

Genetic influences on social network characteristics

Matthew O. Jackson¹

Department of Economics, Stanford University, Stanford CA 94305; and External Faculty, Santa Fe Institute, Santa Fe, NM 87501

Who becomes the most central individual in a society and why? What determines how many friends a given individual has? What determines how clustered or tightly-knit the friendships in a society are? In a set of important and original new findings reported in this issue of PNAS, Fowler, Dawes, and Christakis (1) provide evidence that network characteristics such as those mentioned above are heritable: that is, they show that an increase in the overlap in genetic material in twins corresponds to an increase in the covariation of some of their social network characteristics.

The heritability of network characteristics is important because of its implications for how networks form. Given that social networks play important roles in determining a wide variety of things ranging from employment and wages to the spread of disease (2–4), it is important to understand why networks exhibit the patterns that they do. Although it is well established that personal characteristics and behaviors play critical roles in determining who interacts with whom (5, 6), Fowler *et al.*'s analysis (1) suggests that genetic traits may influence individuals in terms of their social behavior, for instance, by having genetic predispositions regarding things like the tendency of an individual to introduce his or her friends to each other.

As a basis for a more detailed discussion, let me start with a brief overview of their analysis. Fowler *et al.* (1) examine the social network characteristics of 1,110 twins from the Adolescent Health Dataset (*SI Text* in ref. 1), which is based on interviews of high school students. They use standard techniques from twins studies that have been useful in identifying the heritability of a variety of traits and behaviors (7). The method is based on comparing covariation in outcomes for same-sex twins who are monozygotic (identical twins from 1 egg who share all of their genes) with that for twins who are dizygotic (fraternal twins from 2 eggs who share on average half of their genes). Presuming that the social environment that twins share is not influenced by whether they are monozygotic or dizygotic, if network characteristics are significantly more correlated among monozygotic twins than dizygotic twins then there is evidence for a genetic role in network formation.

The network characteristics that Fowler *et al.* (1) investigate are: in-degree (how many students name a given student as a friend), out-degree (how many students a

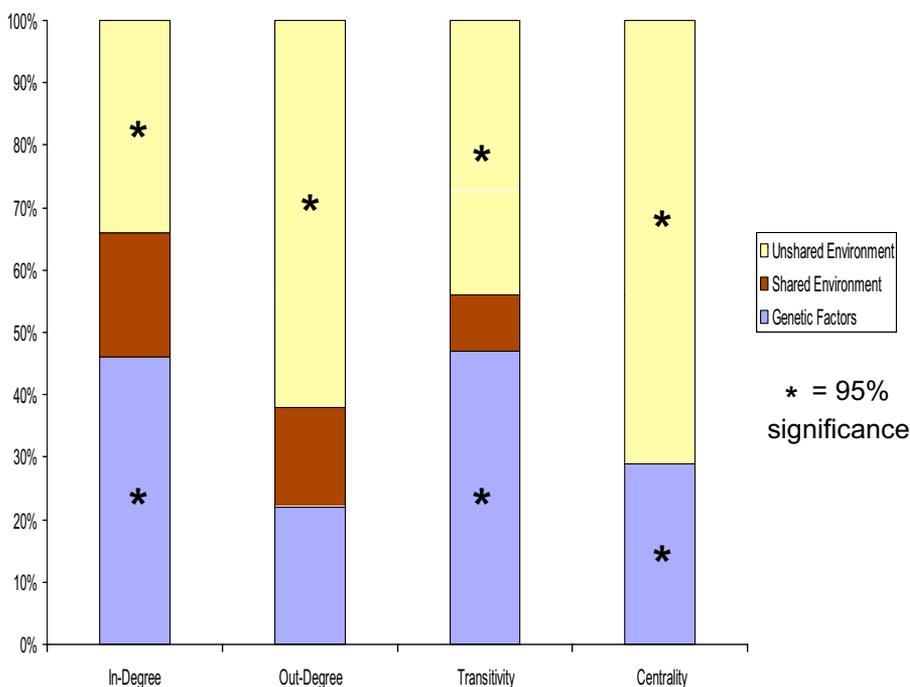


Fig. 1. Percentage of variance in network characteristics explained by genetic and environmental factors. Based on *SI Text* and Table S1 of Fowler *et al.* (1).

