A behavior genetic analysis of the tendency for youth to associate according to GPA

J.C. Barnes\textsuperscript{a,}\textsuperscript{*}, Kevin M. Beaver\textsuperscript{b,d}, Jacob T.N. Young\textsuperscript{c}, Michael TenEyck\textsuperscript{a}

\textsuperscript{a} The University of Texas at Dallas, School of Economic, Political & Policy Sciences, 800 West Campbell Road, Richardson, TX 75080, United States
\textsuperscript{b} Florida State University, College of Criminology and Criminal Justice, 634 West Call Street, Tallahassee, FL 32306, United States
\textsuperscript{c} Arizona State University, School of Criminology and Criminal Justice, 411 North Central Avenue, Suite 600, Phoenix, AZ 85004, United States
\textsuperscript{d} King Abdulaziz University, Center for Social and Humanities Research, Jeddah, Saudi Arabia

\textbf{A R T I C L E   I N F O}

\textbf{Keywords:}
Homophily
Peer network GPA
Behavior genetic
Active gene–environment correlation (rGE)

\textbf{A B S T R A C T}

Behavior genetic research has revealed that many “environmental” variables are partially influenced by genetic factors. Known as gene–environment correlation (rGE), this line of scholarship provides insight on how and why individuals select into certain environments. Juxtaposing this body of evidence with research on peer group homophily—the tendency for peers to resemble one another on certain traits such as academic ability—raised two research hypotheses: (1) youth will associate with peers who receive grades similar to themselves (i.e., homophily for GPA); and (2) a portion of the variance in peer group GPA (i.e., the peer network average GPA) will be explained by individuals’ genetic self-selection into the peer group (rGE). The results supported both hypotheses by showing a strong predictive relationship between the target individual’s GPA and that of his/her peers and by revealing that 72\% of the variance in peer group GPA was explained by genetic influences.

\textcopyright 2014 Elsevier B.V. All rights reserved.

Since the publication of Lazarsfeld and Merton’s seminal work on peer group processes in 1954, social scientists have devoted considerable efforts to understanding peer group homophily. Homophily is typically defined as the tendency for individuals with shared interests to form friendship (or network) ties. A popular colloquialism is commonly used to define homophily: “birds of a feather flock together” (McPherson et al., 2001). Much research has been devoted to understanding the origins (Rushton et al., 1984; Young, 2011; Young et al., 2013) and effects of homophily (McPherson et al., 2001; Rivera et al., 2010), but there are certain aspects of peer group formation that have yet to be analyzed. The current study will shed light on one aspect of adolescent peer group homophily that has, for the most part, been overlooked by social scientists. Specifically, this analysis will examine whether youth tend to associate with peers who perform similarly at school and whether genetic factors are able to explain the phenomenon.

Drawing on a nationally representative data source—and a large subsample of twin pairs—the current study will test two hypotheses dealing with academic ability and peer group homophily. The first hypothesis is that adolescents will show a strong correlation with their peers on grade point average (GPA). Recognizing that a correlation between one’s GPA and his or her peers’ GPA could be the result of several causal mechanisms, we propose a second hypothesis: peer group GPA is structured by genetic factors. A nascent body of evidence indicates that genetic factors correlate with peer group formation and selection (Beaver et al., 2008; Fowler et al., 2009; Rushton et al., 1984; Yun et al., 2011), lending theoretical support for our second hypothesis. Below, we expand our discussion of these two topics and lay the theoretical groundwork for both hypotheses.

