

## Genetic Effects on Physical Health: Lower at Higher Income Levels

Wendy Johnson,<sup>1,2</sup> and Robert F. Krueger<sup>1,2</sup>

Received 9 Sep. 2004—Final 15 Feb. 2005

---

Given the robust finding that people in higher income groups tend to experience better physical health, there is interest in identifying mechanisms underlying this gradient. Using a nationwide sample of 719 twin pairs from the National Survey of Midlife Development in the United States, we investigated the possibility that gene-environment interaction underlies the income-health gradient. We observed that genetic variance associated with 2 measures of physical health, number of chronic illnesses and body mass index, each declined significantly with increasing income. This interaction effect could not be removed by adjusting income for the presence of health insurance coverage and education, suggesting that the interaction is not simply a result of differences in levels of those characteristics with income.

---

**KEY WORDS:** Environmental influences; genetic influences; income; physical health; twin study.

People in higher income groups tend to have better physical health. This relationship is well established throughout history, across geographical boundaries, and for almost every disease and condition (Adler *et al.*, 1994; Antonovsky, 1967). Though it may seem obvious that severe poverty might erode physical health through the effects of poor nutrition, crowded and dirty living conditions, and inadequate medical care, the association exists across the income range. Not only do those just above the poverty line have better health than those in poverty, but those in the highest income levels have better health than those just below them (Marmot *et al.*, 1991). In addition, the relationship cannot be explained by lack of access to health care, and it exists even in populations with universal access to medical care (Adler and Snibbe, 2003). The effect is greatest for the poorest groups, yet in the United States it has the greatest impact on the middle classes because the largest numbers of

people are in those income ranges (Adler and Snibbe, 2003).

Adler *et al.* (1994) summarized findings to date about the relationship between income and health, and examined possible explanations for the basis of the association. The first possibility they considered was that the relationship results from common underlying genetically based factors. They suggested that, for example, physical size or intellectual capacity could contribute directly to both income and physical health, resulting in a spurious association between the two effects. They dismissed this possibility as unlikely, however, pointing out that the association between job status and health persists after adjustment for height and body mass index (Marmot *et al.*, 1991), and that intellectual capacity does not appear to be reliably linked to health. In addition, they noted that any genetic predispositions involved would probably be important only when environmental and behavioral factors impinged on them.

Since then, most research on possible mechanisms accounting for the relationship has focused on another explanation suggested by Adler *et al.* (1994): that income influences biological functions that in turn affect health status. Generally, the process that

---

<sup>1</sup> Department of Psychology, University of Minnesota – Twin Cities, 75 East River Road, Minneapolis, MN 55455, USA.

<sup>2</sup> To whom correspondence should be addressed at Department of Psychology, 75 East River Road, Minneapolis, MN 55455. Fax: +1-612-626-2079; e-mail: john4350@umn.edu; krue9038@umn.edu.

subsequent research has begun to articulate (Adler and Snibbe, 2003) centers around the increasing demands and decreasing resources for dealing with those demands associated with lower levels of income. This results in greater exposure to stress at lower income levels, as well as in greater psychological response to that stress. As Gallo and Mathews (2003) have suggested, the combination of greater stress and stress reactivity among individuals of lower income reduces reserve capacities for responding to environmental challenges, and, over time, may make them more vulnerable to disease. Identified sources of greater stress and environmental challenge among lower income groups include more physically risky jobs, greater exposure to toxins, pathogens, noise, crowding, conflict and crime, and decreased access to social support, recreational facilities, health care, and healthy food (Adler and Snibbe, 2003). Identified sources of greater stress reactivity include lower levels of perceived control and mastery (Lachman and Weaver, 1998), reduced reasons for optimism (Fiscella and Franks, 1997) and greater reasons for hopelessness and hostility (Gump *et al.*, 1999), in part because of peoples' perceptions of relative social standing (Adler *et al.*, 2000).

Though these associations seem plausible, they do not explain the physiological mechanisms that link the psychological experiences associated with low income to the experience of physical health. McEwen (1998) identified a combination of indicators of biological dysregulation he termed "allostatic load" resulting from such stressors that over time may contribute directly to disease. The indicators involved include systolic and diastolic blood pressure, waist-to-hip ratio, HDL and LDL cholesterol, blood glycosylated hemoglobin (an indicator of glucose levels over the past 2–3 months), and the hormones cortisol, dehydroepiandrosterone, epinephrine, and norepinephrine. Allostatic load among older adults appears to increase with decreasing income, and to be associated with general physical and cognitive decline, cardiac events, and mortality (e.g., Karlamangla *et al.*, 2002; Singer and Ryff, 1999).

The indicators involved in allostatic load are generally acknowledged to be under genetic influence (e.g., Komaroff, 1999) as well, and this is also true of many of the most common chronically debilitating physical health conditions, including heart disease, arthritis, many cancers, and diabetes. These are also among the diseases for which the income-health relationship is strongest (Adler and Snibbe, 2003). This highlights the fact that genetic influences are likely involved in the process by which income

influences biological functions that in turn affect health status. In turn, it suggests that, though Adler *et al.* (1994) may have been correct to dismiss the relatively straightforward possibility of common genetic factors underlying the relationship between income and physical health, the actual manner in which genetic influences are involved is more complex.

Genetic influence on health is commonly thought of in two basic ways. First, where specific genes have been identified (e.g., in the cases of diabetes, sickle cell anemia, or breast cancer), the presence of those genes in an individual's genotype and the frequency of those genes within a population subgroup are considered to confer increased liability to illness. Second, most analytic models of quantitative genetic processes that do not specify individual genes are based on the assumption that genetic and environmental factors operate independently and thus add together to produce an overall risk of illness. Both of these ways of thinking about genetic influence gloss over the possibility of gene-environment interaction: that genes may moderate the pathogenic influence of environmental stressors. When this is true, the frequently discussed single heritability statistic is no longer adequate to describe the situation for the entire population, as genetic effects depend on the environments of the individuals within that population. For example, overall heritability of 50% might conceal a situation in which, for half the population with one set of environmental circumstances, the condition is completely genetically determined, while for the other half with a different set of environmental circumstances, the condition is completely environmentally determined.

