiant disorder behavior, based on the Teacher's Rating Forms; P-CD = parent's report of conduct disorder behaviors, based on DICA–R symptom counts; T-CD = teachers' report of conduct disorder behavior, based on the Teacher's Rating Forms; P-GR = parent's report of child's grades; T-GR = teachers' report of child's grades.

The standard univariate quantitative genetic model is based on the understanding that the observed phenotypic variance (Vp) is a linear additive function of genetic (A) and shared (C) and nonshared (E) environmental variance, respectively. Symbolically,

$$Vp = A + C + E.$$

Under this model, the shared environmental variance represents experiential factors common to the members of a twin pair and operating to make them similar. They include such experiences as growing up in the same neighborhood and socioeconomic status. Nonshared environmental influences are those experiential factors unique to each member of a twin pair and operating to make them different. Such experiences may include injuries and illnesses, attending different schools, and participating in different leisure activities, such as sports. The distinction between the two is subtle. For example, two children in the same family may experience the same event (e.g., parental divorce), but that event is only a shared environmental influence to the extent that it makes the children similar. The nonshared environmental component also includes variance attributable to measurement error. Genetic variance can be additive in the sense that if multiple genes influence the trait, they do so independent of each other. It can also be nonadditive, reflecting dominance and other polygenic effects. The standard univariate quantitative genetic model is not identified if all four components of variance (additive and nonadditive genetic and shared and nonshared environmental) are estimated at once; one of the components must be dropped. Given only additive genetic effects, the expected covariance (COV) between any two members of a twin pair as a function of the variance components given above can be specified as

$$COV_{(MZ)} = A + C$$

$$COV_{(DZ)} = .5{*}A + C.$$

We did not fit models that included nonadditive genetic effects because examination of the basic MZ and DZ twin correlations provided no evidence for their appropriateness.

The standard univariate model can be extended to multivariate situations by modeling the covariance between one twin's score on one variable and the other twin's score on another variable in a manner directly analogous to the univariate case. We made use of a latent factor version of the Cholesky model, which can be used to establish baseline parameter estimates of genetic and environmental influences. Implementation of this model relies on the fact that any positive definite covariance matrix can be decomposed (uniquely, except for transformations of sign) into the product of a lower triangular matrix and its transpose. This can be done separately for the genetic and environmental portions of the latent phenotypic variance described above, which makes it possible to estimate the extent to which genetic and environmental influences are correlated across variables. When genetic correlations are high, there is evidence that genetic effects on one variable contribute to genetic effects on the other, and similar statements can be made for shared and nonshared environmental correlations. The model we used is shown in Figure 2. We used a similar model to decompose the residual variance common to each of the two raters, parents and teachers, mainly for the purpose of accounting accurately for the variance associated with these factors as they were not the focus of this study. We compared results for girls and boys at both the phenotypic and genetic and environmental levels by comparing model fit when the latent variable paths were constrained equally across the two groups.

Results

Descriptive Statistics

Table 1 shows the means and standard deviations for each of the measures we used, separately for girls and boys. As expected, there were significant mean differences between girls and boys for most of the measures, and the effect sizes of several of the differences, particularly for inattention, were moderate to large. Most were in the directions indicated by prior research, although others have not found consistent differences in ODD (e.g., Lahey et al., 2000). There were no significant mean differences between MZ and DZ twins for either girls or boys. In order to reduce skewness, all the inattention and disruptive behavior variables were log transformed prior to further multivariate analysis.

Table 2 shows the correlations among the transformed variables we used, separately for girls and boys. The correlations between

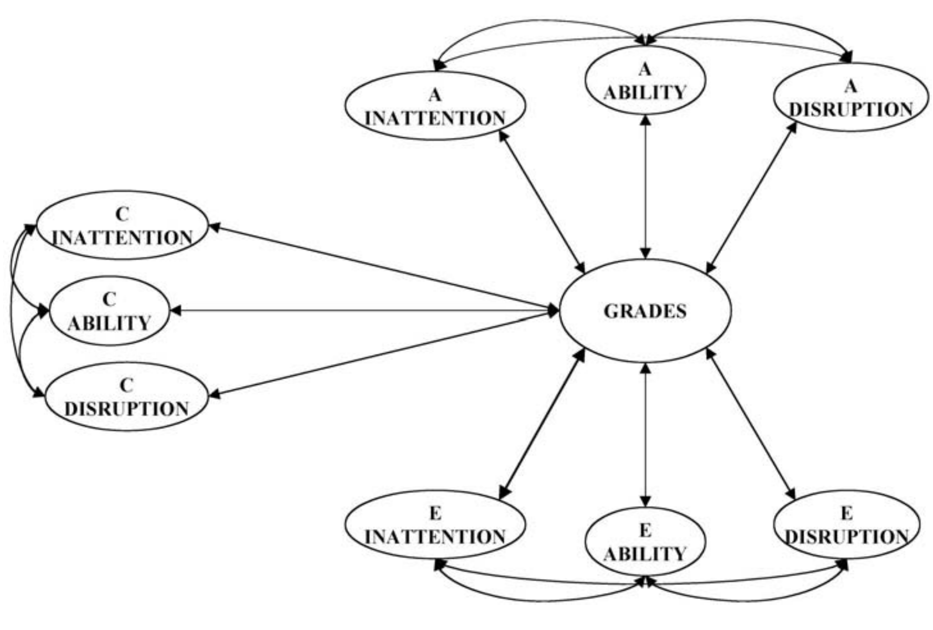

Figure 2. Basic gene–environment model. Genetic and environmental influences on common rater effects in the residuals were included in the model but are eliminated here for clarity. A = genetic influences; C = shared environmental influences; E = nonshared environmental influences.

