GENETIC INTERACTIONS WITH PRENATAL SOCIAL ENVIRONMENT:
EFFECTS ON ACADEMIC AND BEHAVIORAL OUTCOMES

Dalton Conley
Emily Rauscher

Genetic Interactions with Prenatal Social Environment: Effects on Academic and Behavioral
Outcomes
Dalton Conley and Emily Rauscher
NBER Working Paper No. 16026
May 2010
JEL No. I1

ABSTRACT

Caspi et al. (2002, 2003), Guo et al. (2008a), and Pescosolido et al. (2008) all claim to have demonstrated
allele-by-environment interactions, but in all cases environmental influences are potentially endogenous
to the unmeasured genetic characteristics of the subjects and their families. Thus, gene-gene interactions
cannot be ruled out as an alternative explanation. Second, these studies have not deployed adjustments
for multiple hypothesis testing—always an issue, but particularly so for GE studies with multiple alleles
and outcomes. Using data from the National Longitudinal Survey of Adolescent Health (Add Health),
we address these limitations of previous studies by taking advantage of a natural experiment that randomizes
a particular environmental influence – fetal position, resulting in birth weight discordance within monozygotic
twin pairs (validated with dizygotic twins as well). Whether or not we use corrections for multiple
statistical tests, we find no support for the GE interactions (or for main effects of genes or birth weight)
found in past research and, in fact, the only significant allele-birth weight interaction we reveal works
in the opposite direction of Caspi et al.'s classic finding on 5-HTT and maltreatment.

Dalton Conley
New York University
6 Washington Square North #20
New York, NY 10003
and NBER
conley@nyu.edu

Emily Rauscher
295 Lafayette Street 4th Floor
NY NY 100012
ekr220@nyu.edu

Background

Studying genetic-environmental (GE) interactions has long been a goal of social scientists fond of touting the dependence of genetic expression on social structure. However, how do we get from the adage that "a gene for aggression lands you in prison if you're from the ghetto, but in the boardroom if you're to the manor born" to a serious empirical research agenda on the study of GE interactions? Recently, genetic markers on specific loci—such as single nucleotide polymorphisms (SNPs)—have seemed to offer hope for those interested in an explicit research program aimed at specifying and measuring gene-environment interactions for complex traits (what geneticists call Quantitative Traits). Polymorphisms are genetic variants that occur within a species. These could include base-pair substitutions—among one of the four nucleotides that make up our genetic code (G, guanine; C, cytosine; A, adenosine; and T, thymine)—which may: 1) affect the amino acid produced out of that codon (a triplet of nucleotides that determine which amino acid should come next when the messenger-RNA is translated into a protein) if the polymorphism is in an open reading frame (ORF) of a gene (i.e. the protein-related coding region) and is non-synonymous; 2) truncate the protein by causing the transcription machinery to stop there (by producing a stop codon); or 3) do nothing (what are called silent or synonymous mutations) since multiple three-letter codes may result in the same amino acid being produced (though, perhaps at different efficiency levels, something called codon-bias). Hence, these non-lethal polymorphisms, which result from mutations, may present an opportunity to study how specific environments—social or biophysical—may result in different outcomes depending on an individual's genotype.

The basic logic is the following: A certain proportion of a population sample is found to have a variant of a particular allele. If this allele is shown to be randomly distributed across

demographic subgroups (or, for example, within a particular subgroup such as an ethnic group), and, likewise, it is found to be associated with a specific social outcome or tendency (such as addictiveness, shyness, schizophrenia, to name a few) *within* that same population (or subgroup as the case may be), then researchers may try to look for specific environmental conditions which seem to magnify or mitigate its effect—such as family structures, parents' behavior, or simply socioeconomic status.

For example, Guo and Stearns (2002) claim that genetic influence on intellectual development depends on parental employment and ethnicity. Shanahan and colleagues (2008) find that the "risky" DRD2 genotype reduces the likelihood of post-secondary school attendance for boys and that this risk is moderated by social capital. Yet they also find that boys with the risky DRD2 genotype are less likely to have high social capital environments. This is a fundamental flaw; it could be that the allele(s) are interacting not with differential social environments, but rather with other, non-randomly distributed genes (even if the principal gene in question is indeed randomly distributed).[1] To provide other examples of this fallacy, Guo et al. (2008a) and Pescosolido et al. (2008) both claim to have revealed conditional genetic effects depending on family support or friendship networks. Guo et al.

---

[1] Recent genome-wide association studies (e.g., Beauchamp et al. 2009) are particularly subject to this criticism of genetic population stratification. Beauchamp et al. use a method developed by Price et al. (2006) to attempt to control for subpopulation differences. Using principle components analysis, they identify the 10 strongest axes of variation in the sample. They then control for these underlying subpopulations in GWAS models with the hope of preventing spurious results. However, while principle components analysis accounts for as much variability as possible, it is impossible to know whether it successfully identifies all important subpopulations. Beauchamp et al. drop outliers and there may be more than 10 important subpopulations, just to give one example of the potential problems. Within-family studies avoid concerns about population stratification. If allele variation is studied within families (i.e. across siblings) then it does indeed offer a potential way to measure specific genetic influences with some certainty. One would then compare the expression of that allele—as compared to the sibling without the polymorphism, for example—in families of various demographic or economic backgrounds.

indicate that regular family meals eliminate the delinquent tendencies associated with the "risky" DRD2 genotype. Similarly, family support reduces the genetically-influenced risk of alcohol dependence (Pescosolido et al. 2008). However, neither use within-sibling set differences to control unmeasured environmental differences that may covary with genes; and neither attempts to address the unmeasured genetic variability across subjects. Indeed, the environmental mediators in both cases include family behaviors, which could easily reflect genetic rather than exogenous environmental differences. That is, family support and closeness could be related to the very genes in question. Even if they are not, other genetic differences could explain the apparent gene-environment interaction.

A paper by Caspi et al. (2002) that has become a classic in this area of research claims to have uncovered a GE-interaction by comparing male children who have a particular functional polymorphism in the MAOA gene (monoamine oxidase A)—an enzyme which breaks down various neurotransmitters once they are chaperoned out of the synaptic cleft—with those who do not among a longitudinal sample of 1,037 white New Zealanders followed from ages 3 to 26. Those individuals who showed a variable number tandem repeat (VNTR) in the promoter region of the gene (the area that precedes the actual coding portion but which is important to transcriptional activation and regulation) putatively transcribe (and by extension translate) MAOA at a lower rate than those without this polymorphism on their X-chromosome. In turn, MAOA activity as indicated by this genetic difference was interacted with degree of maltreatment the respondents experienced between the ages of 3 and 11 to predict an index of anti-social behavior that included four measures ranging from criminal convictions to anti-social personality disorder criteria of the DSM-IV. They argue that while there do exist other MAO genes that may compensate for deficiencies in MAOA (in particular

MAOB), among children these are not yet fully expressed, thus making MAOA particularly important with respect to moderating the effect of maltreatment during early childhood.

Eight percent of the sample experienced severe maltreatment, 28 percent experienced "probable" maltreatment and 64 percent experienced no maltreatment. In a multiple regression context, the main effect of maltreatment level on the antisocial behavior index was significant, whereas the main effect of MAOA activity level was not, but an interaction effect between the two measures was statistically significant at the $\alpha=.01$ level. They argue that this is a true genetic-environmental interaction effect since the MAOA genotypes were not significantly differently distributed across maltreatment levels—suggesting that this genotype did not itself influence exposure to maltreatment (i.e. the environment is not standing in for the genotype).

In a follow-up study (2003), they use the same cohort to examine the interaction of stressful life events with alleles of the serotonin transporter gene (5-HTT) linked promoter region (5-HTTLPR). Specifically, individuals who have a short 5' (i.e. upstream) promoter may show more propensity than those with a long promoter toward depression. However, previous studies had come to conflicting results; namely many replications had failed to produce results claimed in earlier linkage studies. Some researchers despaired that psychiatric and other behavioral phenotypes were controlled by so many quantitative trait genes that modeling genetic effects in a robust, direct way would not be possible and/or would account for little of the variation (see, e.g., Hamer 2002). Caspi et al. (2003) instead argue that, rather than complicated gene-gene interactions, the muddle of results could result from GE interactions. This muddle motivated their search for an interaction effect of stressful life events and 5-HTTLPR allele.