Data from humans and from studies of experimental plants and animals not only corroborate the possibility of such gene-environment interactions, but also suggest their direction. *Arabidopsis thaliana*, a flowering plant, shows greater genetic variance for leaf number and flowering date in conditions of low light (Stratton, 1998). More directly relevant, genes for metabolic efficiency enabling adaptation to biologically stressful environments play a primary role in affecting lifespan and, by implication, health in both humans and experimental animals (Parsons, 2003). In *Drosophila melanogaster* (the common fruit fly), for example, genetic variability for fitness, and especially mortality, increases in situations of high biological stress (Parsons, 2002). Genes for disease susceptibility also show amplified effects in the presence of triggering environmental risk factors (Tiret, 2002). For example, possession of a specific gene in certain rats is

associated with significantly greater adiposity, glucose intolerance, circulating leptin levels, and blood pressure during high-fat-diet feeding, but not during normal-diet feeding (Pauzova *et al.*, 2003). Similar effects have been observed with respect to genes associated with diabetes in humans (Weiss *et al.*, 2002). It thus seems reasonable to propose that a gene-environment interaction could underlie the observed phenotypic relationship between income and physical health. That is, we hypothesize that genetic variance associated with physical health decreases with increasing levels of income. The purpose of this study was to investigate this possibility in a nationwide US sample of twins aged 25–74. As part of this investigation, we also measured the extent to which common genetic and environmental influences contribute to both income and physical health, thus testing for the existence of common genetic influences whose possibility has been dismissed in the past (Adler *et al.*, 1994).

## METHOD

### Sample

The twin sample used in this study was gathered as part of the MacArthur Foundation Survey of Midlife Development in the United States (MIDUS). Two research organizations, ICR/AUS Consultants and Bruskin Associates, recruited twin pairs by making telephone calls and asking respondents whether they or any of their immediate family members were members of intact twin pairs. About 50,000 households, constituting a representative national sample, were screened in this manner. The 14.8% of respondents who reported the presence of a twin in the family were then asked whether it would be acceptable for the research team to contact the twins to solicit their participation in the survey. The 60% of respondents who gave such permission were referred to the overall MIDUS recruitment process, with comparable participation rates (see Kendler *et al.*, 2000, for additional details). The sample is highly select when compared to the originally contacted 50,000 households, but few of these even had twins among their relatives, and many of the ones that did hesitated to provide contact information for relatives outside their households to the research recruiters. The resulting MIDUS twin sample has demographic characteristics generally comparable to those of the main MIDUS sample.

Both members of each twin pair participating in the study met both overall study eligibility criteria and criteria specific to the twin sample. These criteria included (1) being at least first degree relatives of the original contact or his or her spouse or partner, (2) being between the ages of 25 and 74, (3) having a residential telephone number, (4) living in the continental United States, (5) speaking English, and (6) being mentally and physically able to complete the interview and questionnaires. The base MIDUS twin sample resulting from this process consists of 998 pairs.

Zygoty for each twin pair was determined using self-report questions regarding information such as similarity of eye and hair color and degree to which others were confused as to their identity during childhood. Such techniques are generally more than 90% accurate (Christensen *et al.*, 2003; Lykken *et al.*, 1990), though some pairs in this sample were not considered classifiable on the basis of the self-reports. We made use of the 719 same-sex pairs for which we had income, health, and zygosity data for the current study, resulting in 172 monozygotic (MZ) male pairs, 195 MZ female pairs, 138 dizygotic (DZ) male pairs, and 214 DZ female pairs. We thus excluded 262 opposite-sex pairs and 17 pairs with missing or indeterminate zygosity information from the full MIDUS twin sample of 998 pairs. We excluded the opposite-sex pairs because the sample was too small to test for sex differences in expression of the genetic influences on the variables under study. Table I shows demographic information for the sample we used for this study.

### Measure of Income

The self-administered mailed questionnaire booklets that were part of the MIDUS survey included several questions regarding income. To answer these questions, participants indicated into which of 36 income categories their own personal income, their spouse's or partner's income, their income from Social Security retirement benefits, their income from government assistance programs, and the income from all other sources fell. A total household income composite was computed based on the income reported from these five sources. The income distribution for the sample used is summarized in Table II. Our sample was slightly more wealthy than the country as a whole, but about 30% had income below the national median at the time the survey was conducted, and very low incomes were

**Table I.** Sample demographic information

Category	Percent
<i>Sex</i>	
Male	43.1
Female	56.9
<i>Race</i>	
White	81.2
Black/African American	3.7
Others	1.7
Not reported	13.5
<i>Age</i>	
25–34	23.2
35–44	28.6
45–54	26.5
55–64	13.7
65–74	7.9
<i>Education</i>	
Under 12 years	9.8
12 years	31.2
Over 12 years	55.1
Not reported	3.8
<i>Marital status</i>	
Married	68.8
Not married	27.0
Not reported	4.2
<i>Insurance coverage</i>	
Yes	79.8
No	20.2

substantively represented. The relative sparsity of participants living in poverty had the effect of making it more difficult to detect the hypothesized relationships. To make the income distribution more normal, we log-transformed the income variable. We made use of household income rather than own personal

**Table II.** Frequency distribution of income

Income level	N	%
\$0-\$10,000	74	5.1
\$10,001–\$20,000	124	8.6
\$20,001–\$35,000	247	17.2
\$35,001–\$50,000	353	24.5
\$50,001–\$75,000	350	24.3
\$75,001–\$100,000	131	9.1
\$100,000–\$150,000	85	5.9
\$150,000+	74	5.1
Total	1438	100.0
Mean	\$57,347	
Median	47,097	
National mean	\$44,938	
National median	34,076	

Note: National data are from the 1995 Current Population Survey.

income because we believed that household income more accurately represented the social milieu in which the participants lived.

