Genes and Social Networks

James H. Fowler
UC San Diego
Biology, Politics, and the Emerging Science of Human Nature

James H. Fowler* and Darren Schreiber

In the past 50 years, biologists have learned a tremendous amount about human brain function and its genetic basis. At the same time, political scientists have been intensively studying the effect of the social and institutional environment on mass political attitudes and behaviors. However, these separate fields of inquiry are subject to inherent limitations that may only be resolved through collaboration across disciplines. We describe recent advances and argue that biologists and political scientists must work together to advance a new science of human nature.
Turnout in a Small World

Social Logic of Politics 2005

The Watts-Strogatz (WS) Model

Order → Chaos

“Real” Social Network

Frequency

Total Change in Turnout

0% 5% 10% 15% 20%
Genetic Variation in Political Participation

JAMES H. FOWLER  University of California, San Diego
LAURA A. BAKER  University of Southern California
CHRISTOPHER T. DAWES  University of California, San Diego

\[ \rho = 0.71 \]

\[ \rho = 0.50 \]
Two Genes Predict Voter Turnout

James H. Fowler  University of California, San Diego
Christopher T. Dawes  University of California, San Diego

Partisanship, Voting, & the Dopamine D2 Receptor Gene

Christopher T. Dawes, James H. Fowler
Friendships Moderate an Association Between a Dopamine Gene Variant and Political Ideology

Jaime E. Settle, Christopher T. Dawes, Peter K. Hatemi, Nicholas A. Christakis, James H. Fowler
The Spread of Obesity in a Large Social Network over 32 Years

Nicholas A. Christakis, M.D., Ph.D., M.P.H., and James H. Fowler, Ph.D.

N ENGL J MED 357;4   WWW.NEJM.ORG   JULY 26, 2007
A Common Genetic Variant Is Associated with Adult and Childhood Obesity

Alan Herbert, Norman P. Gerry, Matthew B. McQueen, Iris M. Heid, Arne Pfeuffer, Thomas Illig, H.-Erich Wichmann, Thomas Meitinger, David Hunter, Frank B. Hu, Graham Colditz, Anke Hinney, Johannes Hebebrand, Kerstin Koberwitz, Xiaofeng Zhu, Richard Cooper, Kristin Ardlie, Helen Lyon, Joel N. Hirschhorn, Nan M. Laird, Marc E. Lenburg, Christoph Lange, Michael F. Christman

Obesity is a heritable trait and a risk factor for many common diseases such as type 2 diabetes, heart disease, and hypertension. We used a dense whole-genome scan of DNA samples from the Framingham Heart Study participants to identify a common genetic variant near the INSIG2 gene associated with obesity. We have replicated the finding in four separate samples composed of individuals of Western European ancestry, African Americans, and children. The obesity-predisposing genotype is present in 10% of individuals. Our study suggests that common genetic polymorphisms are important determinants of obesity.

Obesity is associated with an increased risk of type 2 diabetes mellitus, heart disease, metabolic syndrome, hypertension, stroke, and some forms of cancer (1). It is commonly assessed by calculating an individual’s body mass index (BMI) [weight/(height)²] in kg/m² as a surrogate measurement. Individuals with a BMI ≥ 25 kg/m² are classified as overweight, and those with a BMI ≥ 30 kg/m² are considered obese. Having a BMI over 25 kg/m² increases the risk of death (2). Presently, 65% of Americans are overweight and 30% are obese (3). Genetic factors contribute significantly to the etiology of obesity (4, 5), with estimates of the heritability of BMI ranging from 30 to 70% (6–9).

To identify common genetic variants associated with elevated BMI, we have studied individuals from the National Heart, Lung, and Blood Institute (NHLBI)–Framingham Heart Study (FHS) (10). The participants were enrolled from the community without being selected for a particular trait or disease and were followed over 24 years (table S1). In this population, heritability estimates for BMI range between 37 and 54% (11, 12).