This is an autosomal gene—meaning that individuals of both sexes have two copies—so they compared individuals with the homozygous long genotypes and heterozygotes (long / short) to those who were homozygous for the short alleles. They found that in the subsample who had experienced no stressful life events between ages 21 and 26, there was no difference between the three genotypes in the propensity to depression. However, as the number of self-reported stressful life events increased, the genotypes diverged with respect to their likelihood of clinical depression at age 26. They interpret this as a GE interaction.[2]

However, as was the case with Guo (2002 & 2008a) or Pescosolido (2008), it could still be possible that Caspi et al. (2002; 2003) were actually uncovering gene-gene interactions in both studies, because they did not have an exogenous source of environmental variation. In the latter case, those with the "at risk," short alleles were, in fact, more likely to report stressful events than those who had long alleles. We may conclude, then, that measured genotype did influence the measured environmental measure. The researchers try to get around this by reversing the time order: Measuring stressful life events between ages 21 and 26 and measuring depression at age 21 (i.e. prior). When they do this, they do not find the significant interaction they did in the "correctly" ordered model. However, it still may be the case that depression was induced by a gene-gene interaction since an underlying unmeasured gene could cause the phenotype of "negative life events" to emerge in one's early 20s: Imagine a gene that

---

[2] Animal-based studies also find evidence of environmental effects conditional on genotype at these locations. For example, Murphy et al. (2001) studied mice with a disrupted 5-HTT gene and found that those with risky alleles were more fearful and had higher stress hormone levels in response to stress, but no differences by genotype without environmental stress. Research on rhesus macaques found different biological reactions depending on 5-HTT genotype for those raised in stressful environments, but no differences among those raised normally (Bennett et al. 2002). Thus, animal-based experiments suggest gene-environment interactions may explain substantial variance in health and behavioral outcomes, but similar human experiments are not ethical.

causes excessive thrill-seeking and risk-taking, which, in turn, manifests as negative events during one's early adulthood. As for the MAOA interaction, we face the same issue: While measured maltreatment did not vary by MAOA status, it could very well have varied by other genes (present in the parents and potentially passed on to the children). Thus, it would not be the maltreatment that interacted with MAOA status but rather the underlying, unmeasured genotype, which, in combination with given MAOA alleles, causes both parents and offspring to act anti-socially.

In fact, supporting the notional importance of gene-gene interactions (and offering a competing model to GE interactions) is recent genetics research that has shown that among the genes studied in humans (or other [model] organisms such as the fruit fly, *drosophila melanogaster*, or the nematode worm, *Caenorhabditis elegans*), the vast majority of known genes are linked in a single network component when measured by either protein-protein interactions, regulatory relationships, or phenotypic co-variation (Jeong et al. 2001; Stelzl et al. 2005). This suggests that, indeed, one cannot conceptualize the perturbation of one gene as unrelated to the impact of other genes. Conversely, the embeddedness of this network suggests that genomic systems are highly redundant and robust and that other genes may be up- (or down-) regulated to compensate when a given gene is non-functional (or hypertrophic).[3]

Gene-environment research is not alone in its failure to account for unmeasured genetic differences. Low birth weight is consistently found to have developmental consequences and has

---

[3] The real rub is that, since we can plausibly postulate second-, third-, fourth- and, ultimately, Nth-order interactions across alleles, there simply would not be enough degrees of freedom in the approximately seven billion human beings currently occupying the planet to properly test a fully specified model ($21,000! = 9.58 \text{ E } 81648 > 7,000,000,000$). The discovery of about 21,000 genes—a figure much lower than originally hypothesized—is good news in that it is a tractable number of alleles for geneticists to study. However, the irony lies in the fact that, if this lowly number of genes explains the development of human beings in all their forms, then gene-gene interactions are probably quite important

been associated with a wide variety of outcomes in later life, including cognitive ability, behavioral problems, and health (Lowe et al. 2009; Hayes and Sharif 2009; Schlotz and Phillips 2009; Cooper et al. 2009). However, much of the apparent effect of birth weight may reflect inherited differences. While some research has estimated the effect of birth weight using identical twins (Behrman and Rosenzweig 2004; Conley et al. 2003), most studies have relied on sibling (Conley and Bennett 2000; 2001) or even between-family comparisons (Hack et al. 2002; Rich-Edwards 1997; Sorensen et al. 1997), which cannot rule out genetic explanations for differences in outcomes; such an approach may create spurious effects of low birth weight if the same genes driving birth weight are also driving educational or developmental outcomes. For instance, among singleton siblings or fraternal twins, genetic advantage could influence fetal nutrition, making any apparent effects of birth weight spurious. Further, birth weight could affect only those with certain alleles – putting them at risk for poor health outcomes for example. Therefore, even studies using identical twins may be lumping vastly heterogenous treatment effects of birth weight into one average treatment effect.

To address these concerns in both the GE interaction and birth weight literatures, in the present study, we deploy a novel approach: We use both MZ and DZ twin differences in birth weight to predict educational, mental health and behavioral outcomes using nationally representative data from the National Longitudinal Survey of Adolescent Health (Add Health). We deploy MZ twins in order to hold constant genetic differences that may influence both birth weight and the outcome of interest. We then see if the treatment effect of low birth weight varies across twin sets that are divergent on measured alleles. Since this strategy leaves open the possibility that loci across which we are stratifying are correlated with other unmeasured environmental or genetic differences due to population stratification (or linkage

disequilibrium), we also present estimates that compare the effects of that allele (and its interaction with birth weight, as well as a main effect for birth weight), on those same outcomes using intra-sibling comparisons among dizygotic twins. In this way, with the exception of other genes that may be linked to the gene in question through linkage disequilibrium during recombination, all other genes are orthogonal to the measured genetic difference. (However, in the DZ models, we cannot say that birth weight differences are not endogenous to unmeasured genetic characteristics that vary between the twins.) Thus, we present findings that are robust to both these estimation strategies *and* which pass a stricter statistical threshold that accounts for multiple hypothesis testing (another failing of most earlier GE work).

To preview our results: We purport to establish—for the first time in humans outside the laboratory—that environmental stress can interact with alleles that otherwise appear to have no direct effect and cause behavioral phenotypes to emerge. Environmental conditions, even as early as in the womb, can moderate the effect of alleles on social outcomes but in ways that contradict the previous literature. Namely, neither birth weight nor the alleles we measure appear to have any direct effects on the outcomes we study. However, birth weight interacts in both DZ and MZ twin models such that decreased birth weight (previously considered a risk factor) results in lower risk of depression—but only for those who have the "risky" serotonin transporter promoter region allele. These results—standing in contrast to a whole host of previous findings—should fuel debate and future research in social genetics and provide a methodological framework for moving that debate forward.

Methodological Approach

Birth weight is orthogonal to genes for identical twins; that is, genes are shared completely so they cannot affect nutritional advantage relative to one's identical twin in the womb. MZ twin pair fixed effects exploit random variation in birth weight to identify a gene-environment interaction (with the notable exception of unique, de novo mutations post zygotic cleavage). Low birth weight and particular alleles have previously been associated with behavioral and health outcomes – including depression, delinquency, and educational outcomes. Previous research predicts a greater return to birth weight for those with a genetic risk for depression, delinquency, and poor educational performance. Out of concern for population stratification, we also show results for dizygotic twins. That is, risky genotypes may not be randomly distributed by demographic subgroups, in which case we could spuriously attribute an environment-environment interaction in MZ twin models to a GE interaction if G is acting as a proxy for E. Fixed effects model the differences in outcomes within twin pairs.

$$Y_{ij} = a + bBirthweight_{ij} + cRiskyAllele_{ij} + dBirthweight_{ij} * RiskyAllele_{ij} + eSex_{ij} + fTWINSET_j + \varepsilon_{ij}$$

where Y_{ij} is the outcome for a given twin i in pair j, b is the effect of birth weight within twin pair j that has the so-called normal allele, c is the main effect of differences in the risky gene, and $b + d$ is the effect of birth weight (within twin pairs) for those with the risky allele, e represents differences due to sex (for dizygotic twin pairs only), and ε_{ij} is the sibling-specific (i.e. idiosyncratic) error, assumed to be unrelated to genes, birth weight, and control factors. Monozygotic twins share the same genes and sex, so c and e drop out. Individuals are compared to their twin who is the same age, obviating the issue of age differences.