### Measures of Physical Health Problems

The MIDUS self-administered questionnaire booklets also included several questions regarding physical health. We elected to make use of two indicators of overall health in order to assess the degree to which the gene-environment interaction effect that we hypothesized could be replicated within our sample. Participants indicated whether or not they had been treated for each of 29 chronic health problems in the past 12 months. The health problems included asthma, high blood pressure, diabetes, thyroid disease, migraine headaches, tuberculosis, hay fever, arthritis, stomach problems, multiple sclerosis, stroke, hernia, gall bladder trouble, chronic sleeping problems, and ulcer. Only 3.3% of the twins lived together, so the co-twins of participants with infectious chronic conditions such as tuberculosis were not likely to be at

**Table III.** Frequency distributions of physical conditions

Physical condition	N	%
<i>Chronic Illnesses</i>		
0	501	34.8
1	300	20.9
2	230	16.0
3	156	10.8
4	80	5.6
5	60	4.2
6	47	3.3
7	23	1.6
8	17	1.2
9	10	0.7
10+	14	1.0
Mean	1.9	
Median	1.0	
Standard deviation	2.2	
<i>Body mass index</i>		
20 or less	88	6.1
20+ to 23	243	16.9
23+ to 25	210	14.6
25+ to 27	231	16.1
27+ to 29	171	11.9
29+ to 31	112	7.8
31+ to 33	67	4.7
33+	116	8.1
Missing	200	13.9
Mean	26.2	
Median	25.6	
Standard deviation	5.1	

unduly elevated risk. We summed positive responses to form a total score. To make the variable more normal, it was log-transformed. Participants also indicated current height and weight. We used this to compute body mass index (BMI;  $\text{kg}/\text{m}^2$ ). The distributions of these variables are summarized in Table III. There were 200 participants who did not report height and/or weight, making it impossible to compute BMI. As we modeled the raw data, however, this loss was not as critical as it might have been because we still could make use of these participants' data on number of chronic illnesses and income.

Our rationale for the use of BMI as a measure of physical health problems was as follows. Though not necessarily an indicator of current physical health, excessive weight is a well-known risk factor for the development of a number of physical health problems including diabetes, high blood pressure, heart disease, certain kinds of cancer, and stroke (Komaroff, 1999). BMI above 25 is generally considered to be overweight; above 30 is generally considered obese, and health risks increase with increasingly excessive weight. Within normal ranges, health risks associated with greater BMI are not well agreed upon, nor are the existence and nature of health risks associated with low BMI. Still, given the problem of obesity in the United States in general and the fact that 56% of our sample could be considered overweight, it seemed a simple and relatively objective measure of overall health. In particular, only 6.1% of our sample had BMI's less than 20, and low BMI would generally not be expected to contribute to poor health until it dropped below at least 18. Thus there was little risk of systematic reversals in the relationship between BMI and health problems in our sample. The correlation between the number of chronic illnesses and BMI was 0.17, ( $p < 0.001$ ). The relatively low correlation was probably at least partly because the chronic illness measure did not include treatment for heart disease (it also did not include cancer). The two measures thus tapped two aspects of physical health, and neither one assessed it completely. The imperfections in our measurements had the effect of making it more difficult to detect the hypothesized relationships.

### Measures of Insurance Coverage and Education

In completing the MIDUS questionnaire booklets, participants indicated whether or not they had health insurance coverage from various sources, including private insurance, insurance through a cur-

rent or former employer of either the participant or spouse, insurance through a union, medicare and Medicaid, and insurance plans for military personnel or veterans. We collapsed these responses into a single variable reflecting presence or absence of health insurance coverage from any source. Participants also indicated their attained level of education, choosing from categories ranging from less than high school education to attainment of graduate/professional degrees.

### Analytical Approach

*Age adjustment.* Because the co-twins in our sample are the same age and sex, age and sex effects act to increase twin similarity (McGue and Bouchard, 1984). To correct for this in the analysis of genetic and environmental influences on the twin data, we regressed out the effects of age, age<sup>2</sup>, sex, and sex  $\times$  age on each variable prior to fitting our biometric models. Prior to making this adjustment, we regressed numbers of chronic illnesses and BMI on sex, income, and the interaction of sex and income. The interaction terms were not significant.

*Biometric modeling.* The standard univariate quantitative genetic model is based on the understanding that the observed (phenotypic) variance ( $V_p$ ) in a trait is a linear function of genetic ( $A^2$ ) and shared ( $C^2$ ) and non-shared ( $E^2$ ) environmental variance, respectively. Genetic variance reflects variation in genotypes transmitted directly from parents to offspring that result in phenotypic variance in the trait. Shared environmental variance (e.g. variation in neighborhoods) reflects variation in environments that affects all children growing up within a family to the same degree and differentiates among families. Such effects may clearly extend beyond childhood. Non-shared environmental variance (e.g. differential parental treatment) reflects variation in environments that has different effects on individual family members. Error variance is also included with non-shared environmental variance because the nonshared environmental variance terms are the residuals after the effects of additive genetic and shared environmental influences have been estimated. Symbolically,

$$V_p = A^2 + C^2 + E^2. \quad (1)$$

Under this model, each of the variance components is independent of the others. To investigate the possibility that genetic variance associated with physical health might decline with increasing levels of income, we needed a model in which the variance components themselves vary as a continuous function

of income. That is, rather than modeling physical health problems using the simple linear equation shown above, we made use of variance component models for gene-environment interaction described by Purcell (2002) to express the variance associated with each of the three components as a linear interaction with income, leading to the equation

$$V_p = (A + \beta_a * \text{Income})^2 + (C + \beta_c * \text{Income})^2 + (E + \beta_e * \text{Income})^2 \quad (2)$$

In this equation, genetic variance  $A^2$  from Equation (1) is expressed as  $(A + \beta_a * \text{Income})^2$ ,  $C^2$  is expressed as  $(C + \beta_c * \text{Income})^2$ ,  $E^2$  is expressed as  $(E + \beta_e * \text{Income})^2$ . The  $\beta$ 's are parameter estimates that allow the variance components to differ across income levels. Thus, for example, if genetic variance ( $A^2$ ) at the mean level is 0.5 and  $\beta_a$  is  $-0.1$ , then people with incomes 1 standard deviation above the mean would have genetic variance of  $0.37(\sqrt{0.50} - 0.1)^2$ , or while those with incomes 1 standard deviation below the mean would have genetic variance of 0.65. The model for one twin is illustrated in Figure 1.