Using families from this sample, we performed a genome-wide association analysis, using a testing strategy for quantitative traits in...
The Collective Dynamics of Smoking in a Large Social Network

Nicholas A. Christakis, M.D., Ph.D., M.P.H., and James H. Fowler, Ph.D.

N ENGL J MED 358;21  WWW.NEJM.ORG  MAY 22, 2008
Genetic influence on smoking--a study of male twins

*D Carmelli, GE Swan, D Robinette, and R Fabsitz*

Abstract

BACKGROUND. The results of twin and family studies suggest that heredity has a small influence on smoking behavior. METHODS. We conducted a genetic analysis of several aspects of smoking behavior among subjects in the National Academy of Sciences-National Research Council Twin Registry. The registry includes male twins who were born in the United States between 1917 and 1927 and who were members of the armed services during World War II. Information on smoking history was available for 4775 pairs of twins, who were first surveyed in 1967 through 1969, when they were 40 to 50 years old, and then re-surveyed in 1983 through 1985, when they were 58 to 68. Eighty percent of the subjects in this cohort had smoked at some time in their lives, 60 percent were smokers in 1967 through 1969, and 39 percent were smoking in 1983 through 1985. Similarities between twins in smoking habits at baseline and at the second follow-up 15 years later were examined. The comparison of concordance for smoking between monozygotic and dizygotic twins was used to assess the relative contribution of familial and genetic factors. RESULTS. In 1967-1969 survey the ratio of observed to expected concordance for smoking was higher among the monozygotic twins than among the dizygotic twins for those who had never smoked (overall rate ratio, 1.38; 95 percent confidence interval, 1.26 to 1.54), for former smokers (overall rate ratio, 1.35, 95 percent confidence interval, 1.25 to 1.55), for current cigarette smokers (overall rate ratio, 1.16; 95 percent confidence interval, 1.11 to 1.26), and for current cigar or pipe smokers (overall rate ratio, 1.60; 95 percent confidence interval, 1.22 to 2.06). The data also suggest genetic influences on quitting smoking. Monozygotic twins were more likely than dizygotic twins to be concordant for quitting smoking (overall rate ratio, 1.24; 95 percent confidence interval, 1.06 to 1.42). CONCLUSIONS. In this cohort of adult male twins, there were modest genetic influences on lifetime smoking practices.
Dynamic spread of happiness in a large social network: longitudinal analysis over 20 years in the Framingham Heart Study

James H Fowler, associate professor,¹ Nicholas A Christakis, professor²

BMJ 2008;337:a2338
HAPPINESS IS A STOCHASTIC PHENOMENON

David Lykken and Auke Tellegen

University of Minnesota

Abstract—Happiness, or subjective well-being, was measured on a birth-record-based sample of several thousand middle-aged twins using the Well-Being (WB) scale of the Multidimensional Personality Questionnaire. Neither socioeconomic status, educational attainment, family income, marital status, nor an indicator of religious commitment could account for more than about 3% of the variance in WB. From 44% to 52% of the variance in WB, however, is associated with genetic variation. Based on the retest of smaller samples of twins after intervals of 4.5 and 10 years, we estimate that the heritability of the stable component of subjective well-being approaches 80%.

did not reach the eighth grade, whereas others have doctorates; they live on farms, in small towns, in big cities, and in foreign lands; their socioeconomic levels are representative of Minnesota-born adults.

METHOD

A self-rating questionnaire was administered to 2,310 members of this twin registry. One of the questionnaire items read as follows:
A simple rule for the evolution of cooperation on graphs and social networks

Hisashi Ohtsuki\textsuperscript{1,2}, Christoph Hauert\textsuperscript{2}, Erez Lieberman\textsuperscript{2,3} & Martin A. Nowak\textsuperscript{2}
Heritability of cooperative behavior in the trust game

David Cesarini*, Christopher T. Dawes†, James H. Fowler‡‡, Magnus Johannesson§, Paul Lichtenstein‖, and Björn Wallace‡

