

VERY PRELIMINARY – DO NOT CITE OR QUOTE

**Sibling, Peer and Schoolmate Correlations as Indicators of the Importance of
Context for Adolescent Development**

Greg J. Duncan
Northwestern University

Kathleen Mullan Harris
University of North Carolina at Chapel Hill

Johanne Boisjoly
University of Quebec at Rimouski

February 23, 1998

*Corresponding author: Greg J. Duncan, Institute for Policy Research, Northwestern University, 2040 Sheridan Road, Evanston, IL 60208, e-mail: greg-duncan@nwu.edu. Prepared for presentation at the biennial meetings of the Society for Research on Adolescence, San Diego, February 28, 1998. We are grateful to the Family and Child Well-being Research Network of the National Institute of Child Health and Human Development for supporting Duncan with this research. We also acknowledge support to Harris from the Add Health project, a program project designed by J. Richard Udry (PI) and Peter Bearman, and funded by grant P01-HD31921 from the National Institute of Child Health and Human Development to the Carolina Population Center, University of North Carolina at Chapel Hill, with cooperative funding participation of 17 other federal agencies. We thank Jan Hendrickson-Smith of the Carolina Population Center for her dedicated and superb assistance with the data analysis for this paper.

Sibling, Peer and Schoolmate Correlations As Indicators of the Importance of Context for Adolescent Development

Abstract

We use nationally representative data on behavior and attitudes of adolescents from the National Longitudinal Study of Adolescent Health to calculate correlations between genetically differentiated siblings within a family, between grademates within a school, and between peers as defined by the adolescent's "best friend" nominations. We argue that these correlations provide upper-bound estimates of the respective potential influence of family, school and peer context. We find the largest correlations among siblings, especially identical twins. Consistent with the literature on school effects, our grademate correlations were generally quite small, although tended to be somewhat larger for boys than girls. While somewhat larger, the peer-based correlations were also surprisingly small, suggesting a smaller scope for causal peer influences on achievement and, especially, risk behavior than might be expected.

Sibling, Peer and Schoolmate Correlations As Indicators of the Importance of Context for Adolescent Development

I. Introduction

Developmentalists have long recognized the potential influence of the many contexts in which adolescents develop (Bronfenbrenner, 1979), particularly the family, neighborhood, school and peers. Comprehensive assessments of the influence of any one of these contexts are difficult and therefore rare; rarer still are attempts to assess the relative importance on developmental outcomes of these various contexts and possible interactions among them.

An overriding problem in global assessments of contextual effects is that of measurement; the plethora of ways in which a context such as the family, a school or peers may influence development requires a dauntingly extensive set of measurements to capture all of the potentially important influences. Investigating family influences is particularly problematic given the potential role of as-yet-unmeasurable genetic factors (Plomin et al., 1990).

This paper adopts a different approach to the task of global assessment of the importance of contextual effects. We use correlations between siblings within a family, between grademates within a school, and between peers as defined by a set of “best friend” nominations as respective upper bound estimates of the potential influence of family, school and peers. Behavioral geneticists have used correlations in this way for nearly a century in studying family influences (Plomin et al., 1990), as have sociologists for the last quarter century when examining school effects (e.g., Jencks and Brown, 1975; Coleman, 1966) and peer effects (e.g., Kandel, 1978; Billy and Udry, 1985); in recent years economists have also used correlations in this way in analyzing neighborhood influences (Solon et al., 1997). We apply correlational methods to a number of important adolescent outcomes using data from a single, very large and nationally-representative data set – the National Longitudinal Study of Adolescent Health – to provide a comparative look at upper bound estimates of the relative strength of contextual influences.

II. Background

When modeling the effects of context on adolescent development, there are good reasons to prefer direct measurement of the potentially consequential elements of context. Most important, direct measures of theoretically important features of context provide a basis for assessing the theoretical processes involving those features. But even exhaustive inventories of questionnaire items, or, in the case of some dimensions of neighborhood and schooling contexts, administrative data barely begin to capture the dimensions of context that might matter. Incomplete measurement can lead both to an understatement of the global impact of a context and to bias in assessing the role of measured components.

An alternative approach to assessing the strength of contextual effects relies on either the correlations between children who are siblings, neighbors, classmates or peers or, equivalently, on the fraction of variation in children's outcomes of interest that lies between families, neighborhoods, schools or peer groups to provide an upper bound on the possible effect of these contexts.

