2013

Candidate Genes and Voter Turnout: Further Evidence on the Role of 5-HTTLPR

Kristen Diane Deppe
University of Nebraska-Lincoln, kddeppe@gmail.com

Scott F. Stoltenberg
University of Nebraska-Lincoln, sstoltenberg2@unl.edu

Kevin B. Smith
University of Nebraska-Lincoln, ksmith1@unl.edu

John R. Hibbing
University of Nebraska-Lincoln, jhibbing1@unl.edu

doi:10.1017/S0003055413000087

Candidate Genes and Voter Turnout: Further Evidence on the Role of 5-HTTLPR

KRISTEN DIANE DEPPE, SCOTT F. STOLTENBERG, KEVIN B. SMITH and JOHN R. HIBBING *University of Nebraska-Lincoln*

Recently in this journal, Charney and English (2012) presented an extensive critique of candidate gene association studies using the widely noted Fowler and Dawes (2008) article on the relationship between self-reported voter turnout and both 5-HTT (serotonin transporter) and MAOA (monoamine oxidase A) as the driving example of their evaluation. Reanalysis of the Fowler and Dawes data by Charney and English, based on four critiques of candidate gene studies, led to the conclusion that neither polymorphism is related to variations in turnout. We add to this empirical debate by conducting an independent test using an original dataset containing 5-HTT data and two separate participation variables: self-reported participation and actual voting records. Our results confirm the original conclusions by Fowler and Dawes on 5-HTT, but also support several of the critiques suggested by Charney and English. We conclude by offering suggestions for the way candidate gene association studies should be interpreted by the discipline and processed by journal editors.

Candidate gene association (CGA) studies test for the possibility that variation in the nucleotide sequence at a certain prespecified locus in the genome correlates with individual differences in a particular phenotype (e.g., a trait or behavior) in a population. Tens of thousands of CGA studies have generated important information on disease, obesity, drug addiction, reading proficiency, cancer risk, personality characteristics, and many other attributes. Though popular and potentially powerful, these studies face several challenges. To date, they clearly indicate that "usual suspect" polymorphisms do not account for substantial portions of the variance in a large range of phenotypes. Moreover, for virtually all phenotypes, from breast cancer to depression, findings from CGA studies replicate poorly. These results promote a growing sense that much of the explanatory "action" in genetics may not be, as originally expected, in the most common polymorphisms, but elsewhere—perhaps in rare polymorphisms (Goldstein 2009), epigenetics (Jablonka and Raz 2009), alternative splicing (Luco et al. 2010), mRNA regulation (Poliseno et al. 2010), or the rate of copy number variants (Mefford et al. 2008), to name just a few possibilities.

Though CGA studies are prevalent in biology, anthropology, sociology, psychology, and behavioral genetics, at this writing only two original published CGA studies plus a single reanalysis using the same data have examined purely political phenotypes (Charney and English 2012; Fowler and Dawes 2008; Settle et al. 2010). Despite this extremely small sample size, CGA studies on political phenotypes fit the pattern evident in other disciplines. They have produced moderately encouraging initial results that inconsistently replicate in subsequent analyses. This record, particularly for a discipline unfamiliar with the history of CGA studies, is likely to promote confusion or wild swings between equally unwarranted optimism and pessimism regarding the connection of candidate genes to political phenotypes. Our intended contribution to this forum is to provide a balanced perspective on the possibilities and limitations of political CGA studies.

FOWLER AND DAWES; CHARNEY AND ENGLISH

In 2008, using the National Longitudinal Study of Adolescent Health (known as Add Health), Fowler and Dawes published the first CGA study in political science. Add Health data include self-reported voting in the (then) most recent presidential election (2000) and allelic information on six commonly studied genetic polymorphisms (available for approximately 2,300 individuals), including MAOA and 5-HTT. The former relates to monoamine oxidase A, an enzyme important in breaking down neurotransmitters, and the latter to a transmembrane protein involved in serotonin (a neurotransmitter) transport. These are logical candidate genes for political phenotypes, given that numerous studies link them to various dimensions of social behavior. Notably, the transcriptionally less efficient allele for upstream regulatory region polymorphisms in both MAOA and 5-HTT (i.e., MAOA u-VNTR "L" and 5-HTTLPR "S") has been associated with socially challenged and even antisocial behavior (see Meyer-Lindenberg et al. 2006, on MAOA and Bertolino et al. 2005, on 5-HTT). Similarly, previous research suggests that effects of allelic variation may be contingent on environmental factors and gene-environment interactions (Caspi et al. 2002; 2003; 2010). It is also important to note that these polymorphisms have been shown

Kristen Diane Deppe is Graduate Student, Department of Political Science, University of Nebraska-Lincoln (kd.anderson@huskers.unl.edu).