*Department of Economics, Massachusetts Institute of Technology, 50 Memorial Drive, Cambridge, MA 02142; †Political Science Department, University of California at San Diego, 9500 Gilman Drive 0521, La Jolla, CA 92037-0521; §Department of Economics, Stockholm School of Economics, Box 6501, SE-113 83 Stockholm, Sweden; and ‖Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Box 281, SE-171 77 Stockholm, Sweden

Edited by Elinor Ostrom, Indiana University, Bloomington, IN, and approved January 15, 2008 (received for review October 23, 2007)

Although laboratory experiments document cooperative behavior in humans, little is known about the extent to which individual differences in cooperativeness result from genetic and environmental variation. In this article, we report the results of two independently conceived and executed studies of monozygotic and dizygotic twins, one in Sweden and one in the United States. The results from these studies suggest that humans are endowed with genetic variation that influences the decision to invest, and to reciprocate investment, in the classic trust game. Based on these findings, we urge social scientists to take seriously the idea that differences in peer and parental socialization are not the only forces that influence variation in cooperative behavior.
Behaviors
- Participation
- Obesity
- Smoking
- Happiness
- Cooperation

Social Networks

Big Picture
Big Picture

Supergenes
- Dopamine
- Serotonin

Behaviors
- Participation
- Obesity
- Smoking
- Happiness
- Cooperation
**Big Picture**

**Supergenes**
- Dopamine
- Serotonin

?  

**Social Networks**

?  

**Behaviors**
- Participation
- Obesity
- Smoking
- Happiness
- Cooperation
Today’s Piece of the Puzzle

- Twin Study Method
- Network Structure is Heritable
- Reject All Existing Models of Human Social Networks
- Develop Alternative “Attract & Introduce” Model
Twin Studies

• Compare the behavior of
  - monozygotic (MZ) twins (identical)
    • share 100% of their genes
  - dizygotic (DZ) twins (fraternal)
    • share 50% of their genes on average

• Decompose variance
  - A - genetic
  - C - common environment
  - E - unshared environment
Twin Study Method

- **Structural Equations Model (SEM) to estimate**

\[
\begin{align*}
\sigma_P^2 &= \sigma_A^2 + \sigma_C^2 + \sigma_E^2 \\
COV_{MZ} &= \sigma_A^2 + \sigma_C^2 \\
COV_{DZ} &= 0.5\sigma_A^2 + \sigma_C^2
\end{align*}
\]

- **P** = total variance
- **A** = genetic
- **C** = common environment
- **E** = unshared environment
- **MZ** = monozygotic (identical) twins
- **DZ** = dizygotic (fraternal) twins
Criticism of Twin Studies

- MZ and DZ environments may not be comparable
  - MZ twins may be more strongly affiliated than DZ twins

- However
  - Studies of twins reared apart validated by other methods
    Bouchard 1998; Visscher et al 2006
  - Differences between MZ and DZ twins persist even among twins whose zygosity has been miscategorized by their parents
    Bouchard and McGue 2003
  - MZ twins are sometimes in more frequent contact, but this results from rather than causes greater similarity
    Posner et al 1996
  - MZ twins living apart become more similar with age
    Bouchard and McGue 2003
Add Health Data

- National Longitudinal Study of Adolescent Health
  - 3 waves of health, social network, and genetic data
  - In-school sample
    - 90,118 students in 142 schools
  - Name up to 5 male and 5 female friends
  - Oversampled twins
    - 307 MZ pairs
    - 248 same-sex DZ pairs
Egocentric Social Network Measures

- in-degree (0 to N - 1)
- out-degree (0 to 10)
  - 10% name max.
  - mean = 3.8, SD = 3.7
- transitivity (0 to 1, mean=0.18)
  - \( \Pr( j \leftrightarrow k \mid i \leftrightarrow j, i \leftrightarrow k ) \)
- betweenness centrality
  - \( \sum_{i \neq j \neq k} \frac{\sigma_{ijk}}{\sigma_{ik}} \)
## Twin Study Results