The National Longitudinal Study of Adolescent Health (Harris 2009) provides birth weight and sequenced genotype data for three genes putatively related to behavioral and health outcomes conditional on environment (5HTT, DRD2, and MAOA). (Actually, they sequence loci at seven genes, but we focus on the three that have been studied in previous GE interaction research mentioned earlier.) These genes are involved in neurotransmitter (e.g., dopamine and serotonin) transport, receiving and recycling, vital for cognition and behavior. Previous research, including animal-based genetic manipulation, has identified these three loci as influential on a variety of outcomes. For example, Cases et al. (1995) and Shih and Thompson (1999) studied knockout mice (those with the MAOA gene removed) and found they had increased dopamine, serotonin, and norepinephrine levels and increased aggression among males.

Thus far, Add Health has released three waves of panel data for sibling pairs. In 2001-2 (wave 3, the most recent wave that includes genetic data), respondents were ages 18-26 (mode = 22). Siblings of individuals identified as twins in the stratified sample were added, yielding 64% of sibling pairs from the probability sample and 36% from convenience sampling. Buccal swabs were collected in wave 3 from 2612 of the 3139 eligible siblings from wave 1 (a compliance rate of 83%; Harris et al. 2006) for DNA sequencing at the Institute for Behavioral Genetics. Monozygosity was genetically confirmed, requiring complete matches on 11 "highly polymorphic, unlinked short tandem repeat (STR) markers: D1S1679, D2S1384, D3S1766, D4S1627, D6S1277, D7S1808, D8S1119, D9S301, D13S796, D15S652 and D20S481 and the sex determining locus amelogenin" (Harris et al. 2006:992). Our sample includes over 200 twin pairs not missing birth weight, genetic data, or outcome data for either twin (sample size for each model depends on the number of pairs with complete outcome data).

Research has linked polymorphisms in the human genes DRD2, 5HTT, and MAOA with behavior and health outcomes. At the D2 dopamine receptor gene locus (DRD2), the A1 allele is related to fewer receptor binding sites (Pohjalainen et al. 1998). Compared to the A2 allele, possessing the A1 allele has been associated with anxiety, depression, novelty seeking, impulsiveness, lack of inhibition, and substance use (Lawford et al. 2006; Noble et al. 1998; Wiers et al. 1994; Blum et al. 1991). Consistent with previous research, individuals possessing an A1 allele are considered to have the risky genotype.[4] For all three genes, those with no copies of the risky allele are specified in models below, but results are similar with alternative specifications.

Variation at the serotonin transporter gene locus (5-HTT) has been associated with propensity toward depression. Previous research has suggested that individuals with a short allele in the promoter region of 5-HTT have stronger depressive reactions to stressful life experiences (Caspi et al. 2003).

The MAOA gene codes for monoamine oxidase A, which chaperones and breaks down neurotransmitters (e.g., serotonin, dopamine, norepinephrine) and for which variation has been linked to disposition toward aggression in both animals and humans (Guo et al. 2008b; Rowe 2001; Cases et al. 1995; Shih and Thompson 1999; Brunner et al. 1993). Caspi et al. (2002) and Guo et al. (2008a) found that those with fewer repeats (i.e., lower MAOA activity) in the

---

[4] Guo et al. (2008a) suggest a significant interaction between DRD2 and social factors (e.g., family meals and growing up with two biological parents) on delinquent behavior. However, they present an interaction for heterozygotes – with exactly one A1 allele – not for either homozygous type, which makes it difficult to interpret their results. It is unclear whether heterozygosity or having a short allele is driving the results. We distinguish those with no copies of the A1 allele. Results for main and interaction effects of DRD2 variation on delinquency are presented below. Guo and colleagues also found an association between the longer 10R allele of DAT, a dopamine transporter gene, and delinquent behavior, but no environmental interaction. DAT is not presented here because a GE interaction was not found (though in separate analysis we do not find a significant interaction at this locus either).

promoter region of the gene are more susceptible to family environmental influences. Individuals with two copies of the long 4R allele are specified below.

Birth weight is reported by parents, measured in ounces, and logged. Depression is measured using the Center for Epidemiologic Studies-Depression Scale (CES-D). It consists of 20 questions included in the Add Health survey which ask respondents to rate the frequency of a depressive symptom from 0 (never/rarely) to 3 (most/all of the time). The sum of responses for all 20 items indicates the frequency of depressive symptoms. Results are also investigated using an indicator for "any symptoms" and logged scores (after adding one to avoid excluding those with no symptoms). Following Fletcher and Lehrer (2009) and Roberts et al. (1991), age and gender-specific threshold measures of depression are also investigated. Results are largely the same in these specifications and are not presented, but are briefly discussed below.

Delinquency is measured using 12 questions from the Add Health wave 3 survey that ask about deviant behavior in the past 12 months. For example, questions include how often you used someone else's credit/bank card without their permission, deliberately wrote a bad check, stole something, or used a weapon. Sensitivity analyses use an indicator for the presence of any delinquent behaviors and logged scores (after adding one)—results are similar to a linear specification. Educational achievement is measured using cumulative high school GPA gathered from high school transcripts. An indicator of college attendance (as well as a continuous measure for highest grade completed) is also tested, for comparability with Shanahan et al. (2008). In general, results do not differ from those for GPA and thus discussion concentrates on analysis of the continuous measure of achievement (i.e. GPA). A supplemental table presents results for college attendance.

The gene-environment literature has generally reported results without taking multiple hypothesis testing into account. Significance is therefore overstated. Results shown here account for multiple hypotheses, using the relatively conservative Bonferroni correction (α/n) to increase the significance level required. Three genes and three outcomes are investigated below, which could be interpreted as three or nine distinct tests depending on what one considers independent hypotheses. For example, the three outcomes – depression, delinquency, and high school GPA – are likely related. An effect of birth weight on any of them would likely manifest in the others as well. Therefore, tables report general 95% significance levels, with shading to indicate significance after adjustment for the 3 hypotheses of genetic markers. Discussion of results in the text report significance levels with a correction for both 3 and 9 hypotheses, but this is an overly cautious adjustment to the extent that outcomes are related.

Results

Table 1 shows main effects for each outcome. In Panel 1, models predicting each outcome include only the gene which previous research predicts should have effects. Panel 2 includes all three genes. Despite previous evidence that the short 5HTT allele increases depression symptoms, DZ twin fixed effect regressions controlling for a variety of potential confounders show no significant effect of this genotype on depression. Main effects shown in Table 1 are insignificant, whether including all three genes or only 5HTT. Results are similar using the natural log of depression symptoms.

Prior research has also suggested an effect of MAOA on delinquency and claimed that DRD2 affects school continuation. Using Add Health genetic data, for example, Guo et al.

(2008b) find a significant relationship between the short 2R MAOA allele and serious and violent delinquency. Among our analyses of twins, however, main effects of variation at these genetic loci are nearly all insignificant (controlling for gender and birth weight). An exception is identical twin random effects models; having no copies of the risky DRD2 allele is significantly related to highest grade completed and having some college education, but not to high school GPA.[5] These effects, however, emerge only in identical twin random effects models, which estimate the effect of the gene between (as opposed to within) twin pairs, since genotype is always the same for identical twins. Therefore, these significant main effects could reflect population stratification – with genotype potentially non-randomly assigned to different twin pairs. Similarly, as shown in Panel 2 of Table 1, identical twins with two copies of the 4R MAOA allele show significantly lower depression (raw and logged scores), but this relationship disappears when including those with missing birth weight and only holds in identical twin random effects models—i.e. across families.

While rare significant relationships emerge (e.g., lower high school GPA for all and fraternal twins with the "good" MAOA alleles, but only when including those missing birth weight), the overarching pattern is that main genetic effects are insignificant regardless of the specifications tested. In other words, main effects are insignificant in these models (or at least inconsistent across twin types) when comparing those with two copies of the risky allele to all others or when using alternate specifications of the outcome variable.

Similarly, main effects of birth weight are insignificant in all of the identical and fraternal twin models shown (when an interaction with genotype is not included). Exceptions (in models not shown) include a significant *positive* effect of birth weight on the gender-

---

[5] Results predicting some college are presented in supplemental table S1 for comparison with Shanahan et al. (2008).

specific depression cutoff measure among identical twins and, in models including indicators for 5HTT alone or all three genes investigated here together, a significant association with lower likelihood of having any depressive symptoms among fraternal twins. The overwhelming absence of birth weight effects contradicts Conley and Bennett (2000), who use singleton comparisons, which fail to address underlying genetic differences. The absence of main genetic effects suggests previous results may be biased by population stratification.