Scott F. Stoltenberg is Assistant Professor, Department of Psychology, University of Nebraska-Lincoln (sstoltenberg2@unl.edu).

Kevin B. Smith is Professor, Department of Political Science, University of Nebraska-Lincoln (ksmith1@unl.edu).

John R. Hibbing is Foundation Regents Professor of Political Science and Psychology, Department of Political Science, University of Nebraska-Lincoln (jhibbing1@unl.edu).

This research was made possible by a grant from the National Science Foundation (BCS-08-26828).

to affect activation patterns in response to emotional stimuli within the brain systems important in detecting threat (Buckholtz and Meyer-Lindenberg 2008; Hariri, Drabant, and Weinberger 2006).

Drawing on this research and reasoning that voting is indicative of a commitment to social life (Fowler 2006; Fowler and Kam 2007), Fowler and Dawes hypothesize that the less transcriptionally efficient alleles of MAOA u-VNTR and 5-HTTLPR will correlate with reduced levels of voter turnout. They find a direct effect of MAOA u-VNTR genotype on reported voter turnout and an interaction effect of 5-HTTLPR genotype but no direct effect. Individuals who regularly attend religious services and have at least one version of the more transcriptionally efficient (long) allele are significantly more likely to report they voted.

Charney and English (2012) contest Fowler and Dawes' results on four primary grounds and raise two more general points regarding CGA studies. The first concern is phenotypic specification. Fowler and Dawes follow a fairly standard political science approach to constructing their dependent variable by using a survey item asking respondents, "Did you vote in the most recent presidential election?" They classify positive responders as voters and negative responders as nonvoters. Charney and English argue that this dichotomous formulation is unsatisfactory on several counts. First, it provides no indication of voting frequency even though, as Charney and English put it, "voting behavior refers to a quantitative variable;" in other words, "one votes more or less frequently" (2012: 5). Second, it does not take into consideration that some respondents were voting in their first election (Add Health participants were 18–26 years old), and this situation can be different from that facing returning voters. Finally, the measure of voting behavior employed by Fowler and Dawes is reported rather than observed voting, and it is well known that respondents tend to overreport their voting behavior.

The second basis for Charney and English's challenge is the manner in which Fowler and Dawes account for population stratification. Ethnic populations can exhibit substantial allele-frequency differences throughout the genome, including at 5-HTTLPR and MAOA u-VNTR. In a case-control CGA study, participants are sorted into two groups based on the outcome variable of interest (e.g., voters and nonvoters). To determine if a particular allele of a candidate gene is associated with increased "risk" of being a voter, the frequency of that allele in one group is compared to its frequency in the other. Population stratification arises if (1) the two groups vary in ethnic composition, (2) the ethnic groups in question differ in allele frequencies at that locus, and (3) there is an observed association between the allele in question and the outcome of interest. In this case, the ethnic composition of the two groups rather than the particular allele might account for the observed allele-outcome association. Essentially, population stratification is the classic third variable problem. Many CGA studies attempt to eliminate the effects of population stratification, often by restricting analyses to participants of a single ethnic group or by statistically controlling for ethnicity. Fowler and Dawes pool all participants and then include dummy variables for individual racial and ethnic groups in an attempt to control for stratification. Charney and English assert that this approach is insufficient and conduct their analyses separately for Asians, Native Americans, nonwhite Hispanics, African Americans, and whites, even though this practice (and many others urged by Charney and English) greatly reduces the *N*. Both for MAOA u-VNTR and 5-HTTLPR, their results show statistically significant results only for African Americans and even then only at the .08 significance level (2012: 6).

Third, Charney and English question Fowler and Dawes' approach to classifying genotypes. They correctly point out that the connection between particular alleles and transcriptional efficiency for both MAOA u-VNTR and 5-HTTLPR is not fully resolved. Classifying MAOA u-VNTR alleles is complex because the polymorphism is a variable number tandem repeat on the X chromosome. Charney and English note that Fowler and Dawes do not follow the 5-HTTLPR classification practice used in the well-known Caspi et al. study (2003). In that study, Caspi and his colleagues use a three-condition coding for 5-HTTLPR, separating out those who are homozygous short (s/s), heterozygous (L/s), and homozygous long (L/L), but find that the heterozygous and homozygous short groups behave similarly. In contrast, Fowler and Dawes classify L/s genotypes with L/L genotypes. This practice appears to make a difference; Charney and English report that, when the L/s genotype is grouped with the s/s genotype, the 5-HTTLPR/church attendance interaction is not significantly related to reported voter turnout.