<table>
<thead>
<tr>
<th>Model</th>
<th>Proportion of Variance Explained by</th>
<th>Model Fit (-2LL)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Genetic Factors</td>
<td>Common Environment</td>
</tr>
<tr>
<td>In-Degree</td>
<td>0.46 (0.23, 0.69)</td>
<td>0.21 (0.00, 0.40)</td>
</tr>
<tr>
<td>Transitivity</td>
<td>0.47 (0.13, 0.65)</td>
<td>0.09 (0.00, 0.36)</td>
</tr>
<tr>
<td>Betweenness Centrality</td>
<td>0.29 (0.05, 0.39)</td>
<td>0.00 (0.00, 0.19)</td>
</tr>
<tr>
<td>Out-degree</td>
<td>0.22 (0.00, 0.47)</td>
<td>0.16 (0.00, 0.40)</td>
</tr>
<tr>
<td>Out-degree, dropping those who name 10 friends</td>
<td>0.00 (0.00, 0.39)</td>
<td>0.44 (0.09, 0.52)</td>
</tr>
</tbody>
</table>

*Note: 95% confidence intervals indicated in parentheses beneath each estimate. First four models based on 307 monozygotic and 248 dizygotic same-sex twin pairs. The last model drops subjects who named the maximum 10 possible friends, yielding 256 monozygotic and 204 dizygotic same-sex twin pairs. Network measures were transformed to have zero mean and unit variance within each school network to prevent differences between schools from influencing the results.*
A (Very) Brief History of Social Network Models

• Focus on generating variation via endogenous processes acting on identical individuals
  - Physics: scale-free model
  - Economics: connections model
  - Sociology: homophily, structural balance models

• Ignore intrinsic heterogeneity

• Our most intrinsic characteristics are genes
Candidate Social Network Models

- Only models with intrinsic node characteristics can generate significant heritability

- Possibilities include:
  - “Fitness” model (Park-Barbási 2007)
  - “Social Space” model (Boguñá et al. 2004)
  - “ERGM” exponential random graph model (Snijders et al. 2006)
  - Erdos-Renyi model used as a baseline

- How to test?
The Mirror Network Method

1. Generate first set of intrinsic node characteristics
2. Simulate network from first set
3. Generate second set of intrinsic node characteristics
4. Sample one individual from first set and one individual from second set (the “twins”)
The Mirror Network Method

5. Copy node characteristics from sampled twin from first set to sampled twin from second set
6. Simulate “mirror” network from second set
7. Calculate and store egocentric network measures for each twin
8. Repeat 10000 times, calculate twin correlation
Heritability Comparison

Percent Variance Explained by Intrinsic Node Characteristics (Heritability)

Real Data (Add Health)
Social Space Model
Fitness Model
Exp. Random Graph Model
Erdos-Renyi Model

In-Degree
Transitivity
Centrality
Out-Degree
Attract & Introduce Model

- N nodes and E edges
- Each node endowed with two characteristics
  \[
  \Pr\left( p_{\text{attract}}^j \sim \text{Uniform}[0,1] \right) = \alpha \quad \text{and} \quad \Pr\left( p_{\text{attract}}^j = 0 \right) = 1 - \alpha
  \]
  \[
  \Pr\left( p_{\text{introduce}}^j = 1 \right) = \beta \quad \text{and} \quad \Pr\left( p_{\text{introduce}}^j = 0 \right) = 1 - \beta
  \]
- Probability \( p_{\text{attract}}^i \) a social tie from i to j forms
  - If so, then with probability \( p_{\text{introduce}}^i \), i introduces j to all of his “friends”
    - If so, then each friend sends a tie to j with probability \( p_{\text{attract}}^j \) and j sends a tie to each kth friend with probability \( p_{\text{attract}}^k \)
- Repeat until at least E ties are generated
Heritability Comparison

Percent Variance Explained by Intrinsic Node Characteristics (Heritability)

- In-Degree
- Transitivity
- Centrality
- Out-Degree

Heritability Comparison

- Real Data (Add Health)
- Attract & Introduce Model
- Social Space Model
- Fitness Model
- Exp. Random Graph Model
- Erdos-Renyi Model
Degree Distributions