Research by Caspi (2003) and Wilhelm et al. (2006) predicts that those with the short 5HTT alleles should be more sensitive to and benefit more from environmental advantages. However, this research failed to control for a variety of genetic and environmental confounders. Identical twin comparisons address this potential spuriousness and reveal that additional fetal nourishment actually *increases* depression among those homozygous for the risky allele. Weighing an additional pound (above the MZ mean of 90 oz or 5.6 lbs) increases depression symptoms by over 4 points (over 11 standard deviations) for those with two copies of the short allele, but has no effect for others. Specifying the gene differently and comparing those without any short alleles to others shows that birth weight nudges individuals in different behavioral directions depending on genotype. Birth weight increases depressive symptoms for those without any long alleles (2.6 points or 7.5 standard deviations), but decreases symptoms for those with 2 long alleles (by 4.5 points or 13 standard deviations). Figure 1 illustrates this gene-environment interaction.

Results for depression are shown in Table 2. Effects of birth weight depend on 5HTT genotype when including all or identical twins, but not fraternal twins alone. Model 3, limited to identical twins, suggests that an additional pound at birth, compared to one's twin, reduces depressive symptoms by almost 4 ½ points for those with two long alleles. That same pound

increases depressive symptoms by 2.6 points for those with one or two short alleles. The interaction effects shown remain significant even with a Bonferroni correction for 3 or 9 hypotheses. (The main effect of birth weight in model 3 remains only marginally significant [p=.054] with a conservative correction for 9 hypotheses.) Running these regressions separately by genotype suggests that there is no effect among those with two long alleles, while birth weight increases depression among others. Results (not shown) are null using an indicator for any depressive symptoms, but are similar using a logged measure of depression or including twins with missing birth weight (although these specifications are sometimes marginally significant and less robust to Bonferroni corrections). Thus, contrary to findings in previous research, environmental advantages appear to *harm* rather than help those with the risky genotype. Meanwhile, birth weight effects do not depend on MAOA and DRD2 genotype. A marginal interaction between the long MAOA allele and birth weight is found among all twins (Model 4 in Table 2). But this disappears among identical twins and is in the opposite direction as would be expected from the prior literature; that is, it is among those with the "good" allele that birth weight seems to matter.

Caspi and his colleagues (2002) found a significant interaction between MAOA and maltreatment on delinquency. However, correcting for potential genetic and environmental confounders with twin fixed effects, and specifying MAOA in various ways, there is little evidence that MAOA affects delinquent behavior. Further, within identical twin pairs, environmental advantage does not moderate genetic effects (see Table 3). Caspi found a stronger relationship between risky MAOA genotype and delinquency for those exposed to childhood distress (specifically maltreatment). We found no support for this; the random environmental advantage of fetal nutrition did not affect the relationship between MAOA and

delinquency. The results shown in Table 3 are similar when using raw or logged measures of delinquency and including pairs missing birth weight information for one twin. The only exception is a significant estimate for MAOA and its interaction with birth weight when predicting the indicator for any delinquent behavior or logged scores, but only when specifying those with at least one 4R allele (rather than two) and only among all and fraternal twins. (Using logged scores, this interaction only holds among all twins when excluding pairs missing birth weight.) Effects disappear within identical twins and using other specifications of the alleles.

Finally, using Add Health data, Shanahan et al. (2008) find that the A1 allele at the DRD2 gene is related to a lower likelihood of post-secondary school for boys. They indicate that this risk is moderated by social capital, yet they also find that boys with the risky DRD2 genotype are less likely to have high social capital environments. In this case, the risky allele could be interacting with non-randomly distributed genes rather than with social capital (even if the principal gene in question is in fact randomly distributed). Indeed, results from twin comparisons shown in Table 4 suggest previous research was biased by omitted genetic differences. Using the whole sample of twins, models including DRD2 and birth weight interactions suggest that, contrary to Shanahan et al.'s findings, those with no copies of the risky A1 allele have significantly *lower* high school GPAs and experience a boost from fetal nutrition (i.e. are more sensitive to environmental insults). This interaction is marginally significant among fraternal twins and disappears within identical twins (Model 3). When twin pairs missing complete birth weight data are included, the interaction effect is only marginally significant using the whole sample of twins and insignificant in all others. Genetic differences between fraternal twins appear to account for apparent environmental interaction effects with

DRD2. (Supplemental Table S1 shows results for college attendance. No significant interaction effects with DRD2 emerge for years of education or college attendance.)

We find a similar pattern for MAOA when predicting highest grade completed—tables not shown. Fraternal twins with two copies of the long 4R MAOA allele experience significantly higher educational attainment, but experience a slight educational *penalty* for additional birth weight. However, as with results for DRD2, these effects are only significant among fraternal twins. They disappear among identical twins and are only marginally significant for all twins. Effects disappear when pairs missing complete birth weight information are included. Population stratification could account for these apparent GE interaction effects.

Discussion and Conclusions

Caspi (2003) presented evidence that sensitivity to environmental insults increases with each short 5HTT allele – i.e., those with one short and one long allele fall between those with 2 copies of either. There is some evidence for this. Compared to those with two copies of the long allele (the "good" allele), heterozygotes show significantly higher depression with increased birth weight – an additional 3.7 points (over 10 standard deviation units) with each pound. (This remains only marginally significant [p=.066] with a Bonferroni correction of 3). Those with two copies of the short allele show the strongest interaction with this specification, increasing 5.7 points on the depression scale (over 16 standard deviation units) with an additional pound at birth. This effect remains significant for those with two copies of the short allele even with the Bonferroni correction of 9 (p=.027).

Even with the Bonferroni correction, results confirm that effects of 5HTT depend on environmental input, as others have found. However, contrary to previous research, those at most genetic risk for depression (two short alleles) are actually hurt by the environmental boost of fetal nutrition. Individuals with "good" genes show reduced depression with increased birth weight. Unfortunately, this seems to be the old story of cumulative advantage.

The pattern of null findings for gene and gene-environment interaction effects within twin pairs suggest that previous significant findings were likely biased due to population stratification or omitted environmental factors. Genetic differences are distributed non-randomly across the population. The absence of effects within identical twin pairs suggests that previous findings may reflect a failure to account for non-random relationships between environmental and genetic differences.

A potential concern in this research is that the genes in question could influence birth weight or birth weight differences within twin pairs. If the risky MAOA gene, for example, is associated with smaller birth weight differences between DZ twins, the reduced birth weight variation could yield insignificant interactions. However, regressions checking for such an association suggest it is not a concern (see Table 5). Among MZ twin pairs, no genes are associated with twin pair birth weight difference or average. Looking at 5HTT alone does not change results; 5HTT is not significantly related to MZ twin difference in birth weight. Among DZ twins, pairs in which one twin has a risky DRD2 allele have slightly more similar birth weights. This effect is not robust to a Bonferroni correction for 3 hypotheses, is insignificant among identical twins, and is not found when specifying those with two risky DRD2 alleles.

At the individual level, these genes are still unrelated to birth weight. Within fraternal

twin pairs, none of these genes significantly affect birth weight. Identical twins of course share any risky genes, but estimates using random rather than fixed effects show no relationship between genes and birth weight. (Individuals with two copies of the short 5HTT allele have slightly higher birth weight, but this relationship is only marginally significant – even without Bonferroni correction – and is insignificant when specifying those with two copies of the long allele instead). Including those with missing birth weight does not change the results.

A limitation of this study is the potentially selective data on birth weight. Approximately 20% of the sample is missing parent-reported birth weight. An additional 10% is missing birth weight data for their twin. Thus, about a third of all identical twins are excluded. Both identical and fraternal twins missing birth weight information for their twin sibling were born significantly lighter on average than other twins. This suggests parents who did not provide birth weight data may have had lighter babies and excluded twins may weigh less than those included in the study.

Low weight babies experienced the strongest environmental insult and their outcomes may be most sensitive to genotype. Under-representation of low weight babies may therefore cause attenuation bias in estimates of gene-environment interaction effects. The differences are slight, however, amounting to just over an ounce for both MZ and DZ twins, which suggests minimal bias. An alternative story further reduces our concern about missing birth weight. Parents may better remember the birth weight of the lighter rather than the healthier twin. In this scenario, those missing birth weight would be nearer the twin average.