Behavioral geneticists use the pattern of correlations among siblings and parent-child pairs with varying degrees of genetic relatedness and co-residence during childhood to estimate the role of genes and shared and unshared environmental influences (Falconer, 1981). Roughly speaking, a personality trait with a mixture of purely genetic and unshared environmental causes should produce outcome correlations twice as high in monozygotic (MZ, i.e., one-egg) twins as in either dizygotic (DZ, i.e., two-egg) twins or siblings born at different times, since MZ twins have 100% genetic relatedness while DZ and other full siblings share only 50% genetic relatedness. Under these assumptions, the extent to which the MZ correlation is less than perfect reflects the importance of a combination of unshared environmental causes and measurement error.

Allowing for the potential importance of the environments shared by twin and nontwin siblings (e.g., unchanging parenting practices, permanent family resources and, for twins, in utero conditions) complicates inferences from these correlations. If genes were unimportant and environments similar then one would expect similar MZ, DZ and non-twin sibling correlations. Simple behavioral genetics models suggest that the

heritability of a trait equals twice the difference between MZ and DZ correlations and that the role of shared environment can be expressed as twice the DZ correlation minus the MZ correlation.¹

Twin samples are much rarer than more general samples of siblings, since the latter are often an unintended byproduct of procedures adopted by large national surveys. Ignoring issues of decomposition of nature vs. nurture, a host of studies (e.g., Jencks et al., 1977; Griliches, 1979) have used non-twin sibling correlations to estimate the combined importance of the genes and family, neighborhood, school and other environmental experiences shared by siblings. For example, non-twin sibling correlations for years of completed schooling are quite high – around .55 – indicating that there are important elements of the families (including genetic influences), neighborhoods, schools and other aspects of the shared environments of siblings that affect achievement.

The beauty of sibling correlations is that they provide an upper bound on the estimated impact of both measurable and unmeasurable aspects of the environments shared by siblings. Of course, the problem with sibling correlations is that they reveal nothing about the process by which familial and extra-familial influences operate to make siblings more alike than two individuals drawn at random from the population.

Solon et al. (1997) argue that an analogous correlation for children not in the same family but attending the same school or growing up in the same neighborhood indicates the potential importance of the shared environments of classmates or neighbors that lie outside the immediate family. Correlations on achievement or behavioral outcomes between unrelated neighbors or classmates that are similar to those of siblings are suggestive of a strong extra-familial component to the sibling correlations.² In contrast, high sibling correlations coupled with low neighbor and classmate correlations suggest that family effects are driving the sibling correlations and that the scope for pure (i.e., extra-familial) neighborhood or school effects is quite small.

Solon et al. (1997) formalize these observations in the context of an additive model of the effect of family and extra-familial context on some outcome of interest (y):

$$y_{cfs} = A'X_{cf} + B'Z_c + e_{cfs}.$$

Here c references the extra-familial context (e.g., school class, neighborhood, peer group), f references families and s references siblings; X_{cf} is a vector of shared family influences for all siblings within the same family; Z_c is a vector of extra-familial contextual influences (henceforth called “neighborhood” but applicable to other contexts) shared by all siblings and neighboring children; and e_{cfs} is an error term.

They show that the sibling covariance in y_{cfs} can be expressed as:

$$\text{Cov}(y_{cfs}, y_{cfs'}) = \text{Var}(A'X_{cf}) + \text{Var}(B'Z_c) + 2\text{Cov}(A'X_{cf}, B'Z_c),$$

i.e., the sum of shared family variance, shared neighborhood variance and twice the covariance between family and neighborhood factors.

The covariance between neighboring children (or classmates or peers) from different families is:

$$\text{Cov}(y_{cfs}, y_{cfs'}) = \text{Var}(B'Z_c) + 2\text{Cov}(A'X_{cf}, B'Z_c) + 2\text{Cov}(A'X_{cf}, A'X_{cf'}),$$

i.e., the sum of the shared neighborhood variance and twice the covariance between family and neighborhood factors plus twice the covariance in family backgrounds among neighboring children. In comparing sibling and neighbor covariances, it can be seen that shared neighborhood variance and covariance between family and neighborhood factors are common to both. The shared family variance is obviously missing from the non-family neighbor covariance, while the family background covariance of neighboring children is missing from the sibling covariance but is a part of the non-family neighbor covariance.

Accordingly, a neighbor correlation should be viewed as an upper bound on the possible causal role played by neighborhood factors. It is biased upwards because the third, shared family-background component is not a true neighborhood effect. Furthermore, neighbor correlations attribute all of the covariance between family and neighborhood factors to neighborhood effects, which is overly generous. Solon et al. (1997) attempt to deal with the spurious shared-family component by using outcomes defined as residuals of an auxiliary regression of outcomes on observed family characteristics.