1. Homophily in peer groups

A large body of evidence has revealed that members of a peer group tend to resemble one another on a host of traits. Evidence from psychology reveals, for example, that individuals tend to resemble their peers on measures of physical characteristics, attitudes, and aspirations (Caspi and Roberts, 2001; Harris, 1998; McPherson et al., 2001). Similarly, sociological and criminological evidence indicates that peers tend to resemble one another on measures of delinquent and criminal involvement (Akers, 1998; Kandel, 1978; Pratt et al., 2010). Perhaps the most important question underlying all of this research is whether the similarity between peers is due to causation, selection, or a mixture of both. The answer to this question has major implications for theories that cut across social science disciplines and for interventions that target specific causal mechanisms.
Interestingly, current evidence seems to suggest that peer groups may indeed have a causal influence on the behavior of individual members. A string of recent articles by Christakis and Fowler revealed how changes in behavioral tendencies can spread through a peer group, wherein peers tend to resemble one another in obesity (Christakis and Fowler, 2007), sleep loss (Mednick et al., 2010), smoking behaviors (Christakis and Fowler, 2008), drinking behaviors (Rosenquist et al., 2010a), depression (Rosenquist et al., 2010b), and in levels of happiness (Fowler and Christakis, 2008). Studies exploiting random assignment in college dormitories have found that roommates influence one another’s GPA (Sacerdote, 2001; Zimmerman, 2003) and fitness level (Carrell et al., 2009). Importantly, all of these studies reported evidence consistent with a causal explanation—that the behavior of group members is partially endogenous and a consequence of peer influence. For example, evidence indicated that a person’s risk of smoking was decreased if their spouse quit smoking, if their sibling quit smoking, if their friend quit smoking, or if their coworker quit smoking (Christakis and Fowler, 2008). A similar pattern of results was reported in other studies carried out by this group of researchers (e.g., Christakis and Fowler, 2007; Fowler and Christakis, 2008; Rosenquist et al., 2010a,b).

Although the Christakis–Fowler studies have provided compelling support for the causal explanation, their analytical models were unable to account for all causes of peer similarity. Moreover, the question remains as to whether the same pattern of findings would emerge if the outcome were something known to be more stable over time such as cognitive ability or academic performance (Deary et al., 2000; Sameroff et al., 1993). As Sacerdote’s (2001) analysis of randomly assigned college roommates revealed, roommates influence one another’s GPA, raising the question of whether peers (note that roommates may not always perceive themselves as friends) sort themselves according to academic performance or whether the relationship is causal.

2. Cognitive ability as a sorting factor

Findings from one stream of research speak to the intellectual similarity of individuals within a peer group (Bonney, 1942). Two classic studies are particularly relevant when considering the level of similarity between peers on measures of cognitive ability. In the first study, Almack (1922) performed one of the first analyses examining intelligence as a peer-sorting factor. Participants were drawn from grammar and middle schools in San Jose, California and were administered IQ tests. Next, each respondent was asked to identify a single student who they would choose to help them with some work they had been assigned to complete. Respondents were also instructed to identify the student who they would invite to a party. For work assignment and party invitation, the analysis revealed that participants tended to choose students who were similar to themselves in regard to intelligence. Indeed, the correlation between the respondent’s IQ and the peer’s IQ ranged between .304 and .408.

In the second study, Seago (1933) approached the issue in a fashion similar to Almack (1922) by studying a group of 3rd through 8th graders and asking them the following question: “Suppose you were going to a party and could invite just one person to go with you. Suppose it could be either a boy or a girl, but not a member of your family. Whom would you ask to go with you?” Based on the responses to this question, Seago analyzed the similarity of pairs of friends for traits such as height, athleticism, and intelligence. Important for the current focus was that peers exhibited a moderate-to-high correlation for IQ score, with the correlations ranging between .386 and .513 depending on the stability of the friendship bond. Unfortunately, multiple regression analysis was not performed, so it was impossible to rule out other sources of spuriousness such as propinquity. Nevertheless, these findings raise an important issue that begs further attention: do people select their friends based on similar levels of cognitive ability?

Both Almack’s (1922) and Seago’s (1933) early assessments appear to suggest that peer selection is partially influenced by the degree to which two people have similar intelligence levels. More recent evidence also supports this viewpoint but suggests that causal influences may also play a role in peer group academic performance (Lomi et al., 2011; Sacerdote, 2001). For example, Lomi et al. (2011) used stochastic agent-based models to examine the simultaneous relationship between changes in GPA and network transformation. These authors reported, “the academic performance of an individual student tends to become similar to (or ‘assimilate’) the performance of his or her peers (or to remain similar)” (emphasis in original, p. 1514). These results suggest that correlated academic performance may be a product of both selection and causation.

3. Genetic influences on peer group selection

Building on principles of evolutionary biology, Rushton et al. (1984) proposed a genetic similarity theory which states that organisms detect other genotypically similar organisms via the manifestation of genotype through phenotype (e.g., behavior) and exhibit favoritism and protective behaviors toward those similar others. In other words, the genetic influences are expressed as phenotypic outcomes, or signals, that are then noticed by similar others. The more phenotypically (genetically) similar two organisms are, the more likely it is that they will demonstrate altruistic behavior toward one another. In terms of the friendship network, this means that individuals may seek out other phenotypically (genetically) similar individuals and provide mutually supportive environments. Genetic similarity theory suggests that genes are the underpinnings to friendship, in that liking an individual leads to friendship and this friendship then leads to mutually beneficial, cooperative behavior. To be sure, genetic similarity theory argues that there are evolutionary advantages for individuals having the ability to detect genetic similarities in others and then forming a friendship. Namely, such friendship ties are more likely to lead to reciprocal altruism, which could positively impact fitness levels (evolutionarily speaking).