A further concern is the potential relationship between missing birth weight data and outcomes or genetic measures. Table 6 provides comparisons of these two groups: twins with and without complete birth weight data. MZ twins with complete birth weight info have

significantly higher delinquency rates than those without – amounting to about ½ an additional delinquent act such as stealing or damaging property. This difference remains significant with binary (indicating any symptoms) and logged measures of delinquency (after adding one to raw scores to include those with scores of zero). Twins with and without birth weight data show no differences in the specific alleles investigated here.

These birth weight and delinquency differences, though slight, suggest missing birth weight data could be non-random. However, only one (delinquency among MZ twins) remains significant at the 95% level with the multiple hypothesis adjustment. In addition, potential bias due to non-random missing birth weight is investigated for all models using the missing indicator method. Results are similar. Finally, half of those cases missing birth weight have valid information for pounds at birth, but are missing ounces. When these individuals are assigned 7 ounces (the midpoint, plus their birth weight in pounds), results are similar to those presented here.

To summarize, we find no evidence for the GE interaction effects putatively demonstrated by previous scholars using the same dataset (Add Health) or another (the Dunedin sample of Caspi et al.). We argue that this and other previous work has spuriously assumed GE associations are actually GE effects when the environmental measures have been endogenous to unmeasured genotype. In fact, when we use an exogenous source of environmental variation—within twin-set differences in birth weight—the only significant result we obtain works in exactly the opposite direction of that proposed by Caspi et al. Beyond the specific results presented here, we purport to have developed a careful method for assessing GE interaction effects: by deploying both MZ and DZ comparisons, which each complement the other's inferential weakness. (And for direct effect of genetic locus variation,

DZ twin comparisons alone suffice.) We encourage other researchers to deploy the methodology presented here.

Finally, our analysis calls into question the research literature on birth weight. Not only do we find little to no main effects of birth weight on important behavioral and academic outcomes that would seem to be directly in a causal pathway to previous findings regarding birth weight and educational attainment, we find that when interacted with certain alleles, birth weight works in the opposite direction as previously supposed. This, too, calls for further investigation by future researchers.

References

Behrman, Jere R. and Mark R. Rosenzweig. 2004. "Returns to Birthweight." *The Review of Economics and Statistics* 86(2): 586–601

Beauchamp, Jonathan, David Cesarini, J. Niels Rosenquist, James H. Fowler, and Nicholas A. Christakis. 2009. "A Genome Wide Association Study of Educational Attainment." A paper presented at the IZA (Institute for the Study of Labor) Workshop: Genes, Brains, and the Labor Market in Bonn, Switzerland. www.neuroeconomics.nyu.edu/papers/Cesarini1.pdf.

Bennett, A. J., K.P. Lesch, A. Heils, J.C. Long, J.G. Lorenz, S.E. Shoaf, M. Champoux, S.J. Suomi, M.V. Linnoila,   and J.D. Higley. 2002. "Early Experience and Serotonin Transporter Gene Variation Interact to Influence Primate CNS Function." *Molecular Psychiatry* 7**:** 118-122.

Blum, K., E.P. Noble, P.J. Sheridan, O. Finley, A. Montgomery, T. Ritchie, et al. 1991. "Association of the A1 Allele of the D2 Dopamine Receptor Gene with Severe Alcoholism." *Alcohol* 8: 409-416.

Brunner, H.G., M. Nelen, X.O. Breakefield, H.H. Ropers, B.A. van Oost. 1993. "Abnormal Behavior Associated with a Point Mutation in the Structural Gene for Monoamine Oxidase A." *Science* 262, 5133: 578-580.

Cases, Olivier, Isabelle Seif, Joseph Grimsby, Patricia Gaspar, Kevin Chen, Sandrine Pournin, Ulrike Muller, Michel Aguet, Charles Babinet, Jean Chen Shih, Edward De Maeyer. 1995. "Aggressive Behavior and Altered Amounts of Brain Serotonin and Norepinephrine in Mice Lacking MAOA." *Science* 268: 1763-1766.

Caspi, Avshalom et al. 2002. "Role of Genotype in the Cycle of Violence in Maltreated Children." *Science* 297: 851-854.

Caspi, Avshalom et al. 2003. "Influence of Life Stress on Depression: Moderation by a Polymorphism in the 5-HTT Gene." *Science* 297: 851-854.

Conley, Dalton and Neil G. Bennett. 2000. "Is Biology Destiny? Birth Weight and Life Chances." *American Sociological Review* 65: 458–467.

Conley, Dalton and Neil G. Bennett. 2001. "Birthweight and Income: Interactions Across Generations." *Journal of Health and Social Behavior* 42(4): 450-465.

Conley, Dalton, Kate W. Strully, and Neil G. Bennett. 2003. *The Starting Gate: Birth Weight and Life Chances*. Berkeley: University of California Press.

Cooper, C., Harvey N, Cole Z, Hanson M, Dennison E. 2009. "Developmental origins of osteoporosis: the role of maternal nutrition." *Advances in Experimental Medicine and Biology* 646:31-9.

Dobbs, David. 2009, December. "The Science of Success." *The Atlantic*. Retrieved 1/10/2010 from http://www.theatlantic.com/doc/200912/dobbs-orchid-gene.

Druet, C. and K.K. Ong. 2008. "Early childhood predictors of adult body composition." *Best Practice and Research Clinical Endocrinology and Metabolism* 22, 3:489-502.

Ellis, Bruce J. and W. Thomas Boyce. 2008. "Biological Sensitivity to Context." *Current Directions in Psychological Science* 17, 3: 183-187.

Fletcher, Jason M. and Steven F. Lehrer. 2009. "Using Genetic Lotteries within Families to Examine the Causal Impact of Poor Health on Academic Achievement." NBER Working Paper 15148. http://www.nber.org/papers/w15148.

Guo, Guang, Michael E. Roettger and Tianji Cai. 2008a. "The Interaction of Genetic Propensities into Social-Control Models of Delinquency and Violence among Male Youths." *American Sociological Review* 73: 543-568.

Guo, Guang, Xiao-Ming Ou, Michael Roettger, and Jean C. Shih. 2008b. "The VNTR 2 Repeat in MAOA and Delinquent Behavior in Adolescence and Young Adulthood: Associations and MAOA Promoter Activity." *European Journal of Human Genetics* 16: 626–634.

Guo, Guang and Elizabeth Stearns. 2002. "The Social Influences on the Realization of Genetic Potential for Intellectual Development." *Social Forces* 80(3): 881-910.

Hack, M., D. Flannery, M. Schluchter, L. Cartar, E. Borawski, and N. Klein. 2002. "Outcomes in Young Adulthood for Very Low Birth Weight Infants." *The New England Journal of Medicine* 346(2): 149-157.

Hamer, D. 2002. "Rethinking Behavior Genetics." *Science* 298(5591): 71-72.
Harris, Kathleen Mullan, Carolyn Tucker Halpern, Andrew Smolen, and Brett C. Haberstick. 2006. "The National Longitudinal Study of Adolescent Health (Add Health) Twin Data." *Twin Research and Human Genetics* 9, 6: 988-997.

Harris, Kathleen Mullan. 2009. The National Longitudinal Study of Adolescent Health (Add Health), Waves I & II, 1994–1996; Wave III, 2001–2002 [machine-readable data file and documentation]. Chapel Hill, NC: Carolina Population Center, University of North Carolina at Chapel Hill.

Hayes, B. and F. Sharif. 2009. "Behavioural and emotional outcome of very low birth weight infants--literature review." *Journal of Maternal-Fetal and Neonatal Medicine* 22, 10:849-56.

Jeong, H., S. P. Mason, A.-L. Barabási and Z. N. Oltvai. 2001. Lethality and centrality in protein networks. *Nature* 411: 41-42.

Lawford, B.R., R. Young, E.P. Noble, B. Kann, T. Ritchie. 2006. "The D-2 Dopamine Receptor (DRD2) Gene Is Associated with Co-Morbid Depression, Anxiety and Social Dysfunction in Untreated Veterans with Post-Traumatic Stress Disorder." *European Psychiatry* 21: 180-185.

Lowe J, Erickson SJ, Maclean P, Duvall SW. 2009. "Early working memory and maternal communication in toddlers born very low birth weight." *Acta Paediatrica* 98, 4:660-3.

Meas, T., S. Deghmoun, P. Armoogum, C. Alberti, C. Levy-Marchal. 2008. "Consequences of being born small for gestational age on body composition: an 8-year follow-up study." *Journal of Clinical Endocrinology and Metabolism* 93, 10: 3804-9.