Fourth, Charney and English note that the Add Health data come from an unrepresentative sample consisting of numerous sibling pairs, some of them twins. Traditional studies assume the independence of cases and controls; in contrast many Add Health participants are identical by descent and share genes and environments with at least one other person in the sample. Fowler and Dawes correct for this without reducing their sample size, but Charney and English argue that methodological flaws attend this approach. As an alternative, they randomly sample one individual from each family to create numerous samples of unrelated individuals and find no relationships significant at the .05 level.

In addition to identifying these four primary "problem areas" with the Fowler and Dawes study, Charney and English also note that CGA studies in general confront numerous methodological and theoretical issues, highlighting two of them. The first is the need for "reproducibility." Given the complexity of the genome, epigenetic variation, interactions with other genes, interactions with the environment, and the typically weak candidate gene relationships, claimed genotype-phenotype connections frequently do not replicate in other populations. This pattern is true regardless of the model organism and the phenotype of interest (disease, behavior, physical trait, etc.). Charney and English are

correct when they note that "most gene association studies fail to consistently replicate" (2012: 11).

The second general concern pertains to statistical significance. Given the large number of genetic polymorphisms, the multiple options for categorizing variations at a single locus, a virtually limitless range of variables available to interact with genotypes, and a list of behavioral phenotypes that is constrained only by the imaginations of researchers, the number of potential relationships is enormous. Using a standard .05 alpha level, 1 in 20 of these relationships will be found statistically significant by chance alone. To counter the possibility of Type 1 errors, Charney and English recommend Bonferroni correction procedures in which each alpha level is divided by the total number of associations tested.

FURTHER EVIDENCE ON THE CONNECTION OF 5-HTTLPR AND VOTER TURNOUT

The Fowler-Dawes and Charney-English exchange is important for understanding not only the correlates (or lack thereof) of political participation but also the appropriate manner of incorporating genetic information into the study of social and political behavior. Fowler and Dawes are correct when they state that their study presents "the first results ever to link specific genes to political behavior" (2008: 579). Charney and English would probably be quite comfortable if it would have been the last study ever to do so. Despite these differences, both pairs of scholars appear to agree that the underlying dispute is empirical and that empirical disputes require empirical resolution. Dawes and Fowler present empirical findings produced by their analysis of the Add Health data; Charney and English reanalyze those same data and come to quite different empirical conclusions. Our goal is to shed light on this contentious situation by introducing results from a distinct and original dataset that suffers from none of the problems Charney and English identify with Add Health data. No single replication study can offer the final word and no dataset is perfect, but we hope that our findings will advance understanding and perhaps suggest a middle ground regarding perceptions of the connection of allelic variations to voter turnout and, more generally, an appropriate manner for the discipline to handle CGA studies.

In the summer of 2010 we retained the services of a professional survey organization to recruit a representative sample from the population of adult individuals within easy driving distance of our lab. In exchange for a $50 participation fee, 342 people reported for an approximately 90-minute session in which they answered an array of political and personality questions, had their physiological traits assessed, and provided a saliva sample from which their DNA could be extracted and genotyped. This sample was not large, but had several desirable qualities that address concerns pertaining to the Add Health data. Its demographics appeared reasonable: 54.1% female, average age of 45.6, modal family income category of $40,000–$60,000, and modal educational level of "some college." More importantly, our data and analytical procedures allowed us to address all six of the concerns delineated by Charney and English as well as additional concerns we believe should be acknowledged.

With regard to the need to operationalize turnout tendencies more accurately, in contrast to the single, dichotomous self-report item in Add-Health, we had a range of self-reported participation variables (have you worked in a campaign, contacted an elected official, discussed politics with others, etc.) as well as a variable measuring the number of times each participant actually voted in six recent elections (data acquired from the pertinent secretary of state for the primary and general elections of 2006, 2008, and 2010, coded 0 through 6). Our measures had several advantages: they (1) used actual behavior rather than inflated self-reports; (2) were not dichotomous and therefore recorded variation in voting frequency across a range of election types, including primaries, midterms, and presidential elections; and (3) were not derived from a sample in which a large percentage of individuals were voting for the first or second time.