Research has revealed that adolescent peer groups form as a result of various factors, many of which are social/environmental in origin (McPherson et al., 2001; Rivera et al., 2010; Seago, 1933). Emerging evidence is, however, beginning to suggest that adolescent peer groups form as a result of active gene–environment correlation (rGE), a finding that appears to support Rushton et al. (1984) theory of genetic similarity. Behavioral geneticists (Plomin et al., 1977; Scarr and McCartney, 1983) have proffered the term “active rGE” as a way to explain the tendency for humans to select into environments that suit their genetic propensities. Since personality development is at least partially the function of genetic factors (Bouchard et al., 1990) and since peers have been shown to correlate strongly for personality traits (McPherson et al., 2001), scholars have begun to search for genetic influences on peer group selection by drawing on the inferences proffered by Rushton et al. (1984) genetic similarity theory and similar other statements (e.g., Harris, 1998). This line of research has already produced consistent evidence suggesting that peer group selection is partially governed by genetic factors (Beaver et al., 2008, 2011; Cleveland et al., 2005; Fowler et al., 2009; Guo, 2006; Yun et al., 2011).

Though each study has approached the peer group selection issue in a slightly different way, most have hypothesized that genetic factors influence personality development (i.e., the
phenotype), which in turn becomes a selection factor for peer group sorting. Drawing on data from the National Longitudinal Study of Adolescent Health (Add Health), Fowler et al. (2009) analyzed the genetic influences on four dimensions of peer networks: in-degree (the number of times a person was nominated as a friend); out-degree (the number of friends a person has); transitivity (the probability that two of a person’s contacts are friends); and, centrality (the portion of network connections that operate through the target individual). Their analysis revealed that genetic factors explained a statistically significant amount of variance in in-degree, transitivity, and centrality. Along similar lines, Cleveland et al. (2005), also drawing from the Add Health, reported that the majority of the variance in a measure of peer group delinquency was explained by heritable influences. Cleveland et al.’s findings have been supported by two recent studies examining peer group selection using molecular genetic data in the Add Health (Beaver et al., 2008; Yun et al., 2011).

Additional evidence indicating that peer group selection is partially the result of genetic factors came from a recent study examining the genetic underpinnings to the stability in exposure to delinquent peers. In this study, Beaver et al. (2011) calculated bivariate Cholesky models to estimate the proportion of variance in the stability of delinquent peers over time that was attributable to genetic influences. The results revealed that genetic factors accounted for 58% of the covariance in exposure to delinquent peers between two waves of data collection from the Add Health. These findings strongly suggest that genetic factors are associated with peer group selection because, even though the composition of individual members of the peer groups changed over time, genetic factors continued to influence youth to select peer groups that shared similar behavioral patterns.

Directly relevant to the current study is the analysis presented by Guo (2006). Guo analyzed the Add Health sibling files to determine whether siblings who share more genetic material tend to have friends who are more similar to one another as compared to siblings who share less genetic material. For instance, Guo tested to see whether the best friends of identical twins were more similar than the best friends of fraternal twins. The findings revealed that identical twins’ friends were more similar than fraternal twins’ friends across three of the four variables studied: GPA, verbal intelligence, and aggression; the evidence was mixed and not statistically significant for depression. Though Guo did not directly estimate genetic influences, these findings indicate that genetic factors may play a role in the selection of peers for certain characteristics.

Examining the role of genetic factors on peer network characteristics is important for understanding how a variety of outcomes/behaviors may be correlated. A singular conception of homophily as a purely environmental phenomenon may limit the ability to make rich inferences about how relationships unfold during adolescence. Schaefer (2010: 22), for instance, has recently argued that configurations of homophily, defined as “a combination of similarity and dissimilarity across several dimensions,” may be important for understanding how multiple attributes are correlated. Moreover, Rivera et al. (2010: 108) noted that “A second promising avenue for future research is to explore interactions between assortative, relational, and proximity mechanisms” in the genesis and maintenance of social networks. Building on this expanding framework of theoretically influential mechanisms, we propose that genetic factors may also play a role in peer group homophily.