Table 1
Descriptive Statistics and Effect Sizes of Mean Differences and Their Significance for Girls and Boys

	Girls		Boys		Differences	
Measure	*M*	*SD*	*M*	*SD*	Effect size	*p*
Inattention						
P-IATT	1.3	1.5	1.9	1.8	.36	<.001
T-IATT	9.8	3.1	12.9	5.0	.70	<.001
P-ADD	1.0	1.5	1.7	2.7	.30	<.001
Ability						
WISC	102.4	14.0	104.6	13.7	.16	.001
WRAT	6.1	2.1	5.7	2.0	−.18	.003
Disruptive behavior						
P-ODD	6.5	5.5	4.7	4.0	−.35	<.001
T-ODD	8.3	0.9	8.5	1.4	.18	.001
P-CD	0.6	1.6	1.1	1.8	.27	<.001
T-CD	16.4	4.5	18.2	6.6	.30	<.001
Grades						
P-GR	3.12	0.70	2.71	0.84	−.46	<.001
T-GR	2.99	0.46	2.77	0.87	−.25	<.001

Note. Effect size = mean difference divided by pooled standard deviation, stated so that it is positive when the mean for boys is higher; P-IATT = parent's report of attention problems; T-IATT = teachers' report of attention problems; P-ADD = parent's report of attention deficit disorder symptom count, based on the Diagnostic Interview for Children and Adolescents—Revised (DICA–R); WISC = composite IQ from the Wechsler Intelligence Scale for Children—Revised subtests used; WRAT = grade-equivalent reading level, based on the Wide Range Achievement Test; P-ODD = parent's report of oppositional and defiant disorder behavior, based on DICA-R symptom counts; T-ODD = teachers' report of oppositional and defiant disorder behavior, based on the Teacher's Rating Forms; P-CD = parent's report of conduct disorder behaviors, based on DICA-R symptom counts; T-CD = teachers' report of conduct disorder behavior, based on the Teacher's Rating Forms; P-GR = parent's report of child's grades; T-GR = teachers' report of child's grades.

the two reporters of grades were about .68, indicating that parents and teachers were reporting something very similar about the children. There were also strong, although less consistent, correlations between the two reporters on the measures of inattention and disruptive behavior, with the parents and teachers agreeing more about inattention than they did about disruptive behavior. The inattention measures tended to correlate more highly with grades than they did with each other. Together, this suggested

Table 2
Correlations Among Study Variables for Girls and Boys

Variable	1	2	3	4	5	6	7	8	9	10	11
1. P-IATT	—	.47	.47	−.28	−.28	.24	.20	.11	.26	−.47	−.48
2. T-IATT	.47	—	.40	−.35	−.33	.21	.41	.22	.44	−.51	−.61
3. P-ADD	.56	.45	—	−.21	−.20	.21	.14	.19	.19	−.38	−.38
4. WISC	−.26	−.27	−.13	—	.43	−.13	−.10	−.10	−.09	.53	.55
5. WRAT	−.28	−.29	−.21	.48	—	−.14	−.09	−.09	−.09	.45	.52
6. P-ODD	.20	.19	.27	−.05	−.03	—	.10	.32	.11	−.21	−.21
7. T-ODD	.20	.38	.13	−.09	−.08	.21	—	.16	.51	−.14	−.26
8. P-CD	.20	.22	.24	−.08	−.08	.41	.21	—	.16	−.17	−.20
9. T-CD	.30	.57	.26	−.07	−.09	.24	.65	.21	—	−.17	−.15
10. P-GR	−.50	−.50	−.31	.47	.49	−.16	−.18	−.15	−.23	—	.67
11. T-GR	−.49	−.63	−.37	.48	.49	−.13	−.20	−.15	−.30	.69	—

Note. Correlations for girls are presented above the diagonal, and correlations for boys are presented below the diagonal. Significance levels of the correlations vary slightly because of varying amounts of missing data, but a correlation of .10 is significant at approximately $p = .01$, and a correlation of .05 is significant at approximately $p = .05$. P-IATT = parent's report of attention problems; T-IATT = teachers' report of attention problems; P-ADD = parent's report of attention deficit disorder symptom count, based on the Diagnostic Interview for Children and Adolescents—Revised (DICA–R); WISC = composite IQ from the Wechsler Intelligence Scale for Children—Revised subtests used; WRAT = grade-equivalent reading level, based on the Wide Range Achievement Test; P-ODD = parent's report of oppositional and defiant disorder behavior, based on DICA-R symptom counts; T-ODD = teachers' report of oppositional and defiant disorder behavior, based on the Teacher's Rating Forms; P-CD = parent's report of conduct disorder behaviors, based on DICA-R symptom counts; T-CD = teachers' report of conduct disorder behavior, based on the Teacher's Rating Forms; P-GR = parent's report of child's grades; T-GR = teachers' report of child's grades.

substantial rater factors involving grades and inattention, but the disruptive behavior measures and grades did not show the same pattern. The ability measures were also strongly correlated with school grades, on the order of .45. The girls' correlations tended to be slightly higher than the boys'.

Phenotypic Analyses

Because we intended to compare our results for girls and boys, we began by assessing the extent to which we were measuring the same constructs in the two groups (Meredith, 1993). We began by fitting the same factor model in the two groups, allowing all factor loadings and path coefficients to vary freely; $\chi^2(48, N = 1596) = 67.53$, RMSEA = .032, BIC = −219.94. We proceeded to constrain the factor loadings equally across the two groups. If this level of "metric invariance" (Meredith, 1993) cannot be established for a set of measures, comparison of groups on these measures can produce ambiguous findings (Horn & McArdle, 1992). For this model, $\chi^2(55, N = 1596) = 84.95$, RMSEA = .037, BIC = −244.44. Because the chi-square statistic does not reflect model parsimony and tends to produce significant (ill-fitting) results when sample sizes are large (Raftery, 1995), because the RMSEA continued to indicate a close fit, and because BIC favored the metric invariant model, we considered the model with equal factor loadings to provide the best fit, indicating that there was considerable consistency in the assessment of our constructs across gender and providing reasonable grounds for the gender comparisons we planned to undertake. All of the results we present below are based on models with the factor loadings constrained equally across the two groups.

We first reviewed the zero-order correlations and covariances among the four latent variables, shown in Table 3. All of the correlations were significant, as we expected, and the covariances showed that there was more variance in the data for boys than for girls. We were particularly interested in the association between disruptive behavior and school grades because that was the focus of our analysis. For girls, the correlation was −.58, indicating that girls who displayed more disruptive behavior tended to have lower grades. The correlation for boys was −.38. The correlations between ability and school grades were very strong (about .9 for both boys and girls), indicating that the variance in common to the WRAT reading and WISC–R scores almost completely overlapped the variance common to parent and teacher reports of grades. The correlations between disruptive behavior and inattention were not much lower (.85 for girls and .60 for boys), nor were the correlations between inattention and grades (−.89 for girls and −.77 for boys). The correlations between disruptive behavior and ability were the lowest (−.27 for girls and −.20 for boys). Although all of the boys' correlations were lower than the girls', none of the differences was statistically significant.