Murphy, Dennis L., Qian Li, Sharon Engel, Christine Wichems, Anne Andrews, Klaus-Peter Lesch, and George Uhl. 2001. "Genetic Perspectives on the Serotonin Transporter." Brain Research Bulletin 56(5): 487-494.

Noble, E.P., T.Z. Ozkaragoz, T.L. Ritchie, X. Zhang, T.R. Belin, and R.S. Sparkes. 1998. "D2 and D4 Dopamine Receptor Polymorphisms and Personality." *American Journal of Medical Genetics: Neuropsychiatric Genetics* 81: 254-267.

Pescosolido, Bernice A., Brea L. Perry, J. Scott Long, Jack K. Martin, John I. Nurnberger, Jr., and Victor Hesselbrock. 2008. "Under the Influence of Genetics: How Transdisciplinarity Leads Us to Rethink Social Pathways to Illness." *American Journal of Sociology* 114 Suppl.: S171–S201

Price, Alkes L., Nick J. Patterson, Robert M. Plenge, Michael E. Weinblatt, Nancy A. Shadick, and David Reich. 2006. "Principal Components Analysis Corrects for Stratification in Genome-Wide Association Studies." *Nature Genetics* 38, 8: 904-909.

Pohjalainen, T., Rinne, J. O., Någren, K., Lehikoinen, P., Anttila, K., Syvälahti, E. and Hietala, J. 1998. "The A1 Allele of the Human D_2 Dopamine Receptor Gene Predicts Low D_2 Receptor Availability in Healthy Volunteers." *Molecular Psychiatry* 3, 256–260.

Rich-Edwards, J. 1997. "Birth Weight and Risk of Cardiovascular Disease in a Cohort of Women Followed Up Since 1976." *British Medical Journal* 35:396-400.

Roberts, R.E., Lewinsohn, P.M., and J.R. Seeley. 1991. "Screening for Adolescent Depression: A Comparison of Depression Scales." *Journal of the American Academy of Child & Adolescent Psychiatry* 30(1): 58-66

Rowe, D.C. 2001. *Biology and Crime*. Los Angeles: Roxbury.

Schlotz, W. and DI. Phillips. 2009. "Fetal origins of mental health: evidence and mechanisms." *Brain, Behavior, and Immunity* 23, 7:905-16.

Shanahan, Michael J., Stephen Vaisey, Lance D. Erickson, and Andrew Smolen. 2008. "Environmental Contingencies and Genetic Propensities: Social Capital, Educational Continuation, and Dopamine Receptor Gene DRD2." *American Journal of Sociology*, 114 Supplement: S260-S286.

Shih, J. C. and R. F. Thompson. 1999. "Monoamine Oxidase in Neuropsychiatry and Behavior." American Journal of Human Genetics 65:593–98.

Sorensen, H., S. Sabroe, J. Olsen, K. Rothman, M. Gillman, and P. Fischer. 1997. "Birth Weight and Cognitive Function in Young Adult Life: Historical Cohort Study." *British Medical Journal* 315:401-403.

Stelzl, Ulrich, Uwe Worm, Maciej Lalowski, Christian Haenig, Felix H. Brembeck, Heike Goehler, Martin Stroedicke, Martina Zenkner, Anke Schoenherr, Susanne Koeppen, Jan Timm, Sascha Mintzlaff, Claudia Abraham, Nicole Bock, Silvia Kietzmann, Astrid Goedde, Engin Toksöz, Anja Droege, Sylvia Krobitsch, Bernhard Korn, Walter Birchmeier, Hans Lehrach and Erich E. Wanker. 2005. "A Human Protein-Protein Interaction Network: A Resource for Annotating the Proteome." *Cell*. 122: 957-968.

Wiers, R.W., J.A. Sergeant, and W.B. Gunning. 1994. "Psychological Mechanisms of Enhanced Risk of Addiction in Children of Alcoholics: A Dual Pathway?" *Acta Paediatrica Supplement* 404: 9-13.

Wilhelm, Kay, Philip B. Mitchell, Heather Niven, Adam Finch, Lucinda Wedgwood, Anna Scimone, Ian P. Blair, Gordon Parker and Peter R. Schofield. 2006. "Life events, first depression onset and the serotonin transporter gene." *The British Journal of Psychiatry* 188: 210-215.
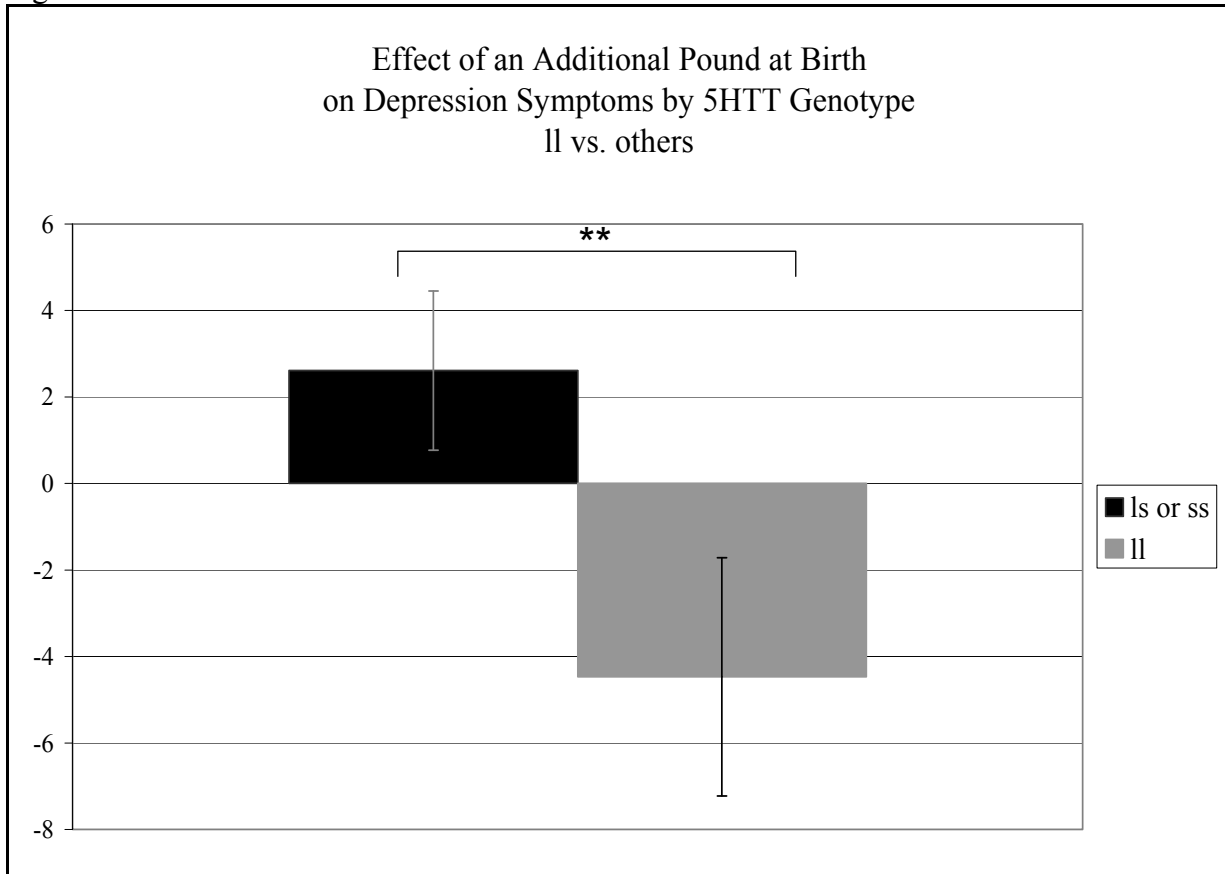
Figure 1

Effect of an Additional Pound at Birth
on Depression Symptoms by 5HTT Genotype
ll vs. others

Table 1: Main Effects of Genes and Birth Weight
Panel 1 – Including Hypothesized Gene