4. The current study

Much attention has been given to the role that peers play in personality development (Harris, 1998) and research findings have consistently shown that all aspects of human development are driven, at least partially, by genetic factors (Turkheimer, 2000). Among these genetically influenced traits is one’s cognitive ability (Plomin, 1999; Plomin and Spinath, 2004). Juxtaposing the knowledge that between-individual differences in cognitive ability are partially the result of genetic influences alongside evidence that properties of one’s peer network are genetically influenced raises three interesting research hypotheses:

Hypothesis 1. Adolescents will tend to associate with peers who have GPAs similar to themselves.

Hypothesis 2. Variance in peer group GPA will be, at least partially, attributable to heritable genetic factors.

Hypothesis 3. A portion of the covariance between an adolescents’ self-reported GPA and that of their peers will be attributable to heritable genetic factors.

The link between heritability of cognitive ability and similarity in friendships is important because a large literature points to the role of peers in affecting GPA and other measures of academic performance. This study will help shed light on how much of the correlation between peers on cognitive ability is due to genetic self-selection and how much may be due to other environmental influences, perhaps even causal influences.

5. Methods

5.1. Data

Data drawn from the National Longitudinal Study of Adolescent Health (Add Health; Harris, 2009) were analyzed in the current study. The Add Health is a prospective, nationally representative study of American adolescents who were enrolled in middle or high school during the 1994–1995 academic year. Sampling began at the school level and, from a list of nearly 30,000 schools, 132 schools were selected for inclusion in the study. On a designated day, all students attending each of the schools were asked to complete a short self-report questionnaire. This round of interviews netted information from more than 90,000 youths and is referred to as the in-school survey.

Beginning immediately after the in-school interviews were completed, a more detailed follow-up interview was conducted with 20,745 students who also participated in the in-school survey. This round of interviews is referred to as the wave 1 in-home survey and will be used to glean information for target respondents. The in-home interview was designed to gather more information than the in-school survey. For instance, participants were asked to report on their personality traits, health concerns, relationships with teachers, activities with peers, and involvement in delinquent behavior.

Two unique features of the Add Health data are capitalized upon in the current analysis. First, the in-school survey data can be used to construct and analyze peer networks (Haynie, 2001). During the in-school interview, each respondent was provided a roster of students enrolled in their school and asked to nominate five male friends and five female friends (nominations could also be made for students who attended a sister school). As a result, the responses from nominated peers can be used to construct peer network measures for target respondents (i.e., wave 1 in-home survey respondents). A portion of target respondents did not have valid peer network data. As a result, the analytical sample for the multivariate models was restricted to target respondents who had valid network data (analytic sample size ranged between 8930 and 11,258).

The second feature of the Add Health that will be utilized is the subsample of twins that were identified during wave 1 interviews.
Specifically, if a respondent identified him/herself as a twin during the in-school interview or during wave 1 in-home interviews, his/her twin was automatically added to the sample. The initial sample of twin siblings included 906 dizygotic (DZ) twins and 578 monozygotic (MZ) twins. After eliminating cases due to missing data on the peer network measures, a final sample of 308 DZ twins and 208 MZ twins was analyzed for the behavioral genetic portion of the study.

5.2. Measures

5.2.1. Peer network grade point average (GPA)

The peer network average GPA measure reflects the mean GPA across four core subjects for all participants who are identified as being within the target respondent's peer network. The peer network measures were preconstructed by Add Health researchers (Carolina Population Center, 2001). For the current study, we utilized the send version of the peer network. The send peer network reflects the peer network as it is perceived by the target respondent. In other words, the send network includes peers who are nominated by the target respondent and does not require that the nominated peer reciprocate the nomination to the target respondent.

To calculate the peer network average GPA, a two-step process was followed. First, peers were asked to self-report their grades in four core subjects (English/language arts, mathematics, history/social studies, and science). Grades were scored according to the following scale: A = 4, B = 3, C = 2, and D (or lower) = 1. Grades for all four subjects were summed and then scaled by the number of courses to calculate each individual peer's GPA. Second, the average GPA for the target respondent's peer network was calculated by summing together the GPAs of all members of the peer network and then dividing by the number of individuals in the network (i.e., the number of nominations sent by the target respondent). The final measure ranged from 1 to 4 with a mean of 2.83 and was appropriately normally distributed. Descriptive statistics for all variables used in the analysis are presented in Table 1 for wave 1 in-home respondents. Importantly, Barnes and Boutwell (2013) reported that twins did not significantly differ from non-twins on a host of measures available in the Add Health data and our own analysis of the data presented here did not reveal any substantive differences between the twin subsample and the full sample of wave 1 in-home participants.