It is important to be clear that the high correlations, particularly between ability and school grades, were between latent variables, not measured variables, and that, at least in girls, the correlations between inattention and disruptive behavior and between inattention and grades were of very similar magnitude. Thus, the high correlations between ability and grades were not unique to that pair of variables. The correlations among the relevant measured variables ranged from .45 to .55. These were certainly substantive correlations, but they in no sense would be considered to represent the same construct, and the measures on which they were based were in actuality very different; they were parents' and teachers' reports of grades and IQ and WRAT scores as administered in our labs. To check that these measures did differ, we added terms allowing for correlation of the residual variances between the IQ and WRAT scores and the grade reports to our phenotypic model. None of these terms was significant. Thus, the high latent correlations actually represent a finding of the study: When measurement error is removed, school grades are very highly related to more objective measures of ability, as well as to inattention, particularly in girls. The high latent correlations unquestionably had some effect on the parameter estimates—if the correlations were lower, the parameter estimates would no doubt also have been lower, but this is as it should be. The magnitudes of the latent correlations, however, did not completely drive the parameter estimates linking the latent ability and inattention variables to the latent grades variable: They were far short of 1.00 for both parameters, and the parameter estimate for the path from inattention

Table 3
Correlations and Covariances Among the Latent Phenotypic Variables

	Girls				Boys			
Variable	1	2	3	4	1	2	3	4
Correlations								
1. Inattention	—				—			
2. Ability	−.69	—			−.63	—		
3. Disruption	.85	−.27	—		.60	−.20	—	
4. Grades	−.89	.94	−.55	—	−.77	.89	−.40	—
Covariances								
1. Inattention	0.92				1.09			
2. Ability	−0.66	1.00			−0.66	1.01		
3. Disruption	0.69	−0.23	0.71		0.69	−0.23	1.32	
4. Grades	−0.83	0.92	−0.45	0.95	−0.83	0.92	−0.45	1.06

Note. Covariances were taken from the common metric completely standardized variance–covariance matrix, so they reflect the differences in latent variable variance between girls and boys.

to grades had a smaller loading relative to the loading for the path from ability to grades than would be indicated by the differences in their latent variable correlations alone.

Figure 3 shows the covariances among the latent inattention, ability, and disruptive behavior variables and the significant parameter estimates predicting the latent grades variable. The phenotypic model fit closely as indicated by the RMSEA. The figure shows the significant parameters for the model with the path coefficients for inattention and ability to grades, as well as the covariances between ability and inattention and between ability and disruptive behavior, constrained equally across the samples of girls and boys. This was possible without significant loss of fit; $\Delta\chi^2(5, N = 1596) = 9.10$, $p = .11$. Before the application of constraints across gender, the path from latent disruptive behavior to latent achievement was not significant for either boys or girls, as we had expected; $\Delta\chi^2(2, N = 1596) = 1.05$, $p = .59$.

The fact that the association between disruptive behavior and school grades could be explained by attention problems and lack of ability suggested that there should be children who displayed high levels of disruptive behavior but who also did well in school because disruptive behavior is not perfectly correlated with either inattention or ability problems. We would expect such children to be of at least normal ability and to manifest few attention problems, but in order to generate ideas about possible mechanisms explaining the resilience of their school performance, we thought it would be useful to identify such children and examine some of their characteristics. There were 7 girls in the sample who received excellent grades (rated as *much better than average* [*As*] on our assessment by their teachers and their parent) but scored in the highest 25% for disruptive behavior as reported across raters and ODD and CD categories. This meant either that they had several symptoms of ODD and/or CD, or that they scored in excess of 20 on the T-CD Scale, or both. These girls had WISC–R IQ scores averaging 111.7, or 0.6 standard deviations above the mean, with 3 in excess of 120. Two of the girls with the highest WISC–R scores read at the 12th-grade level, but 2 with high WISC–R scores read at the 5th-grade level. Overall, the group read one full grade level higher than average for the sample. Their level of engagement appeared to be typical of the 113 girls earning similarly high grades. Only 1 of these disruptively behaving girls (No. 7) appeared to have a significant inattention problem, and the group had fewer attention problems on average than did either the sample in total or the 106 girls in the highest 25% for disruptive behavior, but more than the 113 girls earning similarly high grades. In general, except for their disruptive behavior, they appeared to be rather similar to the high grade group, suggesting that high ability and engagement acted to protect them from the adverse effects of their disruptive behavior. These results are summarized in Table 4.

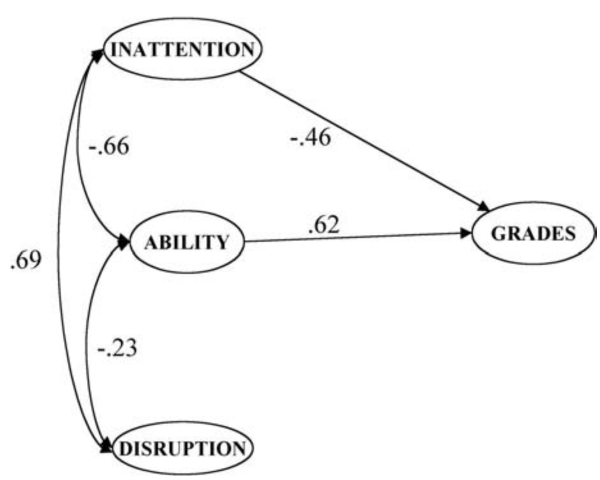

Figure 3. Parameter values from common metric phenotypic models for girls and boys. $\chi^2(62, N = 1596) = 95.19$, $p = .004$, root-mean-square error of approximation = .036. Nonsignificant paths are not shown.