National surveys such as the Panel Study of Income Dynamics (PSID) and National Longitudinal Survey of Youth draw their child samples from a set of families residing in a tightly clustered neighborhood area. Since these clusters approximate neighborhood areas, it is possible to calculate both sibling and neighbor clusters for various outcomes of interest. Solon et al. (1997) calculate such sibling and neighbor correlations with a PSID sample consisting of individuals age 8-16 in 1968. For their outcome measure – years of completed schooling - the sibling correlation (.54) is much higher than the estimated neighbor correlation (.19), suggesting a quite limited scope for the effects of extra-familial contexts.

After removing effects of easily observed socioeconomic characteristics of families (race, family income, family structure and maternal schooling) shared by children living in the same neighborhood, the neighbor correlation dropped further (to .10), suggesting an even more limited scope for unique neighborhood influences. All in all, the results suggest that neighborhood influences are much more limited than family influences in accounting for inter-individual variation in outcomes. Our paper extends the work described above by presenting correlations among schoolmates and peers, as well as breaking down sibling correlations according to genetic relatedness.

There are a number of important qualifications for using correlations among classmates, peers and neighbors as upper-bound estimates of the importance of extra-familial contexts. First, the transitory nature of neighborhoods, schools, and, especially, peer groups (Urberg, et al., 1995) in the lives of children may impart a measurement-error-induced downward bias to our correlations. For example, since residential mobility is quite common in the United States, especially among younger children, children sharing a neighborhood at any given point may have quite different residential histories. However, this bias may not be large since residential moves typically occur between similar neighborhoods.³

Second, the self-selected nature of context, particularly peer groups, will cause the correlations we calculate to overstate, perhaps dramatically, the causal role played by peers (Manski, 1993). In the extreme case, suppose that an adolescent befriends individuals with similar family background, personality and behavior but who play no

causal role in affecting his or her own behavior and achievement.⁴ Peer-based correlations of the kind we calculate would be high but reflect no true causal effect of peer-based contextual factors. Selection factors are present but presumably less important in parental choices of schools and neighborhoods than in an adolescent's choice of friends.

Third, the differential reliability of our various outcome measures will impart correspondingly differential bias to our correlations. We speculate on the nature of this bias when we discuss our results.

Fourth, we advise caution in the interpretation of small contextual correlations (Duncan and Raudenbush, 1998), since they are not inconsistent with effective policy interventions. Effect sizes that program evaluators commonly view as medium or even large translate into small proportions of variance in individual outcomes "explained" by context and into small intra-context correlations (Cain and Watts, 1972; Rosenthal and Rubin, 1982). Consider standardized effect sizes in a school treatment -- that is, standardized mean differences between a set of experimental schools and an equal number of control schools, commonly viewed as small ($d=.2$ of a standard deviation), medium ($d=.4$), large ($d=.6$) or very large ($d=.8$). These effect sizes would give rise to respective intra-school correlations of .01, .04, .08, and .12, respectively.

Thus, even very large effect sizes translate into intra-cluster correlations generally regarded as small. Accordingly, a small correlation between classmates, neighbors or peers does not rule out a large effect size associated with a measured difference between these contexts.

III. Data

Our data come from the National Longitudinal Study of Adolescent Health (Add Health). Add Health is a nationally representative study of adolescents in grades 7 through 12 in the U.S. in 1995, virtually all of whom responded to an inschool, self-administered questionnaire, and a systematic-chosen subset of whom responded to two waves of at-home personal interviews. At the present time, our analysis focuses on data

from the in-school survey. For obvious reasons, we intend to extend our work to include data from the two in-home interviews.

Add Health was designed to help explain the causes of adolescent health and health behavior with special emphasis on the effects of multiple contexts of adolescent life. The study used a school-based design to select a stratified sample of 80 high schools with probability proportional to size. More than 70 percent of the sampled schools were recruited for participation in the study. For each high school, a feeder school was also selected with probability proportional to its student contribution to the high school. The school-based sample therefore has a pair of schools in each of 80 communities.⁵

An in-school questionnaire was administered to every student who attended each selected school on a particular day during the period of September 1994 to April 1995. The in-school questionnaire was completed by more than 90,000 adolescents, and was intended to measure school-context variables; to obtain friendship networks; to measure a variety of health conditions; and to obtain data to make it possible to select special samples of individuals in rare but theoretically crucial categories.