5.2.2. Respondent grades and GPA

During wave 1 in-home interviews, target respondents were asked to self-report their grades in four core courses (English/language arts, mathematics, history/social studies, and science). Note that the questions referenced the same four courses that were asked to the target respondent's peers in the in-school questionnaire. As described above, grades were assigned the following values: A = 4, B = 3, C = 2, and D (or lower) = 1. Next, each target respondent's grades across the four subjects were averaged to generate their overall GPA. In sum, five measures of the target respondent's grades were included: their grade in English/language arts, their grade in mathematics, their grade in history/social studies, their grade in science, and their overall GPA.

5.2.3. Control variables

In order to avoid model misspecification, the target respondent's age, gender (0 = female, 1 = male), and race (0 = non-Black, 1 = Black) were included in the multiple regression analyses.

6. Analysis plan

The analysis unfolded in three interrelated steps. First, the correlation between the target respondents' grades/GPA and the average GPA for the peer network to which they belonged was examined in a series of multiple regression models (Hypothesis 1). These models employed ordinary least squares regression because the dependent variable, peer network average GPA, approximated a normal distribution. All target respondents in the wave 1 in-home data file who had valid network data were analyzed in this portion of the study.

The next step was to determine whether genetic factors explained variance in peer group GPA. To do so, an ACE model was estimated on the peer network average GPA measure using data gleaned from twin pairs (see above). The ACE model—which is presented graphically in the top panel of Fig. 1—was estimated using the statistical program, Mplus. Briefly, the ACE model provides an indication of the degree to which genetic (A), shared environmental (C), and nonshared environmental (E) factors explain variance in the outcome of interest (i.e., peer network GPA). The ACE model decomposed the variance in peer network GPA by analyzing the degree to which twins of different zygosity correlated with each other on that measure. If MZ twins—who are identical twins that share 100% of their DNA—resemble one another more closely than DZ twins—who only share, on average, 50% of their distinguishing DNA—then genetic factors are likely to play a role in the selection of peer network GPA. Thus, to the extent that genetic factors explain variance in peer network GPA, support for Hypothesis 2 will be gained. To estimate these models, the genetic subsample of MZ and DZ twins was analyzed.

Finally, the third portion of the analysis involved estimating a bivariate Cholesky model, which is presented in the bottom panel of Fig. 1. The Cholesky model is similar to the ACE model in that it will estimate the influence of genetic and environmental factors on variance in observed factors. Where the Cholesky model diverges from the ACE model is that the former can simultaneously estimate the impact of genetic/environmental factors on variance in two observed outcomes, as well as the genetic/environmental influences on the covariance between the two outcomes. Thus, the Cholesky model will be utilized to estimate the degree to which the correlations highlighted in step one of the analysis are due to genetic self-selection or environmental factors (Hypothesis 3).

7. Findings

Presented in Table 2 are the results from five regression models where the peer network average GPA served as the dependent variable and all in-home participants with non-missing data at wave 1 were analyzed. This table reveals one main finding: the target respondent’s grade in all four subjects and the target respondent’s overall GPA is positively and significantly related to the peer network average GPA. Standardized regression coefficients are presented in brackets to allow for the observation of the strength of the relationship between the target respondent's grades/GPA and the
peer network average GPA. For example, the target respondent’s GPA emerged as the strongest predictor of the peer network average GPA ($\text{Beta} = .41$). In summary, the findings presented in Table 2 indicate that youth have a tendency to associate with peers whose school performance is similar to their own.