There were 9 boys in the sample who received excellent grades according to all reporters (the same criteria existed as for girls) but scored in the highest 25% for disruptive behavior as reported across raters and ODD and CD categories (again, the same criteria existed as for girls). This meant either that they had several symptoms of ODD and/or CD, or that they scored in excess of 22 on the T-CD Scale, or both. These boys all had high WISC–R IQs—they ranged from 112 to 156, or 0.7 to 3.8 standard deviations above the mean, with an average of 131. Their average reading level was two grade levels above the average for the full sample, and the same as that for the 62 boys earning similarly high grades. They were somewhat less engaged than the high grade group, but more engaged than the 94 boys in the highest 25% for disruptive behavior and more engaged than the overall average. Again, only 1 of these boys (No. 2) appeared to have a significant inattention problem, and the group had fewer attention problems on average than did either the sample in total or the 94 boys in the highest 25% for disruptive behavior, but more attention problems than the 62 boys earning similarly high grades. There were two twin pairs in the group of 9 boys: one MZ and one DZ. In general, these boys were extremely bright and less engaged than others earning similarly high grades. Thus, it appeared that for boys, high ability conferred the resilience effect on their school performance. The results for boys are also summarized in Table 4.

Analyses of Genetic and Environmental Influences

Table 5 shows the intraclass MZ and DZ twin correlations for each measure used in our quantitative genetic model fitting, separately for girls and boys. Except for the teachers' report of ODD problems for girls, all of the MZ correlations were higher than the DZ correlations, suggesting genetic influence on all of the latent constructs the measures represent. At the same time, except for the parents' report of ADD symptoms, all of the DZ correlations were at least half the MZ correlations, suggesting shared environmental influences rather than nonadditive genetic influences. There was no evidence for differences in extent of genetic influences between girls and boys.

The indicated proportions of latent variable variance attributable to genetic and shared and nonshared environmental influences, based on the model shown in Figure 2, are shown in Table 6. It was possible to constrain the genetic, shared, and nonshared environmental variances equally, as well as the rater factor loadings (but not the residual variances) between girls and boys, $\Delta\chi^2(37, N = 798) = 21.73$, $p = .98$, so we present the estimates resulting from application of these constraints. All of the latent constructs showed a high level of genetic influence, about 70%–75%. Overall, something less than 20% of the variance could be attributed to shared environmental influences, leaving less than 10% to nonshared environmental influences. The 95% confidence intervals shown in

Table 4
Characteristics and Questionnaire Responses of Children With Disruptive Behavior Scores in the 75th Percentile or Above Who Also Earned High Grades (Reported as "A" or "Corresponding to A")

Participant	WISC IQ	WRAT reading	Attention problems	Engagement	Disruptive behavior	Grades
Girls						
1	123	12	9.3	14.9	15.0	4
2	108	6	9.3	15.7	16.7	4
3	97	5	9.5	14.9	23.0	4
4	114	12	8.0	14.3	15.5	4
5	120	5	8.0	13.1	15.7	4
6	97	5	8.0	15.1	23.7	4
7	123	5	13.5	15.4	31.0	4
A. Subgroup mean	111.7	7.1	9.4	14.8	20.1	4.0
B. High-grade group mean	115.5	7.3	8.4	14.8	15.9	4.0
C. Disruptive group mean	101.2	5.9	10.4	13.4	17.3	3.0
Difference between A and B	−0.27	−0.10	0.32	0.00	1.56	0.00
Difference between A and C	0.75	0.57	−0.32	0.99	1.04	2.17
Boys						
1	112		11.5	15.4	19.0	4
2	120	7	25.7	10.9	37.0	4
3	142	8	11.0	14.6	26.0	4
4[a]	156	8	13.0	15.4	22.5	4
5[a]	147	8	14.0	14.0	22.5	4
6	121		13.0	13.7	19.3	4
7[b]	131	8	18.0	11.7	24.0	4
8[b]	119	8	12.8	12.0	18.5	4
9	128		15.5	12.9	30.5	4
A. Subgroup mean	130.7	7.8	14.9	13.4	24.4	4.0
B. High-grade group mean	118.6	7.8	12.5	14.1	16.2	4.0
C. Disruptive group mean	103.6	5.6	20.7	12.3	22.2	2.7
Difference between A and B	0.88	0.00	0.48	−0.40	2.05	0.00
Difference between A and C	1.98	1.10	−1.16	0.63	0.55	1.49

Note. Attention problems and disruptive behavior were based on teachers' reports for illustration, although placement in the disruptive group reflected parent and child reports as well. Engagement was based on the child's report. The high-grade group refers to those for whom all reported grade were well above average (4s; girls, $n = 113$; boys, $n = 62$). The disruptive group refers to the highest 25% for disruptive behavior across reporters (girls, $n = 106$; boys, $n = 94$). Grades were averaged across reporters. Differences between A and B and between A and C were effect sizes; for example, (A − B)/overall *SD*, for the variable from Table 1. For engagement, the overall standard deviation was 1.41 for girls and 1.75 for boys. WISC = composite IQ from the Wechsler Intelligence Scale for Children—Revised (WISC–R) subtests used; WRAT = grade-equivalent reading level, based on the Wide Range Achievement Test.
[a] Monozygote twin pair. [b] Dizygote twin pair.

the table were relatively tight for such intervals, suggesting that we were able to make these estimates reasonably accurately.

Table 7 shows the significant genetic and environmental correlations resulting from the constrained model. The confidence intervals associated with genetic and environmental correlations are notoriously large (Carey & DiLalla, 1994), so most of the environmental correlations did not differ significantly from zero as a result of the much smaller proportions of variance associated with the environmental components of variance. The genetic correlations between ability and disruptive behavior did not differ significantly from zero. The genetic correlation between disruptive behavior and grades was significantly different from zero (−.37), but it was also significantly lower than that between inattention and grades (−.87).

Discussion

In this study, we addressed three questions involving the contemporaneous phenotypic, genetic, and environmental relationships among inattention, academic ability, disruptive behavior, and school grades by using a large population-based twin sample of 11-year-olds and comparing results for girls and boys. Although theory and empirical data have suggested that the long-observed association between disruptive behavior and later school achievement difficulties originates with the disruptive behavior, there is evidence that inattention and ability problems can account for the longitudinal association (Fergusson & Horwood, 1995; Fergusson et al., 1993; Frick et al., 1991). We have added to the evidence that the association between disruptive behavior and achievement difficulties is spurious by showing that inattention and ability problems can account for the association contemporaneously as well. In particular, at age 11, at the transition between elementary and secondary school, it is not disruptive behaviors but inattention and ability problems that are independently adversely associated with school grades, which measure perceived day-to-day academic performance. This provides both a strong test of the ability of inattention and ability problems to account for the association between