There are two things that are important to note about this model. First, the value of Income generally

differed for twins within a pair because we made use of adult household income, and, unless the twins actually lived together at the time of the MIDUS survey, most would differ. In fact, only 15 pairs within the sample reported the same income. The model incorporates the individual data for each member of each pair directly, in the same way that it incorporates individual scores on the physical health and BMI measures. Second, though previous researchers (e.g., Adler *et al.*, 1994) have dismissed it, the possibility remains that genes are involved in the relationship between physical health problems and income not through gene-environment interaction (or genetic control of sensitivity through the environment), but through gene-environment correlation (or the genetic control of exposure to different environments). The model can be adapted to allow for the measurement of gene-environment correlation of this form. The adaptations allow for gene-environment correlation by explicitly reflecting the possibility that each component of variance (each linear term that is squared in the equation above) can be further decomposed into genetic and shared and nonshared environmental components common to income and physical health, and similar components unique to physical health. The resulting model thus provides

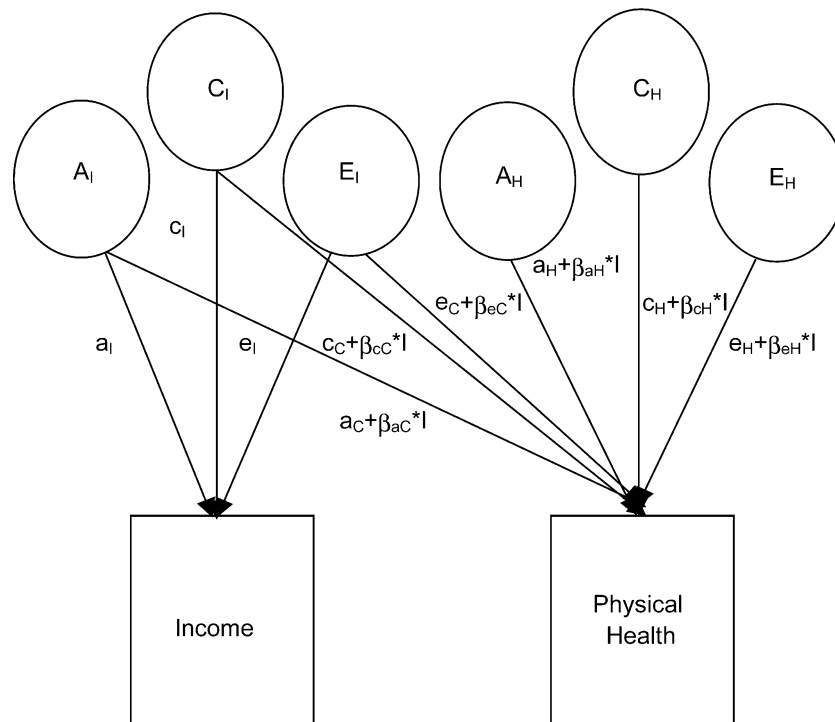


Fig. 1. Partial path diagram of the gene-environment interaction model for one twin. C refers to effects common to both income and health, and I refers to income.

estimates of the genetic and environmental variance common to both income and physical health and the extent to which these vary with income, and the genetic and environmental variance unique to physical health and the extent to which these vary with income. In the presence of gene-environment interaction, the gene-environment correlation will vary as a function of income. Its genetic component can be expressed as

$$r_G = (A_C + \beta_{aC} * Income) / ((A_C + \beta_{aC} * Income)^2 + (A_H + \beta_{aH} * Income)^2)^{1/2}, \tag{3}$$

where the subscripts C and H refer to the components of variance common to the moderator and physical health and unique to physical health, respectively. The shared and nonshared environmental components can be expressed analogously.

Throughout, we conducted our analyses of twin data using maximum likelihood estimation of the structural equation model fit to raw data as operationalized in the computer program Mx (Neale *et al.*, 1999). We assessed model fit using the likelihood ratio comparison of 2 likelihood statistics, which is distributed as  $\chi^2$  and Akaike's Information Criterion (AIC; Akaike, 1983). AIC is defined as the  $-2 * \log$ -likelihood value for the model, plus 2 times the number of parameters. Models having smaller AIC's are preferred.

To address the possibility that any gene-environment interaction we observed merely reflected differences in access to care, use of care, or broad understanding of the importance of care, we conducted our analyses using income, income adjusted for the presence or absence of insurance coverage, and income adjusted for both insurance coverage and level of education.

**RESULTS**

**Preparatory Analyses**

Prior to carrying out our biometric analyses, we verified that the expected phenotypic relationships existed in our sample. That is, when income was stratified into 3 groups consisting of those below 0.5 standard deviations below the mean, those within 0.5 standard deviations of the mean, and those above 0.5 standard deviations above the mean, the group means increased monotonically with decreasing income for both number of chronic illnesses and BMI. Analysis of variance showed that the mean differences were significant ( $F = 19.25$ ,  $df = 2$ ,

1911,  $p < 0.001$ , group means [and variances] for standardized age- and sex-corrected variables 0.28 [1.22] for low income,  $-0.05$  [0.95] for mid-range,  $-0.08$  [0.87] for high income for number of chronic illnesses;  $F = 13.67$ ,  $df = 2$ , 1693,  $p < 0.001$  for BMI, group means [and variances] for standardized age- and sex-corrected variables 0.19 [1.20] for low income, 0.02 [1.01] for mid-range,  $-0.23$  [0.65] for high income). This provided evidence for the decreasing phenotypic variance in physical health with increasing income.