The sample also minimized problems caused by the need for population stratification because it had extremely few nonwhite participants (29; just 8% of the sample). We could control for race/ethnicity as Fowler and Dawes did or exclude all non-white participants to solve the statistical control problem identified by Charney and English. Merely excluding these individuals meant that any generalizations about the association of a candidate gene with voter turnout could not be extended to other racial and ethnic groups, but it is a conservative analytical approach to address the potential problems created by genetic variation across groups. Similarly, because our data were derived from a random sample, we did not have genetically linked individuals as was the case for the Add Health data. As to concerns regarding genotype classification, we should point out that we only genotyped 5-HTTLPR, so we will have nothing to say about a possible relationship between MAOA and voter turnout. With regard to 5-HTTLPR, the genetic information available to us made it possible to run the association as Fowler and Dawes did, as Charney and English recommend, and with a categorization system that incorporated additional genetic information unavailable to either the Fowler-Dawes and Charney-English teams.

The problem of running large numbers of correlations and then skimming off statistically significant relationships was not relevant to our replication study. We did not collect data with the intention of weighing in on a dispute over a CGA with voter turnout, but because this issue was developing into an important controversy and our data were relevant, the first analysis we conducted after obtaining genotypic data was to replicate the Fowler-Dawes study. Our results were thus already Bonferroni corrected because the number of associations we tested was one. The sixth and final way in which our study met the concerns of Charney and English is that they emphasize the need for association studies to be replicated. Certainly, one replication

TABLE 1. Relationship between 5-HTT and Participation

	Model 1		Model 2		Model 3		Model 4		Model 5	
	B (SE)	*p*	B (SE)	*p*	B (SE)	*p*	B (SE)	*p*	B (SE)	*p*
5-HTT	.302*	.099	.347*	.069	.089	.713	−.067	.744	−.199	.256
	(.183)		(.190)		(.243)		(.205)		(.175)	
5-HTT*Attend	.427**	.046	.407*	.071	.324	.260	−.169	.294	.108	.663
	(.213)		(.224)		(.287)		(.161)		(.226)	
Attendance	−.306	.103	−.296	.135	.100	.693	.048	.665	.004	.967
	(.187)		(.198)		(.253)		(.111)		(.105)	
Male	.158	.299	.111	.488	.088	.662	.170	.270	.175	.257
	(.152)		(.160)		(.201)		(.153)		(.154)	
Age	.048***	.000	.048***	.000	.069***	.000	.048***	.000	.048***	.000
	(.006)		(.006)		(.008)		(.006)		(.006)	
Hispanic	.173	.604	—		.393	.421	.201	.550	.184	.582
	(.333)		—		(.488)		(.335)		(.335)	
Black	−.768	.172	—		−1.743**	.026	−.686	.229	−.793	.164
	(.561)		—		(.778)		(.570)		(.568)	
Nat. Am.	−.174	.756	—		−1.062	.170	−.117	.836	−.116	.562
	(.559)		—		(.773)		(.565)		(.562)	
Asian	−.310	.753	—		−.723	.672	−.403	.682	−.412	.675
	(.986)		—		(1.707)		(.982)		(.982)	
Income	.106**	.018	.117**	.014	.208***	.001	.107**	.018	.110**	.015
	(.045)		(.047)		(.060)		(.045)		(.045)	
Education	.547**	.011	.520**	.025	.795***	.009	.559***	.010	.545**	.012
	(.215)		(.230)		(.301)		(.217)		(.217)	
Partisanship	.132	.380	.182	.247	.208	.301	.096	.529	.108	.476
	(.151)		(.157)		(.200)		(.152)		(.153)	
Intercept	−1.479***	.000	−1.548***	.000	−1.192**	.027	−1.167***	.001	−1.199***	.001
	(.382)		(.399)		(.537)		(.360)		(.358)	
R^2	.253		.243		.329		.240		.241	

*$p < .10$; **$p < .05$; ***$p < .01$

is not enough, but it is exactly the sort of contribution necessary if the scientific process is to cumulate.