Table 5
Intraclass Twin Correlations

Measure	Girls MZ (n = 259)	Girls DZ (n = 165)	Boys MZ (n = 253)	Boys DZ (n = 121)
Inattention				
P-IATT	.82	.46	.81	.55
T-IATT	.70	.46	.73	.50
P-ADD	.65	.26	.65	.32
Ability				
WISC	.78	.53	.74	.56
WRAT	.71	.45	.66	.39
Disruptive behavior				
P-ODD	.91	.59	.83	.55
T-ODD	.66	.66	.79	.58
P-CD	.92	.55	.91	.72
T-CD	.59	.33	.74	.43
Grades				
P-GR	.88	.62	.88	.66
T-GR	.75	.46	.73	.48

Note. MZ = monozygotic; DZ = dizygotic; n = number of pairs; P-IATT = parent's report of attention problems; T-IATT = teachers' report of attention problems; P-ADD = parent's report of attention deficit disorder symptom count, based on the Diagnostic Interview for Children and Adolescents—Revised (DICA–R); WISC = composite IQ from the Wechsler Intelligence Scale for Children—Revised subtests used; WRAT = grade-equivalent reading level, based on the Wide Range Achievement Test; P-ODD = parent's report of oppositional and defiant disorder behavior, based on DICA–R symptom counts; T-ODD = teachers' report of oppositional and defiant disorder behavior, based on the Teacher's Rating Forms; P-CD = parent's report of conduct disorder behaviors, based on DICA-R symptom counts; T-CD = teachers' report of conduct disorder behavior, based on the Teacher's Rating Forms; P-GR = parent's report of child's grades; T-GR = teachers' report of child's grades.

disruptive behavior and achievement and evidence for the stability of these relationships over time.

Disruptive but Achieving Children

In reality, the same children tend to display both inattentive and disruptive behaviors, and low ability is also a common problem in this group. This means that it is very hard to distinguish separate etiologies for the three conditions, in part because, to a large degree, the same genetic influences contribute to all of them. The fact, however, that inattention and ability problems can explain the phenotypic correlation of −.55 in girls and −.40 in boys between disruptive behavior and school grades suggests that there should be children who display disruptive behavior and also do well in school. We were in fact able to identify a few such children. Studying these children and others like them may suggest ways to keep children who behave disruptively from complicating their lives further by achieving poorly in and/or dropping out of school.

Overall, the disruptive but achieving children were bright, particularly the boys, who were brighter than other boys earning comparable grades. A couple of the girls appeared to have lower reading levels than might have been expected given their WISC–R IQs. As would be expected, inattention was not a contributing factor for most, nor were ability problems. These children were generally more engaged in school than average, though the boys were less engaged on average than others earning comparable grades. We speculate that many of these children may be bored by a curriculum that does not offer sufficient challenge, particularly because overall their teachers' ratings of their disruptive behavior were especially high. If this is true, they may engage in disruptive behaviors primarily to keep themselves amused. This suggests that greater attention to engaging the interests particularly of bright boys may help to at least minimize disruptive behavior in the classroom. Another way to look at these disruptive but high-achieving boys is that they provide an unusual perspective on the commonly made observation that schools relatively poorly accommodate the kinds of behaviors displayed by boys when they are disengaged.

Genetic and Environmental Influences

At the level of the latent variables, all four of the constructs we explored appeared to be under strong genetic influence in the current environment in which all children are expected to attend school: Overall, 70%–75% of the variance could be attributed to genetic influence, and the estimates of variance attributable to genetic and environmental influence could be constrained equally for girls and boys. At 70%–75%, the percentage of variance attributable to genetic influence is high and may have resulted in part from the strong common genetic relationships between some of the pairs of traits. The genetic correlations suggested that common genetic influences contributed substantially to ability and school grades and to inattention and school grades. The genetic correlations between inattention and ability, inattention and disruptive behavior, and disruptive behavior and school grades were much more moderate, and the genetic correlation between ability and disruptive behavior was not significant. Both shared and nonshared environmental associations appeared to reinforce the genetic associations, as most were in the same direction and were larger in absolute value, though there was little power to estimate them with any precision. To the extent that this observation

Table 6
Indicated Proportions of Latent Variable Variance

Latent variable	A Proportion	A 95% CI	C Proportion	C 95% CI	E Proportion	E 95% CI
Inattention	.75	.54–.88	.11	.02–.26	.14	.08–.21
Ability	.67	.45–.90	.26	.04–.47	.07	.03–.11
Disruptive behavior	.76	.58–.97	.23	.01–.42	.01	.00–.02
Grades	.75	.58–.91	.19	.01–.38	.06	.03–.10

Note. The variances could be constrained equally for boys and girls. A = genetic influences; C = shared environmental influences; E = nonshared environmental influences; CI = confidence interval.

Table 7
Statistically Significant Genetic, Shared Environmental, and Nonshared Environmental Correlations

	Genetic			Shared environmental			Nonshared environmental		
Variable	Ability	Disruption	Grades	Ability	Disruption	Grades	Ability	Disruption	Grades
Inattention	−.54	.50	−.87	—	—	—	−.70	—	−.79
Ability		—	.82		—	.99		—	.99
Disruption			−.37			—			—

Note. Because variances were constrained equal for girls and boys, these correlations were also equal for girls and boys. Dashes indicate nonsignificant correlations. For significant genetic correlations, 95% confidence intervals ranged from .10 to .20 around each point estimate. For significant environmental correlations, 95% confidence intervals ranged from .35 less than the point estimate in absolute value to 1 in absolute value.

reflects reality, it suggests that the school environment as presently structured acts to magnify rather than minimize individual differences associated with school performance. That is, the environment tends to reinforce the expression of genetic vulnerabilities where they exist and to suppress the expression of genetic strengths where they exist. This information could be used to explore the possibility of changes in the school environment that might reverse these effects.

It is important to keep in mind that the presence of large proportions of variance that can be attributed to genetic influence does not mean that outcomes are fixed. Outcomes always result from transactions between genes and environments, and different environments can constrain or enhance genetic expression (Rutherford, 2000). In addition, individuals differing genetically may seek different environments (Kendler & Eaves, 1986) to the extent that they can. Thus, environmental interventions targeted at individuals with particular genetic characteristics can result in major changes in both outcomes and allocations of variance. Although not assessed directly in this study, the pattern of genetic and environmental correlations suggests that correlations between genetic and environmental influences may underlie the relationships among these variables. Such correlations could be thought of as genetic control of exposure to different environments and might include children of high ability who seek out extra credit class work as well as children with attention problems who seek opportunities to engage in disruptive activities to avoid sitting down to complete their assignments. The correlations in the data presented here suggest that one possibly productive avenue for improving the school performance of many children may be for the school environment to offer more ways besides sitting still and concentrating for children to demonstrate mastery of an academic skill.