MZ and DZ double-entered Pearson twin correlations are shown in Table IV. Such correlations on double-entered data closely approximate the intra-class correlation that measures the proportion of variance common to the members of the twin pairs. Comparison of such correlations can be used informally to assess the presence of genetic and environmental influences, but they only provide very approximate estimates of the extent of such estimates. Additive genetic influences reflecting the independent influence of multiple genes are indicated when the MZ correlation exceeds the DZ correlation, and shared environmental influences are indicated when the DZ correlation is more than half the MZ correlation. Nonadditive genetic influences reflecting dominance and other polygenic epistatic effects are indicated when the MZ correlation is more than twice the DZ correlation. As described above, however, such correlations provide information only about overall average levels of genetic and environmental influence in situations involving gene-environment interaction. We present them here because they do

**Table IV.** Double-entered Pearson twin correlations of the physical health, income, insurance, and education variables

Variable	MZ	DZ
Chronic illnesses	0.42	0.14
Top 50% on income	0.39	0.04
Bottom 50% on income	0.44	0.20
Body mass index	0.74	0.41
Top 50% on income	0.63	0.36
Bottom 50% on income	0.81	0.45
Income	0.38	0.13
Insurance	0.48	0.33
Education	0.68	0.54

*Note:* MZ is monozygotic; DZ is dizygotic. Standard errors for the full sample correlations ranged from 0.05 to 0.06. Except for Insurance, variables were age- and sex-corrected, and Income and Chronic illnesses were log transformed. Insurance correlations were tetrachoric.

provide this baseline information. All were significant and all showed substantial evidence of genetic contribution at an overall level. The table also shows the MZ and DZ correlations for the top and bottom halves of the sample stratified by income. These correlations reflect standardized covariance; thus they cannot clearly demonstrate the hypothesized relations because those relations are based on the presumption that neither variance nor covariance is constant across income levels. This was the reason for the implementation of the model described above.

Biometric models are not identified when both additive and nonadditive genetic influences and shared and non-shared environmental influences are modeled at once; one of the parameters must be dropped. As number of chronic illnesses showed no evidence of shared environmental influence and did show evidence of nonadditive genetic influences, we fit models including parameters reflecting additive and nonadditive genetic and nonshared environ-

mental influences to these data. As BMI showed no evidence of nonadditive genetic influences but did show some evidence of shared environmental influence, we fit models including parameters reflecting additive genetic and shared and nonshared environmental influences to these data.

**Biometric Modeling**

The model-fitting statistics for the gene-environment interaction models we applied are shown in Table V. For each measure of physical health problems and for each of income, income adjusted for the presence of insurance coverage, and income adjusted for both insurance coverage and years of education, we estimated two base models for use in developing the likelihood ratio test statistics to evaluate the significance of the interaction parameters. The first base model included estimates of the three possible interaction parameters: additive (and, for chronic illnesses, nonadditive) genetic, shared (for BMI) and

**Table V.** Model-fitting statistics for the gene-environment interaction models of chronic illnesses and body mass index moderated by income and corrected for other variables as indicated

Moderation variable	-2LL	df	LRT	LRT df	p	AIC
<i>Number of chronic illnesses</i>						
<i>Income</i>						
Freely estimated GxE	6640.2	4021				6674.2
Interaction constrained to 0	6654.7	4027	14.5	6	0.025	6676.7
Best-fitting model	6,643.0	4,026	2.8	5	0.731	6,667.0
<i>Income corrected for insurance</i>						
Freely estimated GxE	7,213.7	4,021				7,247.7
Interaction constrained to 0	7,227.1	4,027	13.4	6	0.037	7,249.1
Best-fitting model	7,217.7	4,026	4.0	5	0.549	7,241.7
<i>Income corrected for insurance and education</i>						
Freely estimated GxE	7,247.5	4,003				7,281.5
Interaction constrained to 0	7,257.0	4,009	9.5	6	0.147	7,279.0
Best-fitting model	7,249.2	4,008	1.7	5	0.889	7,273.2
<i>Body mass index</i>						
<i>Income</i>						
Freely estimated GxE	5,956.6	3,882				5,990.6
Interaction constrained to 0	6,002.4	3,888	45.8	6	< 0.001	6,024.4
Best-fitting model	5,964.4	3,886	7.8	4	0.099	5,990.4
<i>Income corrected for insurance</i>						
Freely estimated GxE	6,538.4	3,882				6,572.4
Interaction constrained to 0	6,587.7	3,888	49.3	6	< 0.001	6,609.7
Best-fitting model	6,546.3	3,886	7.94	0.095	6,572.3	
<i>Income corrected for insurance and education</i>						
Freely estimated GxE	6,579.8	3,864				6,613.8
Interaction constrained to 0	6,614.3	3,870	34.5	6	< 0.001	6,636.3
Best-fitting model	6,585.8	3,868	6.0	4	0.199	6,611.8

Note: -2LL is -2\*log likelihood of data; LRT is likelihood ratio test to freely estimated GxE model; df is degrees of freedom; AIC is Akaike Information Criterion. In all cases, the best-fitting model had included a moderating effect on common genetic influences. For BMI, there was also a moderating effect on unique genetic influences.



non-shared environmental influences, as well as main effect on the mean. The second base model constrained all of these interaction parameters to 0. We then estimated each combination of parameters and constraints. The statistics for the two base models and the best-fitting models resulting from this procedure are shown in the table for each measure of physical health problems.

For numbers of chronic illnesses, a model with both additive genetic and nonshared environmental interaction terms constrained to 0 fit best. This is the model labeled “Best fitting” in the table. The non-additive genetic interaction term could not be so constrained without significant deterioration in fit, as shown in the table by comparing both the change in  $\chi^2$  between the best-fitting model and the model with the interaction constrained to 0 and the AIC’s. This remained the case when income was controlled for the presence of health insurance coverage alone, but when income was controlled for both the presence of health insurance coverage and level of education the change in  $\chi^2$  was no longer significant, though the AIC still indicated that this model was best-fitting

and the confidence interval indicated that the interaction term was significant. For BMI, the model with both shared and nonshared environmental interaction terms constrained to 0 fit best, and this is the model labeled “Best fitting” in the table. Again, the genetic interaction and main effect terms could not be so constrained, and this time the full effect remained after controlling income for the presence of health insurance coverage and level of education. The parameters estimated in the best-fitting models, along with their 95% confidence intervals, are shown in Table VI.