Real Data (Add Health)
Attract and Introduce
Social Space
Fitness
ERGM
Erdos-Renyi
Node Degree vs. Transitivity

Mean Node Transitivity vs. Node Degree

- Real Data (Add Health)
- Attract & Introduce Model
- Social Space Model
- Fitness Model
- Exp. Random Graph Model
- Erdos–Renyi Model
Modularity Comparison

\[ Q = \frac{1}{2m} \sum_{ij} \left[ A_{ij} - \frac{k_i k_j}{2m} \right] \delta(g_i, g_j) \]

Degree-degree correlation = 0.18
k-Motif Fingerprint Method

- Use statistical distribution of patterns among k nodes (motifs) to identify networks
  - 16 3-motifs; 218 4-motifs
- Simulate network 100 times
- Calculate mean and variance of each motif
- Generate multivariate beta density
  \[
  \alpha = \frac{\mu^2(1-\mu)}{\sigma^2 - \mu}, \quad \beta = \frac{\mu(1-\mu)^2}{\sigma^2 - (1-\mu)}
  \]
- Calculate likelihood from densities
- Correctly identified candidate network in 10,000 tests
- Multinomial Dirichlet better but convergence difficult
Motif Fingerprint Comparison

3-motifs

4-motifs

Network Size (Number of Nodes)

Adjusted Log–Log Likelihood

Symbols:
- Blue: Real Data (Add Health)
- Green: Attract & Introduce Model
- Orange: Social Space Model
- Red: Fitness Model
- Pink: Exp. Random Graph Model
- Light Blue: Erdos–Renyi Model

Network Size (Number of Nodes)
Discussion

• Evidence rejects homogenous node models
• Attract & Introduce model simple, effective
• We should search for specific genes
  - 5-HT2A associated with “popularity” (Burt 2008)
• We should search for specific intermediating causal mechanisms
  - Personality (extraversion)
Discussion

• Consider coevolution of networks & cooperation in early human society
  - Heritability of network structure means either
    • networks have no impact on fitness (drift)
    • networks affect fitness via frequency dependent selection
  - Natural selection may ensure diversity of network types
• Networks might be a link from genes to behavior
  - What impact does this causal pathway have on fitness?
Before the 17th Century
The Solar System

Galileo Galilei

<table>
<thead>
<tr>
<th>Date</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. Nov.</td>
<td></td>
</tr>
<tr>
<td>3. Nov.</td>
<td></td>
</tr>
<tr>
<td>4. Nov.</td>
<td></td>
</tr>
<tr>
<td>5. Nov.</td>
<td></td>
</tr>
<tr>
<td>7. Nov.</td>
<td></td>
</tr>
<tr>
<td>8. Nov.</td>
<td></td>
</tr>
<tr>
<td>10. Nov.</td>
<td></td>
</tr>
<tr>
<td>11. Nov.</td>
<td></td>
</tr>
<tr>
<td>12. Nov.</td>
<td></td>
</tr>
</tbody>
</table>
The Solar System

Galileo Galilei
Biological Systems

Anton van Leeuwenhoek
Biological Systems

Anton van Leeuwenhoek
Human Social Systems

18th Century
Human Social Systems

- Bourgeoisie
- Proletariat
- Nobility

19th Century
Human Social Systems

20th Century
Human Social Systems

21st Century
Computational Social Science

David Lazer,1* Alex Pentland,2 Lada Adamic,3 Sinan Aral,2,4 Albert-László Barabási,5 Devon Brewer,6 Nicholas Christakis,1 Noshir Contractor,7 James Fowler,8 Myron Gutmann,3 Tony Jebara,9 Gary King,1 Michael Macy,10 Deb Roy,2 Marshall Van Alstyne11

www.sciencemag.org  SCIENCE  VOL 323  6 FEBRUARY 2009
Acknowledgements

Chris Dawes

Nicholas Christakis

Come see us at http://jhfowler.ucsd.edu