Gender Differences

Our data contain some interesting relationships among reporters for effect size differences between girls and boys on the measures of inattention and conduct problems. Specifically, although parents' reports show moderate effect size differences reflecting greater disruptive behavior and inattention in boys than in girls, the teachers' reports show a large (.70) effect size difference reflecting greater inattention in boys than in girls but a much smaller effect size difference (.18 and .30, respectively) for ODD and CD problems. Several reports (e.g., Abikoff, Courtney, Pelham, & Koplewicz, 1993; Stevens, Quittner, & Abikoff, 1998) have suggested that teachers' ratings of inattentive behaviors are often inflated in the presence of conduct problems (though their ratings of conduct problems are not inflated in the presence of inattention), but a recent report based on more objective measurements suggested that the inattentive behaviors are in fact more common in children with conduct problems (Abikoff et al., 2002) in spite of whatever reporting biases may exist. These factors could help to explain our pattern of results, but it is also possible that the teachers were accurately observing an increased level of inattention in boys not stimulated by the school environment. We do not have a good potential explanation for the relatively small effect size difference for the teachers' reports of conduct problems, nor for the reversal of the gender difference for ODD between parents and teachers (parents observed girls to be higher, but teachers observed boys to be higher).

In spite of the substantial effect sizes associated with the mean gender differences, we were unable to uncover any gender differences in the relationships among the variables, nor in the genetic and environmental influences underlying them. There were, however, differences in the phenotypic variances and covariances of the latent variables (see Table 3), reflecting the variance differences in the measured variables contributing to them (see Table 1). These variance differences were captured in the residuals and rater factors in the quantitative genetic model. This suggests that the variance that can be attributable to systematic genetic and shared and nonshared environmental influences is the same in girls and boys but that there is more variance in the behavior that falls outside this model in boys than in girls. This may be because boys' behavior varies more from situation to situation than does girls' (e.g., school vs. home), because people's (e.g., parents and teachers) perceptions of or expectations for boys' behavior varies more, and/or because there is more error variance in our measurements of boys' behavior than of girls'.

Conclusion

As the MTFS is a longitudinal study, we have the opportunity to observe the ways in which the relationships observed in this study evolve as the participants develop. This is particularly important because at age 11, the age of the children in this study, relatively little truly serious disruptive behavior has emerged. In addition, the school curriculum is not generally highly academically differentiated, and grading systems tend to reflect individual progress and effort to a greater degree than they do later (Eccles et al., 1984). Thus, the associations we have observed here may change significantly as the children mature. Such changes, along with the identification of developmental achievement trajectories, will be the subject of future studies as our reassessment of study participants at age 14, which is still underway, is completed. In addition, the MTFS includes data on the twins' relationships with their parents and peers, as well as data on parenting practices. We intend to explore the influences of these factors on the associations we have identified here as well.

This study is subject to several methodological limitations that should be considered when evaluating the significance and generalizability of our results. First, our assessment of behaviors and school grades is based on parents' and teachers' reports rather than on direct observation. Second, our sample is predominantly Caucasian, representative of twin births in Minnesota for the birth years in question. This is an advantage in the sense that it may help to clarify associations within this group, but the generalizability of our findings to other ethnic groups and other time periods needs to be addressed. Finally, our contemporaneous analysis does not allow us to consider school grades to be the outcome of the other three constructs empirically. The causal relationships underlying the associations we have observed can be better evaluated by using the longitudinal data that will become available in this study. As noted above, we intend to do this in future studies. In spite of these limitations, however, the study clearly establishes in a population sample that disruptive behavior has no direct association with school grades at age 11. Rather, the apparent association appears to result from common associations of both with inattention and ability problems. This study also establishes that this pattern of associations exists at the level of genetic influences and is apparently reinforced at the levels of both shared and nonshared environmental influences. This suggests that environmental interventions targeted to suppress the expression of each individual's genetic vulnerabilities while reinforcing the expression of her or his genetic strengths could produce effective results.

References

Abikoff, H., Courtney, M., Pelham, W. E., & Koplewicz, H. S. (1993). Teachers' ratings of disruptive behaviors: The influence of halo effects. *Journal of Abnormal Child Psychology, 21,* 519–533.

Abikoff, H. B., Jensen, P. S., Arnold, L. L., Hoza, B., Hechtman, L., Pollack, S., et al. (2002). Observed classroom behavior of children with ADHD: Relationship to gender and comorbidity. *Journal of Abnormal Child Psychology, 30,* 349–359.

Alarcon, M., Plomin, R., Fulker, D. W., Corley, R., & DeFries, J. C. (1998). Multivariate path analysis of specific cognitive abilities data at 12 years of age in the Colorado Adoption Project. *Behavior Genetics, 28,* 255–264.

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., rev.). Washington, DC: Author.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Barkley, R. A., Fischer, M., Smallish, L., & Fletcher, K. (2002). The persistence of attention-deficit/hyperactivity disorder into young adulthood as a function of reporting source and definition of disorder. *Journal of Abnormal Psychology, 111,* 279–289.

Beutler, L. E., Moleiro, C., Malik, M., Harwood, T. M., Romanelli, R., Gallagher-Thompson, D., & Thompson, L. (2003). A comparison of the dodo, EST, and ATI factors among comorbid stimulus-dependent, depressed patients. *Clinical Psychology and Psychotherapy, 10,* 69–85.

Bouchard, T. J., & McGue, M. (1981, May 29). Familial studies of intelligence: A review. *Science, 212,* 1055–1059.

Browne, M., & Cudeck, R. (1992). Alternative methods of assessing model fit. *Sociological Methods & Research, 21,* 230–258.

Burt, S. A., Krueger, R. F., McGue, M., & Iacono, W. G. (2003). Parent–child conflict and the comorbidity among childhood externalizing disorders. *Archives of General Psychiatry, 60,* 505–513.