In order to gauge the sensitivity of our findings to alternative measurement strategies, we re-estimated each of the analyses discussed above by substituting the send-and-receive version of the peer network GPA measure for the send version that was reported here. Importantly, all substantive findings were identical to those reported in Table 2 when the send-and-receive version of the measure was analyzed. Specifically, the predictive effect of the respondent’s grades on his/her peer network’s GPA was (all are $\text{Beta}$ coefficients and were statistically significant at $p < .01$) .33, .27, .32, .30, and .42 for English Grade, Math Grade, History Grade, Science Grade, and GPA, respectively.
In an effort to explain the potential self-selection mechanisms that are driving the findings observed in Table 2, we analyzed the twin subsample to determine whether genetic and environmental factors played a role in the relationship. First, the findings presented in Table 2 were replicated with the twin data file. It is important to note that the substantive findings presented in Table 2 were replicated when the data were restricted to twins. Specifically, the predictive effect of the respondent’s grades on her peer network’s GPA was (all are Beta coefficients and were statistically significant at p < .01).31, .23, .36, .32, and .41 for English Grade, Math Grade, History Grade, Science Grade, and GPA, respectively. Second, cross-twin correlations for the peer network average GPA measure were examined. These correlations are presented in Table 3 and indicate that genetic factors may play a role in the peer-group selection process. Specifically, the MZ cross-twin correlation coefficient was larger than the DZ cross-twin correlation. These correlations also suggest that environmental factors play a role, especially non-shared environmental factors, since the MZ cross-twin correlation was below unity.

Once again, we replicated the results with the send-and-receive version of the peer network measure. Specifically, when the send-and-receive measure was analyzed, the cross-twin correlation for All Twins was r = .66 (p < .05), the cross-twin correlation for MZ Twins was r = .82 (p < .05), and the cross-twin correlation for DZ Twins was r = .55 (p < .05).

Next, an ACE model was estimated on the peer network average GPA measure. The results from the full ACE model, the AE model, the CE model, and the E model can be found in Table 4. Note that the best-fitting model, as determined by model-fit statistics, is presented in bold. As can be seen, the E model provided the best fit to the data. The estimates from this model revealed that genetic factors accounted for 72% of the variance in the peer network average GPA measure with nonshared environmental factors explaining the remaining 28% of the variance. Substantive results were identical when cases with missing data on the target GPA measure were excluded from the analysis.

The ACE model was estimated on the send-and-receive peer network GPA measure as a sensitivity check. The substantive conclusions were similar to those presented for the send version in that genetic factors were the predominant influence on variance in peer group GPA. The primary point of difference between the findings for the send-and-receive measure versus the send measure was that the full ACE model was the best-fitting model for the former (χ² = 3.14; the next best-fitting model was the AE model χ² = 10.11). Parameter estimates for the ACE model suggested that variance in send-and-receive peer network GPA was accounted for by a combination of genetic factors (A = .53), shared environmental influences (C = .29), and nonshared environmental factors (E = .19).

The final step to the analysis was to estimate whether the covariance between peer network GPA and the target respondent’s GPA (see Table 2) was due to genetic factors operating on both outcomes.1 This analysis was performed by estimating the

---

1 Results not reported in a table revealed the target respondent GPA to be under substantial genetic influence (AE was best-fitting model with χ² = 5.20, A = .68, and E = .32). This was an expected finding given the large literature reporting a genetic influence on cognitive ability (e.g., Plomin et al., 2013).
bivariate Cholesky model [see bottom panel of Fig. 1] on the covariance between the target respondent GPA and the peer network GPA. Results from the Cholesky models are presented in Table 5. The results reveal that the full ACE model best fit the data and that covariance between target respondent GPA and the peer network GPA is explained by a combination of genetic factors (A = .43) and shared environmental factors (C = .60) that are common to both outcomes (the nonshared environmental component was negative and can be interpreted as being non-significant). In other words, peer group homophily for GPA arises due to a combination of genetic and shared environmental influences. The Cholesky model results were replicated with the send-and-receive measure of peer network GPA. The results from the send-and-receive models were substantively similar to the send results reported in Table 5. Specifically, the ACE model was the best-fitting model with genetic factors accounting for 48% of the covariance, shared environmental factors accounting for 49% of the covariance and nonshared environmental factors accounting for the remaining 4% of the covariance.2

8. Discussion

There is little doubt that peers resemble one another on myriad traits. Indeed, much research has reported a correlation between peers on a range of personality characteristics, health outcomes, behaviors, and even physical attractiveness (McPherson et al., 2001). What has yet to be determined is why peers resemble one another so closely. Undoubtedly, much of the extant research can be explained in a causal framework. The work by Christakis and Fowler (2007, 2008) displays a pattern of findings that strongly suggest peers influence one another’s behavior and theorists have long hypothesized that peer effects are among the most powerful influences on personality development (Harris, 1998). While peers may influence each other’s behaviors in some respects—such as smoking, drinking, or eating habits—it is less clear whether they have influence over more stable traits such as cognitive ability and other related outcomes like GPA.