Figures 2 and 3 present graphical descriptions of the relationships implied by these parameters among the sources of variance associated with the physical health variables and income after adjustment for both insurance coverage and level of education. Note that the graphs show that, because genetic variance decreases with increasing income and the other sources of variance remain constant, phenotypic variance decreases with increasing income as well. The graphs also show that, for people with income 1 standard deviation above the mean, for example, genetic variance associated with numbers of chronic illnesses

**Table VI.** Parameter estimates from best-fitting gene-environment interaction models of chronic illnesses and Body Mass Index moderated by income corrected for health insurance and education

	Chronic illnesses		Body mass index	
	Parameter estimate	95% Confidence interval	Parameter estimate	95% Confidence interval
aI	0.28	(-0.34, 0.52)	0.50	(0.39, 0.58)
cI or dI	0.45	(0.10, .60)	0.00	(-0.24, 0.24)
eI	0.76	(0.71, .82)	0.78	(0.73, 0.84)
aC	-0.45	(-0.65, 0.00)	-0.25	(-0.42, -0.08)
cC or dC	0.12	(-0.27, 0.54)	0.00	(-0.42, 0.42)
eC	-0.08	(-0.16, 0.00)	0.00	(-0.06, 0.06)
aH	0.00	(-0.56, 0.56)	0.82	(0.69, 0.88)
cH or dH	-0.46	(-0.70, -0.06)	0.00	(-0.41, 0.41)
eH	0.75	(.70, .80)	0.46	(.43, .50)
$\beta$ aC	0.00	fixed	-0.07	(-0.11, -0.02)
$\beta$ cC or $\beta$ dC	0.00	fixed	0.00	fixed
$\beta$ eC	0.00	fixed	0.00	fixed
$\beta$ aH	0.00	fixed	-0.12	(-0.18, -0.06)
$\beta$ cH or $\beta$ dH	0.11	(0.03-0.27)	0.00	fixed
$\beta$ eH	0.00	fixed	0.00	fixed
$r_a$ at -2sdI	-0.16	(-0.47, -0.05)	-0.11	(-0.35, 0.11)
$r_a$ at 0sdI	-0.33	(-0.64, -0.22)	-0.30	(-0.52, -0.06)
$r_a$ at 2sdI	-0.69	(-1.00, -0.58)	-0.55	(-0.79, -0.32)
$r_c$	N/A	N/A	0.00	N/A
$r_e$	0.14	(-0.28, 0.01)	0.00	(-0.13, 0.13)

Note:  $r_a$  at -2sdI refers to genetic correlation between income (corrected for insurance and education) and the appropriate physical health measure at 2 standard deviations below the mean for corrected income. The entries following can be interpreted analogously. As discussed in the text, we modeled nonadditive genetic influences (D) for chronic illnesses and shared environmental influences (C) for BMI. For chronic illnesses, the genetic correlation includes both additive and nonadditive genetic influences. enetic.

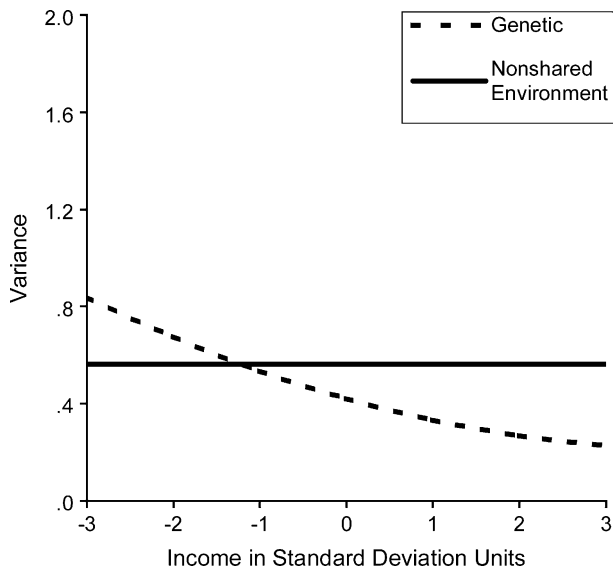


Fig. 2. Variance in log of number of chronic illnesses as a function of log of income corrected for insurance and education, by source of variance.

(adjusted for health insurance coverage and education) was about 0.33, while for those with income 1 standard deviation below the mean, genetic variance associated with numbers of chronic illnesses (adjusted for health insurance coverage and education) was about 1.7 times greater, or 0.55.

Table VI also shows the genetic and environmental correlations between income and chronic illnesses, and income and BMI. Because it was unique genetic variance in physical health that decreased

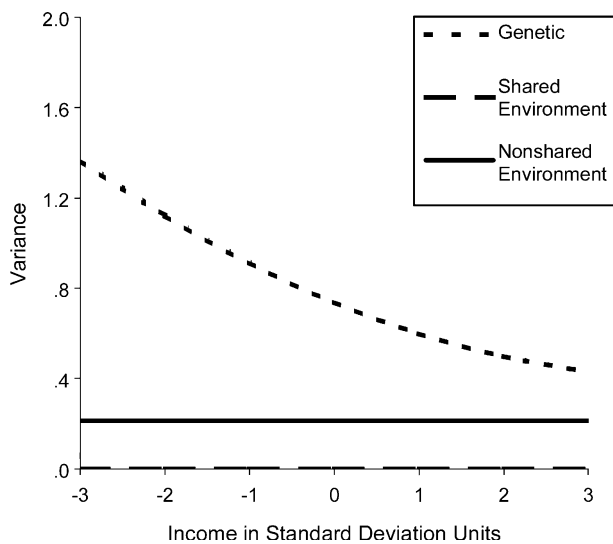


Fig. 3. Variance in body mass index as a function of log of income corrected for insurance and education, by source of variance.

with increasing income (see Equation (3) above), the genetic correlations increased with income. The patterns of correlations for the two measures of physical health were very similar. For both, the genetic correlations with income were small at 2 standard deviations below the mean (and insignificant for BMI), but at 2 standard deviations above the mean the correlations were both significant and substantial ( $-0.69$  for numbers of chronic illnesses and  $-0.55$  for BMI). The environmental correlations were not significant.