Butts, J. A., Snyder, H. N., Finnegan, T. A., Aughenbaugh, A. L., Tierney, N. J., Sullivan, D. P., & Poole, R. S. (1995). *Juvenile court statistics: 1992*. Washington, DC: Office of Juvenile Justice and Delinquency Prevention.

Carey, G., & DiLalla, D. L. (1994). Personality and psychopathology: Genetic perspectives. *Journal of Abnormal Psychology, 103,* 32–43.

Clark, C., Prior, M., & Kinsella, G. (2002). The relationship between executive function abilities, adaptive behaviour, and academic achievement in children with externalizing behaviour problems. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 43,* 785–796.

Cohen, P., Cohen, J., Kasen, S., Velez, C. N., Hartmark, C., Johnson, J., et al. (1993). An epidemiological study of disorders in late childhood and adolescence: I. Age and gender-specific prevalence. *Journal of Child Psychology and Psychiatry, 34,* 851–867.

Conduct Problems Prevention Research Group. (1992). A developmental and clinical model for the prevention of conduct disorder: The FAST Track program. *Development and Psychopathology, 4,* 509–527.

Connors, C. K. (1969). A teacher rating scale for use in drug studies with children. *American Journal of Psychiatry, 126,* 884–888.

Eccles, J., Midgley, C., & Adler, T. (1984). Grade-related changes in the school environment: Effects on achievement motivation. In J. Nicholls (Ed.), *Advances in motivation and achievement* (Vol. 3, pp. 283–331). Greenwich, CT: JAI Press.

Elliott, D. S., Huizinga, D., & Ageton, S. S. (1985). *Explaining delinquency and drug use*. Beverly Hills, CA: Sage.

Farrington, D. P., Loeber, R., Elliott, D. S., Hawkins, J. D., Kandel, D. B., Klein, M. W., et al. (1990). Advancing knowledge about the onset of delinquency and crime. In B. B. Lahey & A. E. Kazdin (Eds.), *Advances in clinical child psychology* (Vol. 13, pp. 383–442). New York: Plenum Press.

Fergusson, D. M., & Horwood, L. J. (1995). Early disruptive behavior, IQ, and later school achievement and delinquent behavior. *Journal of Abnormal Child Psychology, 23,* 183–199.

Fergusson, D. M., Horwood, L. J., & Lynsky, M. T. (1993). The effects of conduct disorder and attention deficit in middle childhood on offending and scholastic ability at age 13. *Journal of Child Psychology and Psychiatry, 34,* 899–916.

Frick, P. J., Kamphaus, R. W., Lahey, B. B., Loeber, R., Christ, M. A. G., Hart, E. L., & Tannenbaum, L. E. (1991). Academic underachievement and the disruptive behavior disorders. *Journal of Consulting and Clinical Psychology, 59,* 289–294.

Gomez, R., Harvey, J., Quick, C., Sharer, I., & Harris, G. (1999). *DSM–IV* AD/HD: Confirmatory factor models, prevalence, and gender and age differences based on parent and teacher ratings of Australian primary school children. *Journal of Child Psychology and Psychiatry, 40,* 265–274.

Grigorenko, E. L. (2001). Developmental dyslexia: An update on genes, brains, and environments. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 42,* 91–125.

Hawkins, J. D., & Lishner, D. M. (1987). Schooling and delinquency. In E. H. Johnson (Ed.), *Handbook of crime and delinquency prevention* (pp. 179–221). New York: Guilford Press.

Hinshaw, S. P. (1992a). Academic achievement, attention deficits, and aggression: Comorbidity and implications for intervention. *Journal of Consulting and Clinical Psychology, 60,* 893–903.

Hinshaw, S. P. (1992b). Externalizing behavior problems and academic underachievement in childhood and adolescence: Causal relationships and underlying mechanisms. *Psychological Bulletin, 111,* 127–155.

Horn, J. L., & McArdle, J. J. (1992). A practical and theoretical guide to measurement invariance in aging research. *Experimental Aging Research, 18,* 117–144.

Iacono, W. G., Carlson, S. R., Taylor, J., Elkins, I. J., & McGue, M. (1999). Behavioral disinhibition and the development of substance-use disorders: Findings from the Minnesota Twin Family Study. *Development and Psychopathology, 11,* 869–900.

Jacobson, K. C., Prescott, C. A., & Kendler, K. S. (2000). Genetic and environmental influences on juvenile antisocial behavior assessed on two occasions. *Psychological Medicine, 30,* 1315–1325.

Johnson, W., Bouchard, T. J., Jr., Segal, N. L., & Samuels, J. (2005). General intelligence and reading performance in adults: Is the genetic factor structure the same as for children? *Personality and Individual Differences, 38,* 1413–1428.

Jöreskog, K. G., & Sörbom, D. (2003). LISREL (Version 8.53) [Computer software]. Chicago: Scientific Software International.

Kagan, J. (1999). The role of parents in children's psychological development. *Pediatrics, 104,* 164–167.

Kazdin, A. E. (1993). Treatment of conduct disorder: Progress and directions in psychotherapy research. *Development and Psychopathology, 5,* 277–310.

Kendler, K. S., & Eaves, L. (1986). Models for the joint effect of genotype and environment on liability to psychiatric illness. *American Journal of Psychiatry, 143,* 279–289.

Kimball, M. M. (1989). A new perspective on women's math achievement. *Psychological Bulletin, 105,* 198–214.

Lahey, B. B., Miller, T. L., Schwab-Stone, M., Goodman, S. H., Waldman, I. D., Canino, G., et. al. (2000). Age and gender differences in oppositional behavior and conduct problems: A cross-sectional household study of middle childhood and adolescence. *Journal of Abnormal Psychology, 109,* 488–503.

Little, R. J. A., & Rubin, D. B. (1987). *Statistical analysis with missing data*. New York: Wiley.

Loeber, R. (1991). Antisocial behavior: More enduring than changeable? *Journal of the American Academy of Child and Adolescent Psychiatry, 30,* 393–397.

Lynam, D. R., Caspi, A., Moffitt, T. E., Wikstrom, P. H., Loeber, R., & Novak, S. (2000). The interaction between impulsivity and neighborhood context in offending: The effects of impulsivity are stronger in poorer neighborhoods. *Journal of Abnormal Psychology, 109,* 563–574.