In a second level of sampling adolescents and parents were selected for in-home interviews. The school rosters of the sampled schools were used as the sampling frame for the in-home sample. From the school rosters, a random sample of some 200 students from each school pair was selected, irrespective of school size, to produce the core in-home sample of about 12,000 adolescents. This creates a self-weighting sample because whether or not the student responded to the in-school survey does not affect his or her inclusion in the main in-home sample. A number of special over-samples were also selected for in-home interviews using screeners from the in-school questionnaires. With varying probabilities, the following special samples were drawn: physically disabled adolescents; black adolescents from highly educated families; several ethnic samples (Cuban, Puerto Rican, and Chinese adolescents); saturated samples in 14 schools; and a genetic sample described below.

The in-home interviews were conducted between April and December 1995, yielding the Wave I data. The core sample plus the special samples produced a sample size of 20,745 adolescents in Wave I. A parent, generally the mother, was also

interviewed in Wave I. The one- to two-hour adolescent interview was interviewer assisted using a laptop computer and audio-CASI for the more sensitive questions. Topics covered in the in-home interview include health status, nutrition, peer networks, decision-making processes, family composition and dynamics, educational aspirations and expectations, employment experience, and involvement in risky behavior, such as sexual activity, substance use, and criminal activities. Adolescents were given the Add Health Picture Vocabulary Test (AHPVT), a computerized, abridged version of the Peabody Picture Vocabulary Test-Revised, at the beginning of the in-home interview.

All adolescents in grades 7 through 11 in Wave I (plus 12th graders who were part of the genetic sample) were followed up one year later for the Wave II in-home interview in 1996. The Wave II interview was generally similar to Wave I. In addition, contextual data on the neighborhoods and communities in which adolescent respondents live were assembled using extant sources and merged with the Wave I and Wave II in-home data. Bearman, Jones, and Udry (1997) provide a more detailed description of the Add Health study.

The genetic over-sample is one feature of Add Health that we exploit in this paper. Based on the information that adolescents report in the in-school questionnaire about siblings living in their household, a number of genetic pairs can be determined. Identical twins, fraternal twins, half-siblings, and stepsiblings were sampled with certainty. In addition, a number of non-related pairs were recruited for the genetic sample. The majority of full-sibling pairs entered into the sample by chance. There are a total of 289 monozygotic (i.e., identical) twins, 453 dizygotic (i.e., fraternal) twins, and 1,251 full-sib non-twin pairs. Not examined in this paper are the 443 half-sibling pairs, 201 cousin pairs (living in the same household), and 461 non-related pairs (stepsiblings, adopted children, adolescents in group homes). Sample sizes for analysis will be less than these numbers, however, depending on whether there are valid data for both adolescents in each pair.

Non-familial relationships between adolescents in Add Health provide another basis on which to analyze the resemblance of pairs of individuals on a number of measures. Non-familial pairs include peers and grademates. In the in-school

questionnaire each adolescent respondent named his or her five best male friends and five best female friends. Friendship pairs (i.e., peers) are constructed on the basis of these nominations. Each adolescent can nominate between 1 and 10 friends, and pairing the nominator with each nominee forms the friendship pairs. Friendship pairs can span school grades, but must be from the same school. In-school and in-home data can be linked to the nominator and each nominee to construct the pair record with data on the same measures for each member of the pair. Reciprocated nominations are included in our file of friendship pairs, with the nominator and nominee in the opposite order for each pair observation. We eliminate all those cases without valid IDs, sex, age, and in-school weight.

We also constructed grademate pairs within grade and school. Because there is no relationship or nomination basis on which to construct these pairs, the universe of all pairs is constructed by simply forming each possible and nonredundant pair of adolescents within grade by school. The shared context of this set of pairs, therefore, is school grade. Recall that the in-school sample contains about 90,000 adolescents.

The analysis in its current form utilizes only data from the in-school interview, which was given to a much larger group of pairs. Future analysis will incorporate data from the in-home interview to analyze such outcomes as AHPVT, height, weight, and involvement in risk behavior.

We imposed some additional restrictions on our sample for the current analysis. We chose to examine adolescents aged 12 to 20 at the time of the in-school questionnaire. While there were a few 10- and 11-year old respondents to the in-school questionnaire, these students are quite unusual in that they are much younger than the typical 7th or 8th grader and have probably been promoted to the upper grade level based on achievement. Because of tremendous space demands, our computer facilities did not permit us to conduct analysis of all grademate pairs. Therefore, we chose a 20 percent random sample of schools in which grademate pairs were analyzed. We are developing a strategy for analyzing the full set of grademate pairs.