### DISCUSSION

In this study, we investigated the extent to which a gene-environment interaction underlies the relationship between income and physical health problems. We found evidence for such an interaction in two measures of physical health problems, numbers of chronic illnesses and BMI, in a nationwide US twin sample ranging in age from 25 to 74. Our results suggested that genetic variance associated with physical health problems decreases with increasing income. As others have observed about the gradient at the phenotypic level, (Adler and Snibbe, 2003), the gene-environment interaction effect was not the result of simple access to insurance coverage, nor was it the result of the direct effects of level of education.

This study had several limitations. First, all of our measures were based on self-report, and are thus subject to the well-known limitations of such measures. People may have only limited understanding of the health treatments they receive and they may have health conditions of which they are unaware and for which they have not been treated. They may be motivated to present themselves either favorably or unfavorably for any of a variety of reasons, a factor we might expect to operate especially strongly for the height and weight measurements used to calculate BMI. At the same time, self-reports of measures such as those we used are economical and straightforward to collect, and generally found to be substantively correlated with physicians' assessments (Idler *et al.*, 1999; Jylha, 1994; Krause and Jay, 1994). Follow-up studies could help to clarify the extent of this limitation by, for example, both collecting self-ratings of health and carrying out home visits to assess physical health problems more directly, though this would, of course, involve considerable additional expense. Another approach would be to collect self-ratings of health in countries such as Canada that have

population-based medical records associated with their universal health insurance coverage that can be accessed without self-report.

Second, neither of our physical health problem measures completely captured overall physical health problems. The chronic illness measure did not include cancer and heart disease. It also did nothing to assess the degree to which the conditions that were included, some of which were relatively minor, debilitated the participant or could be expected to threaten normal lifespan. High BMI, though a risk factor for many serious physical health problems, is not necessarily an indication of current poor health status, and low BMI can sometimes be an indication of poor health. Fourth, our sample had somewhat higher income and was somewhat better educated than the general population and lacked variability particularly at the lower levels of health and income. Most of these limitations, however, should have reduced our ability to detect interaction effects. The fact that we were able to do so probably attests to the strength of the associations.

Third, it is important that this study be replicated, especially in larger samples. This sample was big enough to detect an effect predicted by theory and observed in other organisms, but more complex effects may also exist. For example, though we found no indication of significant sex differences, there is the possibility that there were sex differences in the interaction effects that, in our modestly sized sample, would be relatively difficult to detect.

The relationship between income and physical health has long presented a tantalizing puzzle to researchers and policymakers alike. Researchers have begun to articulate a process centered around increasing demands and decreasing resources for dealing with those demands associated with lower levels of income (Gallo and Mathews, 2003) to explain the relationship, but the underlying physiological mechanisms by which this process might associate income and physical health have remained unclear. The results we present here, though they require replication in other samples, potentially shed light on some of those mechanisms. First, they provide some evidence for the existence of common genetic influences on income and physical health, a possibility that has been largely dismissed in the past (Adler *et al.*, 1994). Second, for both numbers of chronic illnesses and BMI, total variance and genetic variance were smaller and mean levels were lower among those with higher income. The stress theory of aging (Parsons, 2003) holds that genes for metabolic

efficiency enabling adaptation to stressful environments play a primary role in affecting lifespan and, by implication, health. Genetic interactions similar to those we observed in this study for human physical health problems have been documented in *Drosophila melanogaster* for longevity in situations of varying degrees of biological stress (Parsons, 2002). The stress theory of aging suggests that, for humans in the environment of evolutionary adaptation, increased longevity resulted from primary selection for resilience under stress, and an individual's relative fitness (i.e. health) would have been similar throughout the lifespan. In more benign conditions such as those for laboratory animals or humans in the developed world today, a greater proportion of the population can survive to older ages, but this trend is counteracted by reduction in the degree of selection for resilience under stress. The results of this study seem highly consistent with this theory.

The observation of decreasing genetic variance in physical health problems with increasing income, even after adjustment for presence of insurance coverage and education, helps to explain how the genetic influence on physical health problems is involved in its relationship with income. It does not, however, explain how increased income may act to compress the genetic variance. What is it about income that results in changes in genetic variance? Whatever it is, it is important to note that it appears to transcend specific disease categories. We examined two candidates in this study: health insurance coverage and education. The fact that they failed to explain the decreasing variance makes it especially likely that psychological variables may be involved. We suggest that this may be a fruitful area for future research.

## ACKNOWLEDGMENTS

This research was supported by the John D. and Catherine T. MacArthur Foundation Research Network on Successful Midlife Development, and by National Institute on Aging Grant No. AG20166. We thank Tom Bouchard for his helpful comments on earlier drafts of the manuscript.

## REFERENCES

- Adler, N. E., Boyce, T., Chesney, M. A., Cohen, S., Folkman, S., Kahn, R. L., and Syme, S. L. (1994). Socioeconomic status and health. *Am. Psychol.* **49**:15–24.
- Adler, N. E., Epel, E. S., Castellazzo, G., and Ickovics, J. R. (2000). Relationship of subjective and objective social status