given student names as friends), transitivity (if A and B are friends, and B and C are friends, what is the likelihood that A and C are friends), and betweenness centrality (the fraction of shortest paths between other pairs of students that a given student lies on). Their statistical analysis assumes that the variation in a network characteristic can be additively separated into a component that is genetic, a component caused by the environment that would be shared with a twin, and a component caused by the environment that would not be shared with a twin. The covariation between monozygotic twins is then the variance caused by the common environment plus the variance caused by genetic factors, whereas the covariation between dizygotic twins is the variance caused by the common environment plus half of the variance caused by genetic factors. This formulation allows one to solve for the percentage of variation in a given network characteristic that is caused by each of the genetic, common environment, and unshared environment components, as pictured in Fig. 1, which is based on Fowler *et al.*'s *SI Text* and Table S1.

As we see in Fig. 1, Fowler *et al.* (1) find that almost half of the variation in transitivity and in-degree are genetically attributable, and more than a quarter of

betweenness centrality is genetically attributable, but the genetic component of the out-degree variation is too small to be statistically significant. It is also worth noting that the common environment is statistically insignificant in all cases.

Let me point out 2 issues here. First, we should be cautious in drawing conclusions from the observation that genetics show a statistically significant role in determining in-degree but not out-degree. The confidence interval on out-degree extends from 0 to $\approx 50\%$, which indicates that we have little idea of whether genetics really matter in the determination of out-degree. In fact, the 95% confidence intervals on the percentage of in-degree and out-degree variation attributable to genetics [(23%, 69%) and (0%, 47%), respectively] show substantial overlap and one could not reject the hypothesis that they are the same or even reversed. Second, in-degree and out-degree in the Adolescent Health Dataset are somewhat spe-

Author contribution: M.O.J. wrote the paper.

The author declares no conflict of interest.

See companion article on page 1720.

¹E-mail: jacksonm@stanford.edu.

© 2009 by The National Academy of Sciences of the USA

cial. Why might student A name student B as a friend without having B name A as a friend, which actually happens in the data more than half of the time (8)? B might simply have forgotten to name A, or B might already have reached the cap of naming at most 5 male or 5 female friends that was imposed in the Adolescent Health Dataset [each of these caps was reached by approximately a quarter of the subjects, which might not be too problematic but should be kept in mind (8)], or B and A might have different perspectives on their relationship. The first 2 reasons are forms of measurement error, while the third suggests a need to be aware of what in-degree and out-degree represent here. Given these issues, further data and analysis are needed to sharpen the conclusions of what sorts of network characteristics exhibit genetic influences, and we should be careful in interpreting differences between in-degree and out-degree.

Regardless of the issue of in- versus out-degree, the significance of genetic influence on at least some of the network characteristics is of substantial interest and suggests an array of potential explanations. The second main contribution of Fowler *et al.*'s study (1) is to examine which models of network formation produce network patterns consistent with the data. In this regard, I think it is useful to distinguish between 2 ways in which genetics might operate. The first is via passive characteristics of an individual such as "attractiveness" or some other trait that affects how others might perceive and act toward him or her, and the second is via active characteristics of the individual such as how sociable he or she is or other personality traits that might affect the individual's behavior. This is where the observation that transitivity is heritable becomes intriguing. The heritability of transitivity is a real hurdle to models that are built solely on passive characteristics. Passive characteristics can influence how likely someone is to initiate or accept a relationship with a given individual, which can impact both degree and centrality, but cannot (directly) influence how likely a

given individual's friends are to be friends with each other and thus have difficulty generating a relation between genetics and transitivity. Indeed, when Fowler *et al.* (1) test models built on passive traits such as location in a space, fitness, attractiveness, none generate a significant genetic component of the variation in transitivity (see figure 1 in ref. 1). A reason that Fowler *et al.*'s "attract and introduce" model fits the data better than the other models that they examine is that it has both passive and active heritable characteristics. Their model incorporates genetic variation in the passive trait of attractiveness (how likely other individuals are to link to a given individual), which leads to patterns of variation in in-degree and centrality that match the data, and an active trait of the propensity of an individual to introduce his or her friends to each other, which generates genetically attributable variation in transitivity.