The current study analyzed three hypotheses specific to the similarity of peers in terms of GPA. The first hypothesis was that peers would resemble one another in their GPA. A series of multiple regression models were estimated and one consistent finding emerged: youth tend to associate with peers who perform similarly to themselves in school. In short, the analyses revealed much support for Hypothesis 1. The second hypothesis stated that peer group selection for GPA would be explained at least partially by genetic factors. In this way, Hypothesis 2 drew on the concept of active rGE which notes that individuals choose environments that are well suited to their personality and their abilities (Scarr and McCartney, 1983). ACE model results revealed that 72% of the variance in peer group GPA was attributable to genetic self-selection, lending strong support for Hypothesis 2. The remaining variance was attributable to nonshared environmental influences. Finally, the third hypothesis stated that a portion of the covariance between the target respondent GPA and the peer network GPA (Hypothesis 1) would be attributable to genetic factors. Findings from the Cholesky analysis suggested that this was indeed the case and that the covariance between target respondent GPA and the peer network GPA was explained by a combination of genetic and shared environmental factors.

The current findings are consistent with a long line of research showing peers tend to resemble one another in terms of academic ability (see Almack, 1922; Seagoe, 1933; McPherson et al., 2001).

---

2 Justification for selecting the AE model as the best fit to the data for the Cholesky analysis also existed. Note that the AE model was the best fit in the ACE model presented in Table 4 and in footnote 1. Thus, one might argue that the shared environment (i.e., C) cannot explain covariance of two items that have previously been shown to vary based on genetic (i.e., A) and nonshared environmental (i.e., E) influences. We opted to adhere strictly to fit criterion when selecting the best fitting model in all analyses, but we recognize this substantive interpretation as being important as well.
At the same time, they offer a unique explanation for why peers resemble one another in academic ability (Guo, 2006). The analysis suggests that individuals actively select into peer groups based on their own genetic propensities and that these genetic propensities are the same genes that affect one’s own academic abilities, perhaps lending credence to the common colloquialism, “birds of a feather flock together.” It is important to keep in mind, however, that the current results do not suggest that individuals consciously choose peers who are more similar to themselves in terms of academic ability or genetic profile. As noted by Almack (1922:529), “Unquestionably the selecting goes on unconsciously. One is often not analytic in picking out his friends.” Thus, finding an active rGE for peer group GPA does not necessarily imply conscious selection. It is likely the result of a complex arrangement of influences involving the propensity for relationships to form based on actors attributes, shared relationships, propinquity, and social foci (Feld, 1981; Rivera et al., 2010). In an evolutionary framework, our results are supportive of Rushton et al. (1984) theory in which genetic similarity draws individuals together—perhaps below the level of consciousness—leading to friendship that facilitates altruism (see also, Guo, 2006).

It is important that the results from two recent studies (Boardman et al., 2012; Sacerdote, 2001) be kept in mind when placing the present findings into a theoretical/explanatory framework. Boardman et al. (2012) reported that genetic homophily for the DRD2 gene was stronger in schools with greater levels of inequality. The explanation of this finding provided by the authors suggests that schools with high levels of inequality are more likely to implement academic tracking policies in which students are grouped by academic ability. This tracking policy would, in effect, constrain friendship choices because an individual is more likely to befriend those in their class (school-level propinquity) (Rivera et al., 2010). If students share many classes with those who have similar academic ability this would influence their potential pool of friends and, therefore, they would be more likely to befriend those with similar GPAs. In this scenario, the sorting of peers according to cognitive ability would result due to a complex arrangement of factors at the school-level (i.e., tracking policies), as well as at the individual-level (i.e., individual differences in academic performance). Boardman et al. (2012) finding indicates that school environments and policies such as academic tracking may lay the foundation for additional unobserved selection factors, perhaps offering an alternative explanation for the current results. At the very least, the Boardman and colleagues’ findings offer evidence to suggest peer group homophily is a complex phenomenon caused by factors operating on multiple levels/dimensions (Schafer, 2010). For example, our findings may be integrated with Boardman et al.’s if the following causal pathway is realized: genes → grades/GPA → academic tracking → peer group formation.