- with psychological and physiological functioning: Preliminary data in healthy white women. *Health Psychol.* **19**:586–592.
- Adler, N. E., and Snibbe, A. C. (2003). The role of psychosocial processes in explaining the gradient between socioeconomic status and health. *Curr. Directions Psychol. Sci.* **12**(4), 119–123.
- Akaike, H. (1983). Information measures and model selection. *Bull. Int. Stat. Inst.* **50**:277–290.
- Antonovsky, A. (1967). Social class, life expectancy and overall mortality. *Millbank Memorial Fund Quart.* **XLV**:31–73.
- Bailis, D. S., Segall, A., Mahon, M. J., Chipperfield, J. G., and Dunn, E. M. (2001). Perceived control in relation to socio-economic and behavioral resources for health. *Social Sci. Med.* **52**:1661–1676.
- Christensen, L., Frederiksen, H., Schousboe, K., Skytthe, A., Wurmbs-Schwark, N., von, Christensen, K., and Kyvik, K. (2003). Age- and sex-differences in the validity of questionnaire-based zygosity in twins. *Twin Res.* **6**:275–278.
- Fiscella, K., and Franks, P. (1997). Does psychological distress contribute to racial and socioeconomic disparities in mortality. *Social Sci. Med.* **45**:1805–1809.
- Gallo, L. C., and Matthews, K. A. (2003). Understanding the association between socioeconomic status and physical health: Do negative emotions play a role? *Psychol. Bull.* **129**:10–51.
- Gump, B. B., Matthews, K. A., and Raikkonen, K. (1999). Modeling relationships among socioeconomic status, hostility, cardiovascular reactivity, and left ventricular mass in African American and white children. *Health Psychol.* **18**:140–150.
- Harris, J. R., Pedersen, N. L., Stacey, C., McClearn, G. E., and Nesselroade, J. R. (1992). Age differences in the etiology of the relationship between life satisfaction and self-rated health. *J. Aging Health* **4**:349–368.
- Idler, E. L., Hudson, S. V., and Leventhal, H. (1999). The meanings of self-ratings of health: A quantitative and qualitative approach. *Res. Aging* **21**:458–476.
- Jylha, M. (1994). Self-rated health revisited: Exploring survey interview episodes with elderly respondents. *Social Sci. Med.* **39**:983–990.
- Karlamangla, A. S., Singer, B. H., McEwen, B. S., Rowe, J. W., and Seeman, T. E. (2002). Allostatic load as a predictor of functional decline: MacArthur studies of successful aging. *J. Clin. Epidemiol.* **55**:696–710.
- Kendler, K. S., Thornton, L. M., Gilman, S. E., and Kessler, R. C. (2000). Sexual orientation in a US national sample of twin and sibling pairs. *Am. J. Psychiat.* **157**(11), 1843–1846.
- Kessler, R. C., Gilman, S. E., Thornton, L. M., and Kendler, K. S. (2004). Health, wellbeing, and social responsibility in the MIDUS twin and sibling subsamples. In O. G. Brim, C. D. Ryff and R. C. Kessler (eds.), *How Healthy Are We? A National Study of Wellbeing at Midlife*. Chicago: University of Chicago Press.
- Komaroff A. L. (Ed.) (1999). *Harvard Medical School Family Medical Guide*. New York: Simon & Schuster.
- Krause, N. M., and Jay, G. M. (1994). What do global self-rated health items measure? *Med. Care* **32**:930–942.
- Lachman, M. E., and Weaver, S. L. (1998). The sense of control as a moderator of social class differences in health and well-being. *J. Pers. Social Psychol.* **74**:763–773.
- Lykken, D. T., Bouchard, T. J., McGue, M., and Tellegen, A. (1990). The Minnesota Twin Family Registry: Some initial findings. *Acta Genetica Medica Gemellol* **39**:35–70.
- Marmot, M. G., Smith, G. D., Stansfield, S., Patel, C., Head, J., White, I., Brunner, E., and Feeney, A. (1991). Health inequalities among British civil servants: The Whitehall II study. *Lancet* **337**:1387–1393.
- McEwen, B. S. (1998). Protective and damaging effects of stress mediators. *N Engl. J. Med.* **338**:171–179.
- McGue, M., and Bouchard, T. J. Jr. (1984). Adjustment of twin data for the effects of age and sex. *Behav. Genet.* **14**:325–343.
- Neale, M. C., Boker, S., Xie, G., and Maes, H. H. (1999). *Mx: Statistical Modeling*. Richmond, VA: Medical College of Virginia Department of Psychiatry.
- Parsons, P. A. (2002). Aging: The fitness-stress continuum and genetic variability. *Exp. Aging Res.* **28**:347–359.
- Parsons, P. A. (2003). From the stress theory of aging to energetic and evolutionary expectations for longevity. *Biogerontology* **4**(2), 63–73.
- Pauzova, Z., Sedova, L., Berube, J., Hamet, P., Tremblay, J., Dumont, M., Gaudet, D., Pravenec, M., Kren, V., and Kunes, J. (2003). Segment of rat chromosome 20 regulates diet-induced augmentations in adiposity, glucose intolerance, and blood pressure. *Hypertension* **41**:1047–1055.
- Purcell, S. (2002). Variance component models for gene-environment interaction in twin analysis. *Twin Res.* **5**(6), 554–571.
- Roberts, G. (1999). Age effects and health appraisal: A meta-analysis. *J. Gerontol.* **54B**(1), S24–S30.
- Singer, B. H., and Ryff, C. D. (1999). Hierarchies of life histories and associated risks. In N. E. Adler, M. Marmot, B. S. McEwen and J. Stewart (eds.), *Socioeconomic Status and Health in Industrial Nations*. New York: New York Academy of Sciences, pp. 96–115.
- Stratton, D. A. (1998). Reaction norm functions and QTL-environment interactions for flowering time in *Arapidopsis thaliana*. *Heredity* **81**:144–155.
- Strawbridge, W., and Wallhagen, M. I. (1999). Self-rated health and mortality over three decades. *Res. Aging* **21**:402–416.
- Tiret, L. (2002). Gene-environment interaction: a central concept in multifactorial diseases. *Proc. Nutr. Soc.* **61**:457–463.
- Weiss, E. P., Brown, M. D., Shuldiner, A. R., and Hagberg, J. M. (2002). Fatty acid binding protein-2 gene variants and insulin resistance: gene and gene-environment interaction effects. *Physiol. Genomics* **10**(3), 145–157.

Copyright of Behavior Genetics is the property of Springer Science & Business Media B.V.. The copyright in an individual article may be maintained by the author in certain cases. Content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.