Given these restrictions, the sample sizes for the sets of pairs that we analyze are: 159 monozygotic twins, 206 dizygotic twins, 722 full non-twin sibs, 302,576 friendship

pairs, and 636,527 grademate pairs. Sample sizes still vary across outcome measures because of missing values for any member of a pair.

Measures

Grade Point Average. Grades from the most recent marking period were reported for the following subjects: English/Language Arts; Mathematics; History/Social Studies; and Science. The grade point average (GPA) is constructed by assigning a '4' to each A, a '3' to each B, a '2' to each C, and a '1' to each D or lower grade reported. Because Fs are part of the “D or lower” response category, the GPA measure is biased upward with a range from 1 to 4. GPAs are computed on the basis of the number of subjects for which grades were reported.

GPAs across all students who responded to the GPA questions in the in-school survey averaged 2.81, with a standard deviation of .81 (Table 1). We used the subset of 13,077 students who also responded to the in-home, personal interview to obtain a kind of test-retest reliability measure of GPA. Since the average time that elapsed between the two surveys was 7.6 months, almost certainly spanned marking periods and sometimes spanned school years, there legitimate reasons to expect a less-than-perfect correlation. In fact, the correlation was .68.

Mental Health Index. We constructed a mental health index based on seven items dealing with symptoms of emotional distress such as, having trouble with eating, difficulty falling asleep, feeling depressed or blue, crying a lot, and being afraid. Response categories range from 0, never, to 4, every day. We constructed an additive index by adding the response to each of the 7 items, resulting in a continuous index with a range of 0, indicating no emotional problems, to 28, indicating excessive symptoms of depression or poor mental health.

Just under 10 percent of the in-school sample reported no symptoms; scale scores averaged 6.8. The internal reliability of the scale, as revealed by Cronbach’s alpha, was .83 (which is based on 78,399 students with valid data).

Expectations. A set of questions asked adolescents about their expectations in a number of life-course domains. On an 8-point scale from "no chance" to "it will happen,"

adolescents were asked about the chance that they will: i) graduate from college; ii) live to age 35; and iii) be married by 25. We transformed the 8-category response scale into percentage chance by assigning, zero to a '0' on the scale (labeled "no chance"), .25 to a '2' ("some chance"), .50 to a '4' ("About 50-50"), .75 to a '6' ("pretty likely"), and 1.0 chance to a response of '8' ("it will happen"). The four unlabeled response categories between these values were assigned the midpoint of these scores.

Although two of these expectations items were asked in both the in-school and in-home waves, the response scales were not identical. This, as well as other measurement errors and true changes in expectations, produced relatively low correlations across the two waves -- .32 for "live to age 35" and .41 for "married by 25."

Number of school activities. A simple count of the number of clubs, organizations, and teams in which adolescents participated constitutes our activities outcome. School activities range from French Club to Computer Club, from Debate team, Band, Newspaper, Honor Society, Yearbook, Student Council to various sports. Adolescents were asked about 33 school activities; the in-school sample's scores on this measure ranged from 0, for no activities, to 33. Since a handful of students appeared to erroneously mark participation in all activities, we top-coded the number of activities that were greater than 10 to 10, which produced a small (.1) drop in the unweighted mean. For the overall in-school sample, the mean reported number of activities was 2.2; about one-quarter of the students reported engaging in no extra-curricular activities.

Risk Behavior. Two risk behaviors were measured in the in-school interview. Respondents were asked how often in the last 12 months they did "something dangerous because they were dared to" and "got drunk." We converted the responses based on the occurrence in a year to the number of times the adolescent engaged in each risk behavior in a month (i.e., a zero to "never", a 2.5 to "2 or 3 days a month", etc). The range on risk behavior varies from 0 (never engaged in the risk behavior) to 28 times per month ("nearly every day").

Some 61% of students reported never engaging in dangerous acts in the past year, and 69% reported never getting drunk. The correlation between in-school and in-home

reports of number of times drunk was low -- .24. (The “dangerous acts” item was not included in the in-home survey.)

In all cases we adjusted for age and gender differences in the values of these outcomes by regressing each outcome on a set of single-year age dummies using data from the entire in-school samples of boys and girls. The difference between an adolescent’s reported score and his or her age-based gender-specific predicted score is our dependent variable. Correlations for pairs for which there is no natural ordering (i.e., all but the self-nominated friends) are calculated using data in which each pair appears twice, once for each pair ordering, using methods described in Snedecor and Cochran (1967).⁶ Finally, pair data for grademates and peers are weighted by the inverse of the number of pairs of grademates or peers in the grade or friendship network.