Should we conclude that the observed patterns of networks cannot be generated without some active heritable characteristic? Although the Fowler *et al.* (1) study is suggestive in this direction, there are still things that need to be investigated before reaching such a conclusion. Some social settings exhibit a negative relationship between an individual's degree and the extent to which his or her friends are friends with each other (9, 10). That is, individuals with higher degrees tend to have lower levels of transitivity among their friends. This suggests taking another look at the Adolescent Health Dataset to determine, for example, whether the genetic component of the variation in transitivity disappears once the variation in transitivity that corresponds to the variation in degree is accounted for. The results reported in Fig. 1 examine one network characteristic at a time and do not incorporate the covariation among the network characteristics themselves, which could further refine the picture of which network characteristics are heritable and to what extent.

Finally, it may be useful to further distinguish types of heritable traits that could impact network characteristics. Both the

"attract" and "introduce" parts of the model are traits that directly influence network characteristics. There are other potentially heritable traits that could indirectly influence network characteristics. For example, attitudes and preferences appear to be at least partly heritable (11–13), and in a high school that could influence which classes an individual takes, which clubs he or she joins, which extracurricular activities he or she engages in, whether or not he or she works outside of school, and so forth. Similarly, heritable passive traits such as an individual's physical and mental capabilities may affect his or her choice of activities. A variety of heritable traits could then lead monozygotic twins to have greater overlap in their activities than dizygotic twins and thus greater overlap in their common environment. This idea is precluded in the Fowler *et al.* (1) specification where the common environment is assumed similar across types of twins, which has been a debated assumption in the literature on twin studies, with some evidence that monozygotic twins end up in greater contact with each other than dizygotic twins (see the *SI Text* in ref. 1). Given that friendship formation is influenced by contact (5), a student who joins a small club might tend to have fewer friends and more transitive friendships than one who joins a large club. If monozygotic twins are more likely to have the same activities than dizygotic twins, it could lead to increased correlation in their network characteristics, including transitivity. This is just one indirect way in which genetics could influence network characteristics; showing that there are further (complementary) mechanisms that deserve exploration.

Fowler *et al.* (1) provide an exciting new perspective on network formation that informs us about the scope of factors that influence friendship formation, opens a new set of questions about the roles of heritable traits in social network formation, and suggests new types of models to be explored.

ACKNOWLEDGMENTS. This work is supported by National Science Foundation Grant SES-0647867.

1. Fowler JH, Dawes CT, Christakis NA (2009) Model of genetic variation in human social networks. *Proc Natl Acad Sci USA* 106:1720–1724.
2. Ioannides YM, Datcher Loury L (2004) Job information networks, neighborhood effects, and inequality. *J Econ Lit* 42:1056–1093.
3. Barrat A, Barthélemy M, Vespignani A (2008) *Dynamical Processes in Complex Networks* (Cambridge Univ Press, Cambridge, UK).
4. Jackson MO (2008) *Social and Economic Networks* (Princeton Univ Press, Princeton).
5. McPherson M, Smith-Lovin L, Cook JM (2001) Birds of a

feather: Homophily in social networks. *Annu Rev Sociol* 27:415–444.

6. Kandel DB (1978) Homophily, selection, and socialization in adolescent friendships. *Am J Sociol* 14:427–436.
7. Neale MC, Cardon LR (1992) *Methodology for Genetic Studies of Twins and Families* (Kluwer, Dordrecht, The Netherlands).
8. Moody J (2001) Race, school integration, and friendship segregation in America. *Am J Sociol* 107:679–716.
9. Goyal S, van der Leij MJ, Moraga-Gonzalez JL (2006) Economics: An emerging small world. *J Polit Econ* 114:403–412.
10. Jackson MO, Rogers BW (2007) Meeting strangers and friends of friends: How random are social networks? *Am Econ Rev* 97:890–915.
11. Olson JM, Vernon PA, Harris JA, Jang KL (2001) The heritability of attitudes: A study of twins. *J Personality Soc Psychol* 80:845–860.
12. Reed DR, Bachmanov AA, Beauchamp GK, Tordoff MG, Price RA (1997) Heritable variation in food preferences and their contribution to obesity. *Behav Genet* 27:373–387.
13. Fowler JH, Baker LA, Dawes CT (2008) Genetic variation in political participation. *Am Polit Sci Rev* 102:233–248.