OUR RESULTS

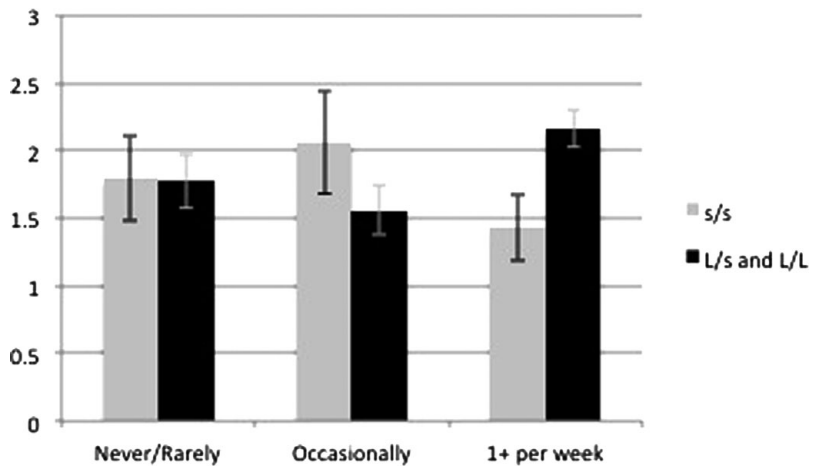
In our data, allele frequencies for 5-HTTLPR ($N = 333$, L = .55, S = .45) and rs25531 ($N = 335$, A = .93, G = .07) were consistent with previous reports, including Fowler and Dawes, and genotypes were in Hardy-Weinberg equilibrium (both $\chi^2 < 1.7$, n.s.). Our genotype call rate was ≥97%, and no discrepancies were observed when we re-genotyped a randomly selected 10% of the participants.

Our Model 1 (see Table 1) substantively replicated Fowler and Dawes's most complete model (their Model 6). Eleven of the 13 independent variables were the same (as mentioned earlier, we did not include a variable for MAOA u-VNTR genotype nor a measure of verbal intelligence—which Fowler and Dawes found to be irrelevant anyway). The only real difference of note between our model and theirs concerned the dependent variable. Their measure of political participation was the aforementioned dichotomous, self-reported vote. Ours was a broader, 0–6 variable indicating the number of the following political activities in which participants claim to have engaged: worked in a campaign, contacted a government official, contributed money to a political cause, attended a political meeting or rally, held political office no matter how minor, and discussed politics with others.

This model's results contain good news for Fowler and Dawes. Even with a completely different dataset and an improved measure of political participation, we found, as they did, that the 5-HTTLPR genotype interacts with church attendance to predict political participation. As was expected, the control variables of age, income, and education were also strongly significant. Our Model 1 even gave evidence of a marginally significant direct effect for the 5-HTTLPR genotype ($p <$ .10) such that "s" allele homozygotes (i.e., s/s) are less likely to report participating. These results suggest that the relationship that Fowler and Dawes report may not be due simply to the peculiarities of the Add Health data, or to running so many interactions that a statistically significant relationship was bound to materialize somewhere, or to their limited, dichotomous measure of political involvement. This replication of the effects of the interaction between 5-HTTLPR genotype and church attendance on political participation lends credence to Fowler and Dawes' original finding.

Figure 1 is a visual display of the relationship, showing the means for reported participation levels by the individual's amount of church attendance and

FIGURE 1. Mean Participation by Church Attendance for 5-HTT Groups (Fowler and Dawes Coding)



5-HTTLPR genotype, combining the L/s and L/L genotypes into one group following Fowler and Dawes' coding. As can be seen, to the extent that the L allele is associated with greater participation, the relationship appears to be confined to individuals who report attending church regularly, at least once a week. As Fowler and Dawes suggest, the combination of an L allele and frequent church attendance may foster self-reported political participation.

That said, further analyses indicate, as suggested by Charney and English, that this relationship is fragile. We continued to examine the association of 5-HTTLPR and participation using Fowler and Dawes' full model, though Charney and English ran their analyses without partisanship, income, and education as controls. The first change we implemented was to drop out the 29 individuals in our sample who identified themselves as something other than "white, non-Hispanic." Recall that Charney and English assert that merely inserting dummy variables does not sufficiently stratify the population. As they note, the simplest way to proceed is to analyze each racial/ethnic group individually. Given the nature of our sample, we could do this individual analysis, but only for whites, without losing many degrees of freedom. The results are reported in Model 2 and show that restricting the analysis to white, non-Hispanic participants again largely replicates the Fowler and Dawes findings, though the significance level for the key interaction term slips from .05 to .10.