IV. Results

Correlations calculated over the combined samples of boys and girls are presented in Table 2. Correlations for male-male, female-female and cross-sex pairs are detailed in Table 3.⁷

Consistent across all but one of the outcomes (expectations of marriage by age 25), correlations are relatively highest for identical twins. The difference between correlations for monozygotic and dizygotic twins (as well as other sibling types) is often quite large in these data, particularly in the case of college expectations (but less so for GPA), mental health, and engaging in either positive (i.e., extra-curricular activities) or negative (i.e., dangerous and drinking) acts. Since twice the difference in correlations between MZ and DZ twins is presumed in simple genetic models to reflect heritability, these data are consistent with a hypothesis of high heritability. We return to the issue of heritability in our discussion section.

Relatively low MZ twin correlations in the case of two of the expectation measures (living to age 35 and being married by age 25) are suggestive of either a large non-shared environmental component or measurement error. The low test-retest (in-school vs. in-home) correlations for these items described earlier reinforce a worry about

measurement error. However, expectations of college graduation produce the highest MZ twin correlations, which argues against blanket conclusions of a method-based measurement error for the expectations questions.

Almost without exception, grademate correlations are the lowest of all. This is consistent with the large literature showing the limited scope of school effects. That grademate correlations are quite low (.03) for the dangerous-acts and times-drunk measures suggests their universality across the very different kinds of schools in these nationally-representative data. The same- and cross-sex correlational data presented in Table 3 suggest that school context may matter relatively more for boys than girls for academic performance, college expectations and engaging in dangerous acts. However, we have yet to determine whether these relatively small differences might be attributable to sampling error.

Given the large element of self-selection in friendship networks as well as the potential for peers to exert causal effects on attitudes and behavior, we had expected quite large peer correlations. Although always at least twice as large as the grademate correlations, in only two cases (GPA for all types of pairs, extra-curricular activities for girls) was the peer correlation larger than .25. In only a couple of cases were peer correlations larger than the respective correlations of non-twin siblings.

Peer correlations for committing dangerous acts and getting drunk are surprisingly low, ranging from .03 to .13 across the various pairs of same- and cross-sex friends. Since one might expect adolescents to choose their friends partly on the basis of their willingness to engage in these kinds of risk behaviors, the residual scope for causal peer effects appears remarkably limited.⁸

V. Discussion

We have used correlations across a nationally-representative sample of siblings, grademates and peers to estimate upper bounds on the possible casual effects of families, school and friends on attitudes and behaviors. Our results show the largest correlations among siblings, especially identical twins. Consistent with the literature on school

effects, our grademate correlations are generally very small. Although somewhat larger, the peer-based correlations are also surprisingly small, suggesting a smaller scope for causal peer influences on achievement and, especially, risk behavior than might have been thought.

The relatively high correlations for MZ twins are somewhat surprising in light of the literature on genetic components of delinquency and criminality. The review of DiLalla and Gottesman (1990) shows a relatively small difference in mean concordance rates of juvenile delinquency between MZ (mean rate = .87) and DZ (.72) twins and a considerably larger difference in concordance rates for adult criminality (respective MZ and DZ rates = .51 and .22). They attribute the former difference to the frequent but transitory nature of much juvenile delinquency (Gold and Petronio, 1980) which, they argue, masks whatever genetic component there may be for the relatively small subset of individuals who continue to commit criminal acts into adulthood. MZ and DZ correlations for our “dangerous acts” (.62 and .22) and “times drunk” (.60 and .27) were more similar to their adult than adolescent correlations.

We urge caution in drawing conclusions regarding heritability based on these correlations. High MZ correlations could stem from nongenetic causes such as the tendency for parents, teachers and peers to give special, more similar treatment to identical twins; identical twins to participate in delinquent activities together; identical twins to share environmental experiences; or identical twins to respond more similarly to the self-reported questions in the Add Health study. Although some studies (Rowe, 1983) have tested for and failed to find support for some of these possibilities, the evidence to rule out these alternative explanations is hardly conclusive.

The preliminary nature of our work precludes drawing firm conclusions from these data. Planned refinements to the analysis include: i) the calculation of proper sampling errors for our correlations; ii) along the lines of Solon et al. (1997), residualizing peer and grademate correlations from the shared family background factor component in order to narrow the bounds of peer and grademate influences; iii) expanding our list of outcomes to include a receptive vocabulary test score, comprehensive violence and delinquency scales; iv) defining neighborhood contexts

based on adolescents' residential addresses; v) adding data on correlations between adopted and natural siblings, which is an indicator of shared environmental effects; and vi) estimation of more conventional multilevel models of family, school and peer effects using direct measures of the components of these environments.