Lyons, M. J., True, W. R., Eisen, S. A., Goldberg, J., Meyer, J. M., Faraone, S. V., et al. (1995). Differential heritability of adult and juvenile antisocial traits. *Archives of General Psychiatry, 52,* 906–915.

Mau, W., & Lynn, R. (2001). Gender differences on the Scholastic Aptitude Test, the American College Test, and college grades. *Educational Psychology, 21,* 133–136.

McCall, R. B., Evahn, C., & Kratzer, L. (1992). *High school underachievers: What do they achieve as adults* (Vol. 1). Newbury Park: Sage.

McCartney, K., Harris, M. J., & Barnieri, F. (1990). Growing up and growing apart: A developmental meta-analysis of twin studies. *Psychological Bulletin, 107,* 226–237.

McGue, M., Bouchard, T. J., Iacono, W. G., & Lykken, D. T. (1993). Behavioral genetics of cognitive ability: A life-span perspective. In R. Plomin & G. E. McClearn (Eds.), *Nature, nurture, and psychology*. Washington, DC: American Psychological Association.

McGue, M., Wette, R., & Rao, D. C. (1984). Evaluation of path analysis through computer simulation: Effect of incorrectly assuming independent distribution of familial correlations. *Genetic Epidemiology, 1,* 255–269.

Meredith, W. M. (1993). Measurement invariance, factor analysis, and factor invariance. *Psychometrika, 58,* 525–543.

Moffitt, T. E. (1993). Adolescence-limited and life-course persistent antisocial behavior: A developmental taxonomy. *Psychological Review, 100,* 674–701.

Moffitt, T. E., Caspi, A., & Rutter, M. (2005). Strategy for investigating interactions between measured genes and measured environments. *Archives of General Psychiatry, 62,* 473–481.

Moffitt, T. E., & Silva, P. (1988). Self-reported delinquency, neuropsychological deficit, and history of attention deficit disorder: Boys' developmental trajectories from age 3 to age 15. *Child Development, 61,* 893–910.

Nadder, T. S., Rutter, M., Silberg, J. L., Maes, H. H., & Eaves, L. J. (2002). Genetic effects on the variation and covariation of attention deficit-hyperactivity disorder (ADHD) and oppositional-defiant disorder/conduct disorder (Odd/CD) symptomatologies across informant and occasion of measurement. *Psychological Medicine, 32,* 39–53.

Neale, M. C. (1997). *Mx: Statistical modeling*. Richmond, VA: Medical College of Virginia, Department of Psychiatry.

Nelson, J. R., Benner, G. J., Lane, K., & Smith, B. W. (2004). Academic achievement of K–12 students with emotional and behavioral disorders. *Exceptional Children, 71,* 59–73.

Patterson, G. R. (1986). Performance models for antisocial boys. *American Psychologist, 41,* 432–444.

Patterson, G. R., Baryshe, B. D., & Ramsey, E. (1989). A developmental perspective on antisocial behavior. *American Psychologist, 44,* 329–335.

Patterson, G. R., Reid, J. B., & Dishion, T. J. (1992). *Antisocial boys*. Eugene, OR: Castalia.

Petrill, S. A., & Thompson, L. A. (1993). The phenotypic and genetic relationships among measures of cognitive ability, temperament, and scholastic achievement. *Behavior Genetics, 23,* 511–518.

Plomin, R., Fulker, D. W., Corley, R., & DeFries, J. C. (1997). Nature, nurture, and cognitive development from 1 to 16 years: A parent–offspring adoption study. *Psychological Science, 8,* 442–447.

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology, 25,* 111–163.

Rhee, S. H., Waldman, I. D., Hay, D. A., & Levy, F. (2001). Aetiology of the sex difference in the prevalence of *DSM–III–R* ADHD: A comparison of two models. In F. Levy & D. A. Hay (Eds.), *Attention, genes, and attention deficit hyperactivity disorder* (pp. 139–156). Philadelphia: Psychology Press.

Rowe, D. C. (1983). Biometrical genetic models of self-reported delinquent behavior: A twin study. *Behavior Genetics, 13,* 473–489.

Rutherford, S. L. (2000). From genotype to phenotype: Buffering mechanisms and the storage of genetic information. *BioEssays, 22,* 1095–1105.

Rutter, M. (1967). Classification and categorization in child psychiatry. *International Journal of Psychiatry, 3,* 161–187.

Rutter, M., Graham, P., Chadwick, O. F., & Yule, W. (1976). Adolescent turmoil: Fact or fiction? *Journal of Child Psychology and Psychiatry and Allied Disciplines, 17,* 35–56.

Skodak, M., & Skeels, H. M. (1949). A final follow-up of 100 adopted children. *The Journal of Genetic Psychology, 75,* 85–125.

Snow, R. E. (1991). Aptitude–treatment interaction as a framework for research on individual differences in psychology. *Journal of Consulting and Clinical Psychology, 59,* 205–216.

Sternberg, R. J., Grigorenko, E. L., Ferrari, M., & Clinkenbeard, P. (1999). A triarchic analysis of an aptitude–treatment interaction. *European Journal of Psychological Assessment, 15,* 3–13.

Stevens, J., Quittner, A. L., & Abikoff, H. (1998). Factors influencing elementary school teachers' ratings of ADHD and ODD. *Journal of Clinical Child Psychology, 27,* 406–414.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics*. Boston: Allyn & Bacon.

Taylor, J., Iacono, W. G., & McGue, M. (2000). Evidence for a genetic etiology of early-onset delinquency. *Journal of Abnormal Psychology, 109,* 634–643.

Thapar, A. (2003). Attention deficit hyperactivity disorder: New genetic findings, new directions. In R. Plomin & J. C. DeFries (Eds.), *Behavioral genetics in the postgenomic era* (pp. 445–462). Washington, DC: American Psychological Association.

Thompson, L. A., Detterman, D. K., & Plomin, R. (1991). Associations between cognitive abilities and scholastic achievement: Genetic overlap but environmental difference. *Psychological Science, 2,* 158–165.

Welner, Z., Reich, W., Herjanic, B., Jung, K. G., & Amada, H. (1987). Reliability, validity, and parent–child agreement studies of the Diagnostic Interview for Children and Adolescents (DICA). *Journal of the American Academy of Child and Adolescent Psychiatry, 26,* 649–653.

Received April 20, 2004
Revision received February 21, 2005
Accepted April 4, 2005 ■