VARIABLES	(1) Depression	(2)	(3)	(4)	(5) Delinquency	(6)	(7)	(8)	(9) HS GPA	(10)	(11)	(12)
	Twins	DZ	MZ	MZ-RE	Twins	DZ	MZ	MZ-RE	Twins	DZ	MZ	MZ-RE
5HTT ll	-0.31	-0.32	0	0.08								
	(1.10)	(1.14)	(0)	(0.88)								
MAOA both 4R					-0.40	-0.39	0	-0.32				
					(0.39)	(0.44)	(0)	(0.26)				
DRD2 no A1									-0.05	-0.05	0	0.17
									(0.13)	(0.15)	(0)	(0.15)
Log oz	1.86	0.86	3.78	-0.16	-0.49	-0.86	0.24	0.58	-0.02	0.13	-0.30	-0.19
	(2.73)	(3.49)	(4.45)	(2.30)	(0.94)	(1.30)	(1.29)	(0.69)	(0.30)	(0.45)	(0.33)	(0.28)
Male	-0.93	-0.91	0	-0.30	0.84**	0.85**	0	0.43+	-0.45**	-0.46**	0	-0.34*
	(0.80)	(0.83)	(0)	(0.82)	(0.29)	(0.32)	(0)	(0.26)	(0.09)	(0.11)	(0)	(0.15)
Constant	-2.17	2.28	-11.27	6.54	2.65	4.31	-0.46	-2.04	3.07*	2.31	4.16**	3.77**
	(12.32)	(15.76)	(20.04)	(10.33)	(4.24)	(5.86)	(5.81)	(3.09)	(1.37)	(2.03)	(1.49)	(1.24)
Observations	486	280	206	206	456	270	186	186	392	232	160	160
R-squared	0.618	0.593	0.654		0.635	0.611	0.692		0.847	0.815	0.912	
No. of pairs	243	140	103	103	228	135	93	93	196	116	80	80

Panel 2 – Including All Three Genes

VARIABLES	(1) Depression	(2)	(3)	(4)	(5) Delinquency	(6)	(7)	(8)	(9) HS GPA	(10)	(11)	(12)
	Twins	DZ	MZ	MZ-RE	Twins	DZ	MZ	MZ-RE	Twins	DZ	MZ	MZ-RE
5HTT ll	-0.46	-0.47	0	0.54	-0.65	-0.66	0	0.25	0.21	0.22	0	-0.26
	(1.17)	(1.20)	(0)	(0.95)	(0.40)	(0.44)	(0)	(0.30)	(0.14)	(0.16)	(0)	(0.17)
MAOA both 4R	0.36	0.39	0	-1.760*	-0.27	-0.25	0	-0.39	-0.27+	-0.27	0	0.07
	(1.21)	(1.25)	(0)	(0.85)	(0.41)	(0.46)	(0)	(0.27)	(0.14)	(0.17)	(0)	(0.15)
DRD2 no A1	0.75	0.76	0	-0.26	0.014	0.02	0	-0.04	-0.04	-0.04	0	0.25
	(1.12)	(1.16)	(0)	(0.88)	(0.37)	(0.42)	(0)	(0.28)	(0.13)	(0.15)	(0)	(0.16)
Log oz	1.69	0.93	3.16	-0.79	-0.57	-0.98	0.25	0.53	-0.00	0.10	-0.19	-0.18
	(2.81)	(3.57)	(4.63)	(2.40)	(0.96)	(1.31)	(1.34)	(0.73)	(0.30)	(0.45)	(0.32)	(0.27)
Male	-1.19	-1.18	0	-0.36	0.80**	0.81*	0	0.37	-0.40**	-0.41**	0	-0.39**
	(0.87)	(0.89)	(0)	(0.85)	(0.29)	(0.33)	(0)	(0.27)	(0.09)	(0.11)	(0)	(0.15)
Constant	-1.78	1.46	-8.35	10.46	3.18	5.02	-0.49	-1.83	3.02*	2.48	3.68*	3.77**
	(12.67)	(16.10)	(20.84)	(10.83)	(4.32)	(5.90)	(6.01)	(3.29)	(1.36)	(2.03)	(1.43)	(1.22)
Observations	460	270	190	190	446	268	178	178	376	222	154	154
R-squared	0.615	0.592	0.646		0.636	0.617	0.682		0.856	0.823	0.922	
No. of pairs	230	135	95	95	223	134	89	89	188	111	77	77

Standard errors in parentheses
** p<0.01, * p<0.05, + p<0.1

NOTE – If twin pairs missing one birth weight value are included, with a dummy indicator for missing birth weight, genetic effects are still insignificant. Main genetic effects on logged depression and delinquency scores are also null. Variation in birth weight and genes alone do not significantly predict these outcomes.

Table 2: Interaction Effects – Depression

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
					Depression				
	Twins	DZ	MZ	Twins	DZ	MZ	Twins	DZ	MZ
Log oz_5HTT-ll	-12.32**	-8.40+	-27.33**						
	(4.27)	(5.03)	(8.59)						
5HTT-ll	55.38**	37.62	0						
	(19.31)	(22.74)	(0)						
Log oz_MAOA both 4R				-9.65*	-8.90	-12.07			
				(4.58)	(5.45)	(8.96)			
MAOA both 4R				43.80*	40.43	0			
				(20.81)	(24.77)	(0)			
Log oz_DRD2 no A1							-5.36	-7.45	2.48
							(4.49)	(5.26)	(9.19)
DRD2 no A1							25.00	34.45	0
							(20.32)	(23.77)	(0)
Log oz	6.90*	4.13	15.95**	6.44+	5.45	8.82	4.78	5.23	2.61
	(3.21)	(3.98)	(5.73)	(3.48)	(4.35)	(6.14)	(3.66)	(4.60)	(6.25)
Male	-0.95	-0.92	0	-1.01	-1.009	0	-0.91	-0.89	0
	(0.79)	(0.83)	(0)	(0.85)	(0.884)	(0)	(0.81)	(0.83)	(0)
Constant	-24.94+	-12.49	-28.03	-22.99	-18.54	-7.072	-15.90	-18.08	-11.81
	(14.47)	(17.98)	(19.91)	(15.71)	(19.60)	(20.13)	(16.51)	(20.75)	(20.81)
Observations	486	280	206	470	272	198	482	282	200
R-squared	0.631	0.601	0.685	0.623	0.595	0.662	0.616	0.597	0.645

Standard errors in parentheses

** p<0.01, * p<0.05, + p<0.1

Shading indicates significance with Bonferroni correction for 3 hypotheses: p<.05; p<.10

Table 3: Interaction Effects – Delinquency

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
				Delinquency Score					
	Twins	DZ	MZ	Twins	DZ	MZ	Twins	DZ	MZ
Log oz_MAOA both 4R	-1.47	-1.61	-1.01						
	(1.57)	(2.02)	(2.60)						
MAOA both 4R	6.28	6.93	0						
	(7.14)	(9.17)	(0)						
Log oz_5HTT-ll				-2.05	-1.82	-3.27			
				(1.50)	(1.89)	(2.64)			
5HTT-ll				8.75	7.70	0			
				(6.78)	(8.55)	(0)			
Log oz_DRD2 no A1							0.26	0.17	0.89
							(1.81)	(1.99)	(4.10)
DRD2 no A1							-1.14	-0.71	0
							(8.20)	(8.98)	(0)
Log oz	0.20	-0.10	0.72	0.27	-0.48	2.13	-0.91	-1.22	-0.46
	(1.20)	(1.61)	(1.78)	(1.13)	(1.50)	(1.78)	(1.48)	(1.74)	(2.82)
Male	0.86**	0.87**	0	0.78**	0.80*	0	0.80*	0.81**	0
	(0.29)	(0.33)	(0)	(0.28)	(0.31)	(0)	(0.32)	(0.31)	(0)
Constant	-0.48	0.90	-0.40	-0.76	2.64	-4.25	4.39	5.77	0.72
	(5.40)	(7.26)	(5.84)	(5.11)	(6.77)	(6.13)	(6.68)	(7.82)	(9.28)
Observations	456	270	186	472	278	194	468	280	188
R-squared	0.637	0.613	0.693	0.631	0.610	0.682	0.628	0.604	0.654