¹ These conclusions hold for so-called "broad based heritability," which is the sum of additive and nonlinear genetic effects (Falconer, 1981). Factors that render problematic conclusions of simple behavioral genetics models include: i) in utero environmental shocks and twin competition for resources (Devlin et al., 1997); ii) more similar treatments accorded to twins by parents, teacher, and classmates; iii) more similar environments sought by MZ twins relative to other siblings; iv) measurement error; and v) more similar sibling mutual influence for MZ relative to other siblings.

² As we show below, an alternative interpretation is that familial effects that are similar among neighboring families drive the extra-familial correlations.

³ Solon et al. (1997) cite evidence from a national sample of children indicating that a Census-tract-based measure of average family income correlated .99, .90 and .76 for residential locations one, five and ten-years apart. Furthermore, the correlations between ten-year average neighborhood income and income in any one of the ten years were all at least .90 and averaged .94.

⁴ In the language of our earlier discussion of neighbor correlations, this spurious component of variance is analogous to the covariance in family backgrounds among neighboring children.

⁵ Some high schools spanned grades 7 through 12 and therefore served as their own "feeder" school, so the "pair" was in fact a single school. There are 134 discrete schools in the study.

⁶ In the case of twins living with other sibs, we only matched one twin from each twin pair with another sib in the family in our non-twin sib correlations.

⁷ Standard errors for the correlations are calculated under the assumption of simple random sampling. Given that the clustered nature of the sample produces a downward bias in standard errors calculated under the assumption of simple random sampling, we intend to calculate proper sampling errors for the correlations in the near future.

⁸ Note that adolescents who had dropped out of school prior to the 1994-1995 in-school interview are part of neither the in-school nor the in-home samples of respondents and peers. This probably restricts the variation of our two measures of risk behavior and reduces our correlations.

References

- Bearman, Peter S., Jo Jones, and J. Richard Udry. 1997. The National Longitudinal Study of Adolescent Health: Research Design [WWW document]. <http://www.cpc.unc.edu/projects/addhealth/design.html>
- Billy, J.O.G. and J.R. Udry. 1985. "The Influence of Male and Female Best Friends on Adolescent Sexual Behavior." Adolescence 20:21-32.
- Bronfenbrenner, U. 1979. The ecology of human development: Experiments by nature and design. Cambridge, MA: Harvard University Press.
- Cain, G. and Watts, H. 1972. "Problems in Making Policy Inferences from the Coleman Report" American Sociological Review, 35 (2): 228-252.
- Coleman, J. et al., 1966. Equality of Educational Opportunity. Washington, D.C.: GPO.
- Crane, J. 1991. "The Epidemic Theory of Ghettos and Neighborhood Effects on Dropping Out and Teenage Childbearing." American Journal of Sociology, 96(5):1126-1159.
- Devlin, B., Daniels, M. And Roeder, K. 1997. "The Heritability of IQ" Nature, 338, 31 July, 468-471.
- DiLalla, L. and Gottesman, I. 1989. "Heterogeneity of causes for delinquency and criminality: Lifespan Perspectives" Development and psychopathology, 1, 339-349.
- Duncan, G. J. and Raudenbush, S. 1998. "Getting Context Right In Studies of Child Development" Mimeo. Northwestern University, Evanston, IL.
- Falconer, D. 1981. Introduction to Quantitative Genetics. New York: Longman.
- Gold, M. and Petrino, R. 1980. "Delinquent Behavior in Adolescence." In J. Adelson (Ed.) Handbook of Adolescent Psychology, 495-535, New York: Wiley.
- Griliches, Z. 1979. "Sibling Models and Data in Economics: Beginnings of a Survey" Journal of Political Economy 87:S37-S64.

Haveman, R. & Wolfe, B. 1994. Succeeding generations: On the effects of investments in children. New York: Russell Sage.

Jencks, C., Bartlett, S., Corcoran, M., Crouse, J., Eaglesfield, D., Jackson, G., McClelland, D., Mueser, P., Olneck, M., Schwatz, J., Ward, S., and Williams, J. 1997. Who Gets Ahead: The Determinants of Economic Success in America, New York: Basic Books.

Jencks, C. and Brown, M. D. 1975. "Effects of High Schools on Their Students" Harvard Education Review, (45), 273-324.

Kandel, D.B. 1978. "Homophily, Selection, and Socialization in Adolescent Friendships." American Journal of Sociology 84:427-436.

Manski, C. 1993. "Identification of Endogenous Social Effects: The Reflection Problem." Review of Economic Studies, 60:531-542.