The second analysis that should be kept in mind when interpreting the present findings was presented by Sacerdote (2001) and relied on the random assignment of roommates to Dartmouth College dorm rooms. Sacerdote identified a small-to-moderate correlation between roommate GPAs during their freshman year. Because roommate assignment was random, this association is, ostensibly, a causal one (Sacerdote was able to establish that roommate assignments were random and also performed a thorough analysis to rule out plausible spurious associations). Thus, Sacerdote’s findings appear to conflict with the current conclusion that youth select into peer groups with similar academic performances and that this selection is driven by antecedent genetic factors. Note, however, that the current findings do not suggest that the entire association between one’s academic performance and that of his/her peers is due to genetic selection. Rather, environmental influences also emerged as significant, leaving room for causal effects such as those presented by Sacerdote. An obvious complication in comparing our results to those of Sacerdote, however, is that our study analyzed high school students who are less likely to live with one another (as do college roommates). Thus, we may expect fewer, or perhaps qualitatively different “peer effects” in the present sample.

Though the current study is one of the first to analyze the genetic influences on peer group GPA, several limitations must be kept in mind when interpreting the findings. First, all estimates presented in this analysis are likely to be affected by the restriction of friend choice to the individual’s schoolmates. The school environment is unique in that it restricts the variance in any trait to a relatively small group of individuals. The resulting influence may be that peer group selection (i.e., active rGE) is limited due to relatively few options. Noting this possibility early on, Seagoe (1933:31) commented that, “Within the limits set by propinquity, however, we select associates according to the satisfying of our contacts and according to the pleasure we obtain through our associates. Propinquity marks the boundaries within which we may choose; similarity conditions our choices within those boundaries. We enjoy the companionship of persons of the same age and physical maturity, of the same intelligence, and of the same disposition.” Building on these observations, one may expect to find an even stronger genetic influence on peer group GPA if respondents were, say, college students (but note that causal influences may be more prevalent among such a sample [Sacerdote, 2001]).

Recall that the current studied relied on the send measure for the peer network. This measure may be particularly influenced by the school environment. It is important to remember that the send measure only captures peers nominated by the respondent. The measure, therefore, could be tapping school-level factors such as school politics and popularity. In particular, many respondents may be nominating a group of “popular” individuals who did not nominate the respondent as a friend. If this is occurring, the heritability estimates could be reflecting the tendency for students to value other high status students, even if their peer does not reciprocate this relationship. We attempted to mitigate the likelihood that this bias influenced our conclusions by replicating the findings with the send-and-receive version of the measure. Findings were substantively similar across both measures for all of the analyses, limiting the chances of a biased finding.

A third limitation to keep in mind is that our study is unable to identify which genes influence peer group selection. Recent analyses have shown that certain genetic polymorphisms linked to the dopaminergic system are related to peer group deviance (Beaver et al., 2008; Yun et al., 2011), suggesting that these same genetic factors may correlate with peer group GPA. Future research should consider the molecular genetic linkages to peer group GPA because it may provide valuable insight into the neurological underpinnings of peer group homophily given the role of genotype in the functioning of the brain and how the social world is processed. Finally, a fourth limitation is that the nonshared environmental factors that accounted for the remainder of the variance in peer group GPA are undefined. Scholars should prioritize studies that unpack the nonshared environmental component of peer group selection.

Despite these limitations, this is one of the first studies to analyze the genetic influences on peer group GPA (but see Guo, 2006). These findings pave the way for future work to explore the role of genetics in peer group selection and, at the same time, suggest that “peer effects” may not unilaterally reflect causal influences. If peers resemble another in terms of academic ability, and if active rGE underlies the sorting of

---

3 We would like to thank an anonymous reviewer for raising this point.
individuals into certain groups, what impact might this have on long-term outcomes such as educational success, career opportunities, and even mating opportunities? It is our hope that the current findings will spark these types of discussions amongst our peers. Perhaps they can even be extended to help shed light on loosely related phenomena such as the “increasing dominance of teams in [the] production of knowledge” among the scientific community (Wuchty et al., 2007; 1036).

Acknowledgements

This research uses data from Add Health, a program project directed by Kathleen Mullan Harris and designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullen Harris at the University of North Carolina at Chapel Hill, and funded by grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 23 other federal agencies and foundations. Special acknowledgment is due Ronald R. Rindfuss and Barbara Entwistle for assistance in the original design. Information on how to obtain the Add Health data files is available on the Add Health website (http://www.cpc.unc.edu/addhealth). No direct support was received from grant P01-HD31921 for this analysis.

References

Young, J.T.N., Rebellion, C.J., Barnes, J.C., Weerman, F.W., 2013. Unpacking the black box of peer similarity in deviance: understanding the mechanisms linking personal behavior, peer behavior, and perceptions. Criminology 52, 60–86.