Standard errors in parentheses
** $p<0.01$, * $p<0.05$, + $p<0.1$

Table 4: Interaction Effects – Cumulative High School GPA

VARIABLES	(1) Twins	(2) DZ	(3) MZ	(4) Twins	(5) DZ	(6) MZ	(7) Twins	(8) DZ	(9) MZ
				GPA					
Log oz_DRD2 no A1	1.11*	1.41+	0.17						
	(0.51)	(0.71)	(0.67)						
DRD2 no A1	-5.07*	-6.41*	0						
	(2.33)	(3.23)	(0)						
Log oz_MAOA both 4R				0.26	0.30	0.15			
				(0.50)	(0.71)	(0.62)			
MAOA both 4R				-1.40	-1.58	0			
				(2.29)	(3.21)	(0)			
Log oz_5HTT-ll							-0.29	-0.61	0.74
							(0.50)	(0.69)	(0.65)
5HTT-ll							1.49	2.91	0
							(2.25)	(3.11)	(0)
Log oz	-0.64	-0.72	-0.38	-0.15	-0.08	-0.25	0.11	0.38	-0.63
	(0.42)	(0.62)	(0.45)	(0.38)	(0.55)	(0.43)	(0.37)	(0.52)	(0.44)
Male	-0.45**	-0.46**	0	-0.42**	-0.42**	0	-0.45**	-0.45**	0
	(0.09)	(0.11)	(0)	(0.09)	(0.12)	(0)	(0.09)	(0.11)	(0)
Constant	5.85**	6.16*	4.14**	3.72	3.32	3.65*	2.38	1.07	4.62**
	(1.87)	(2.80)	(1.50)	(1.72)	(2.51)	(1.39)	(1.65)	(2.37)	(1.51)
Observations	392	232	160	386	224	162	396	230	166
R-squared	0.851	0.821	0.912	0.856	0.823	0.922	0.848	0.816	0.913

Standard errors in parentheses

** $p < 0.01$, * $p < 0.05$, + $p < 0.1$

Shading indicates significance with Bonferroni correction for 3 hypotheses: $p < .05$; $p < .10$

Table 5: Predicting Birth Weight

VARIABLES	(1) Log oz diff	(2) Log oz diff	(3) Log oz	(4) Log oz
Twin pair	MZ	DZ	MZ – RE	DZ
DRD2 both no A1	-0.019	-0.004		
	(0.021)	(0.023)		
DRD2 one no A1	0	-0.057*		
	(0)	(0.027)		
5HTT both ll	0.026	-0.015		
	(0.022)	(0.024)		
5HTT one ll	0	-0.016		
	(0)	(0.025)		
MAOA both 4R	-0.005	0.002		
	(0.020)	(0.022)		
MAOA one 4R	0	0.007		
	(0)	(0.025)		
Male	0.013	-0.018	0.050	0.019
	(0.020)	(0.019)	(0.031)	(0.022)
Individual				
DRD2 - no A1			-0.018	0.013
			(0.033)	(0.028)
MAOA - 4R			0.010	0.032
			(0.032)	(0.030)
5HTT - ll			-0.051	-0.021
			(0.036)	(0.029)
Constant	0.093**	0.145**	4.500**	4.496**
	(0.019)	(0.026)	(0.031)	(0.026)
Observations	95	136	190	272
R-squared	0.025	0.047		0.798
Number of pairs	95	136	95	136

Standard errors in parentheses
** p<0.01, * p<0.05, + p<0.1

Table 6: Twins with and without Complete Birth Weight Data

	MZ w/ twin BW			MZ w/o twin BW			DZ w/ twin BW			DZ w/o twin BW		
	Mean	Std Error	N	Mean	Std Error	N	Mean	Std Error	N	Mean	Std Error	N
Log oz	4.5*	0.01	208	4.37*	0.05	15	4.52*	0.01	285	4.42*	0.04	28
Cumulative GPA	2.82*	0.05	181	2.62*	0.08	89	2.61+	0.06	252	2.43+	0.08	104
Highest Grade Comp	13.41	0.13	208	13.08	0.18	100	13.31+	0.12	285	12.96+	.16	129
Depression	5.71	0.35	208	5.77	0.46	98	5.54	0.30	284	5.80	0.49	127
Log depression	1.63	0.05	208	1.68	0.07	98	1.58	0.05	284	1.61	0.07	127
Any depression	0.93	0.02	208	0.95	0.02	100	0.89	0.02	285	0.93	0.02	129
Delinquency	0.78*	0.15	202	0.29*	0.09	100	0.69	0.11	283	0.77	0.16	129
Log delinquency	0.32*	0.04	202	0.14*	0.04	100	0.29	0.03	283	0.32	0.05	129
Any delinquency	0.30*	0.03	208	0.15*	0.04	100	0.25	0.03	285	0.27	0.04	129
DRD2 – no A1	0.53	0.04	203	0.58	0.05	100	0.61	0.03	285	0.54	0.04	127
MAOA – both 4R	0.49	0.04	200	0.39	0.05	97	0.49	0.3	280	0.42	0.04	127
5HTT – ll	0.31	0.03	206	0.33	0.05	98	0.33	0.03	284	0.32	0.04	128

Indicates significant mean difference (two-tailed ttest, unequal variance) * $p<0.05$, + $p<0.10$
Bonferroni correction of 3 $p<0.05$, $p<0.10$

Log and Any deliquency measures are significant at $p<.05$ after correcting for 9 hypotheses.

Table S1: Main and Interaction Effects – College Attendance
Panel 1: Main Effects

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	College Attendance – Hypothesized Gene				College Attendance – All 3 Genes			
	Twins	DZ	MZ	MZ-RE	Twins	DZ	MZ	MZ-RE
DRD2 no A1	0.07	0.07	0	0.20*	0.06	0.06	0	0.23*
	(0.08)	(0.08)	(0)	(0.09)	(0.08)	(0.08)	(0)	(0.09)
5HTT ll					-0.04	-0.04	0	0.03
					(0.08)	(0.08)	(0)	(0.10)
MAOA both 4R					0.12	0.12	0	0.03
					(0.08)	(0.09)	(0)	(0.09)
Log oz	0.16	0.18	0.12	-0.07	0.11	0.11	0.12	-0.08
	(0.19)	(0.25)	(0.31)	(0.22)	(0.20)	(0.25)	(0.33)	(0.22)
Male	-0.30**	-0.30**	0	-0.12	-0.30**	-0.30**	0	-0.16+
	(0.06)	(0.06)	(0)	(0.09)	(0.06)	(0.06)	(0)	(0.09)
Constant	-0.00	-0.06	0.01	0.83	0.16	0.21	0.01	0.84
	(0.87)	(1.10)	(1.42)	(0.98)	(0.88)	(1.12)	(1.46)	(1.00)
Observations	484	284	200	200	462	272	190	190
R-squared	0.806	0.788	0.828		0.802	0.789	0.818	
No. of pairs	242	142	100	100	231	136	95	95

Panel 2: Interaction Effects

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
				GPA					
	Twins	DZ	MZ	Twins	DZ	MZ	Twins	DZ	MZ
Log oz_DRD2 no A1	-0.14	-0.11	-0.28						
	(0.31)	(0.37)	(0.63)						
DRD2 noA1	0.71	0.56	0						
	(1.42)	(1.69)	(0)						
Log oz_MAOA both 4R				-0.36	-0.59	0.40			
				(0.32)	(0.38)	(0.63)			
MAOA both 4R				1.72	2.79	0			
				(1.46)	(1.71)	(0)			
Log oz_5HTT-ll							-0.14	-0.10	-0.26
							(0.30)	(0.36)	(0.62)
5HTT-ll							0.60	0.45	0
							(1.37)	(1.62)	(0)
Log oz	0.24	0.24	0.25	0.29	0.40	-0.07	0.22	0.22	0.23
	(0.26)	(0.33)	(0.43)	(0.24)	(0.30)	(0.43)	(0.23)	(0.29)	(0.41)
Male	-0.30**	-0.30**	0	-0.29**	-0.29**	0	-0.30**	-0.30**	0
	(0.06)	(0.06)	(0)	(0.06)	(0.07)	(0)	(0.06)	(0.06)	(0)
Constant	-0.35	-0.35	0.09	-0.62	-1.10	-0.02	-0.22	-0.23	-0.13
	(1.16)	(1.47)	(1.43)	(1.10)	(1.36)	(1.42)	(1.03)	(1.29)	(1.44)
Observations	484	284	200	472	274	198	488	282	206
R-squared	0.806	0.788	0.829	0.807	0.794	0.826	0.806	0.784	0.832

Panel 1: Random effects Models 4 and 8 suggest identical twins with no copies of the "risky" A1 allele are significantly more likely to attend college. Results are similar when including pairs with missing birth weight. However, this between-family estimate could reflect population stratification.
Panel 2: Interaction effects between genotype and birth weight are all insignificant.