Plomin, R., DeFries, J. C., and McClearn, G. E. (1990) Behavioral Genetics: A Primer (2nd Ed.), New York: W. H. Freeman.

Rosenthal, R., and Rubin, D. 1982. "Comparing effect sizes of independent studies." Psychology Bulletin, 92, 500-504.

Rowe, D. 1983. "Biometrical Genetic Models of Self-reported Delinquent Behavior: A Twin Study" Behavioral Genetics, 13, 473-489.

Rowe, D. 1994. The Limits of Family Influence: Genes, Experience and Behavior, New York: The Guilford Press.

Snedecor, G. and Cochran, W. (1967) Statistical Methods (6th Edition), Ames, IA: Iowa State University Press.

Solon, G., Page, M. and Duncan, G. 1997. "Correlations Between Neighboring Children in Their Socioeconomic Status as Adults" Mimeo. University of Michigan.

Urberg, K.A., Degimencioglu, S., Tolson, J., and Halliday-Scher, K. 1995. "The Structure of Adolescent Peer Networks" Developmental Psychology, (31), 540-547.

Table 1: Weighted means and standard deviations for outcomes measures, for all respondents in the in-school survey.

Outcome	Mean	Standard Deviation	Minimum	Maximum	Fraction Zero
Grade point average	2.81	.80	1.0	4.0	
Mental health	6.83	5.47	0	28	9.9
Expectation of graduating from college	78.7	29.1	0	100	4.1
Expectations of living to age 35	82.7	24.6	0	100	2.3
Expectation of married by age 25	53.9	30.7	0	100	8.6
Monthly frequency of dangerous acts in past year	1.67	5.59	0	28	60.6
Monthly frequency of getting drunk in past year	1.06	3.88	0	28	68.6
Number of extra-curricular school activities	2.19	2.12	0	10	25.6
Number of observations	77,446				

Source: Authors' calculations from the Add Health Study

Table 2: Sibling, grademate and peer correlations for various adolescent outcomes

Outcome	MZ twins	DZ twins	Non-twin siblings	Grademates	Closest friends
Grade point average	.58	.44	.30	.09	.33
Mental health	.37	.06*	.10	.01	.08
Expectation of graduating from college	.62	.10*	.33	.04	.18
Expectations of living to age 35	.21	.12	.11	.03	.09
Expectation of married by age 25	.12	.23	.14	.01	.07
Monthly frequency of dangerous acts in past year	.62	.22	.09	.03	.06
Monthly frequency of getting drunk in past year	.60	.27	.16	.03	.10
Number of extra-curricular school activities	.47	.25	.29	.08	.23
Number of observations	109-159	131-206	494-722	458,656- 636,527	221,108- 302,576

Note: An asterisk indicates the correlation is not significantly different from zero at the .05 level.

Table 3: Male-male, female-female and cross-sex sibling, grademate and peer correlations for various adolescent outcomes

Outcome	MZ twins	DZ twins	Non-twin siblings	Grademates	Closest friends
Grade point average	.59	.38	.33	.11	.33
	.57	.63	.31	.07	.36
		.41	.28	.08	.28
Mental health	.29	.14*	.10*	.01	.07
	.41	.19*	.09*	.02	.10
		-.07*	.11	.01	.05
Expectation of graduating from college	.53	.06*	.36	.05	.21
	.74	.26*	.35	.03	.19
		.08*	.30	.04	.15
Expectations of living to age 35	.18	.35	.14	.02	.08
	.22	.33	.03*	.05	.10
		-.13*	.14	.03	.07
Expectation of married by age 25	.18	.44	.15	-.01*	.06
	.05*	.13*	.13	.01	.09
		.17	.14	.01	.06
Monthly frequency of dangerous acts in past year	.67	.46	.07*	.05	.08
	.47	.18*	-.02*	.02	.06
		.01*	.14	.01	.03
Monthly frequency of getting drunk in past year	.73	.60	.09*	.04	.10
	.38	.49	.03*	.03	.13
		.00*	.21	.02	.07
Number of extra-curricular school activities	.55	.26	.22	.09	.22
	.40	.29	.39	.10	.26
		.21	.24	.07	.18
Number of observations	56-79	42-57	149-200	129,966-166,992	60,665-81,585
	53-80	24-48	146-223	114,045-150,897	71,394-99,184
		65-101	199-299	244,291-318,638	89,049-121,807

Note: First entry in cells is male-male correlation; second entry is female-female correlation; third entry is male-female and female-male correlation. An asterisk indicates the correlation is not significantly different from zero at the .05 level.

Source: Authors' calculations based on the Adolescent Health Survey