The next model modification was more substantial. Model 3 repeated our baseline model (Model 1), but changed the participation measure to the one advocated by Charney and English. Instead of the number of self-reported political acts, the model explained variations in the number of times each participant actually voted in six recent elections (0 to 6). This voting frequency measure was quite similar to the variable suggested by Charney and English; it eliminated concerns about self-reporting biases and provided more complete information than a dichotomous formulation. Critically, the results indicate that the 5-HTTLPR genotype is not associated with actual voting frequency either on its own ($p = .71$) or when considered with church attendance ($p = .26$).

Next, in Model 4, we repeated Model 1, but this time the 5-HTTLPR genotype was categorized the way Charney and English recommend. Our reading of the literature is that some studies do combine heterozygous genotypes with homozygous longs as Fowler and Dawes do, but Charney and English are correct in that the more common procedure is to combine heterozygous genotypes with the homozygous shorts (or to treat them separately). When we used the genotype categorization recommended by Charney and English instead of the one used by Fowler and Dawes, the variable for 5-HTTLPR genotype again fails to be significantly related to political participation (we also tried running it against voting frequency, and the results were no better).

Finally, in Model 5, we presented a modification in genotype classification that is not discussed by either Charney and English or Fowler and Dawes, but should have been. Recent thinking with regard to 5-HTT acknowledges that there is more to the polymorphism than the length of 5-HTTLPR. A single-nucleotide polymorphism (SNP) known as rs25531 is now believed to interact with the length of 5-HTTLPR such that individuals with guanine (g) at the site, even if they have the long allele, are no more transcriptionally efficient on average than individuals with the short allele (Hu et al. 2001; Wendland et al. 2006; it may be that the Add Health data do not include the needed information on rs25531). We recategorized the "g"-long combination as short, resulting in 27 additional individuals being included as something other than L/L. With this "triallelic" genotype classification, the most accurate according to the latest research, 5-HTT genotype continues to exhibit no statistically significant relationship with self-reported political participation (see Model 5). When we ran this same model with voting frequency as the

TABLE 2. Effect Sizes for Each Model

	Model 1	Model 2	Model 3	Model 4	Model 5
5HTT	.080	.092	.018	.051	.056
5HTT* Attend	.097	.092	.055	.016	.023

dependent variable (not shown), the relationship even gave some indication of going in the opposite direction from that hypothesized by Fowler and Dawes, though the coefficient was not significant ($p < .13$).

The sensitivity of these findings to changes in the model specification may not be particularly surprising in light of the relatively small *N* and modest coefficients, but it does lead to concerns over effect size and substantive (as opposed to statistical) significance. Previous examinations of the association between 5-HTTLPR and political participation say very little about effect size, and addressing this omission is a necessary step toward fully evaluating the nature of the relationship. Table 2 reports the semi-partial correlations computed from our data, which provide a basic indication of effect sizes for both the main effect of 5-HTTLPR and the interaction effect with church attendance. Even in Models 1 and 2, where we have the strongest results, these effect sizes are quite small and are reduced even further as the model is altered by improving the measurement of key independent and dependent variables. Given these small effect sizes our analyses are statistically underpowered (a power analysis of our models indicates their power is, at best, .45). Thus our analysis is much less likely to commit the Type I errors cautioned by Charney and English than Type II errors. To put it another way, our analyses may be failing to reject a false null hypotheses on the impact of 5-HTTLPR because the effect size is simply too small to be reliably detected in a complex multivariate model. We suspect the same conclusion applies to the results generated by the Add Health data. In sum, though the presence of replication offers some convergence on the association between 5HTTLPR and political participation, the effect size of this relationship suggests that more analyses on larger samples are needed before firm conclusions can be drawn.

DISCUSSION

The Fowler-Dawes finding of a statistically significant relationship between self-reported political participation and the interaction of church attendance and allelic variation in 5-HTTLPR replicates with a completely different dataset that, although relatively small, is not beset by problems ascribed to the Add Health data. Moreover, because ours is a replication study, we had a particular relationship to test. It was the first association that we checked and not one of dozens or hundreds. Yet, though the findings replicate using a model specification very close to that of Fowler and Dawes, any deviations from (including improvements in) that model cause the relationship to decay. Notably, substituting voting frequency for self-reported political participation (Model 3) and using state-of-the-art categorization of the genotype (Model 5) result in the relationship disappearing entirely. This fragility is obviously a concern.

Where does all this leave understanding of a possible relationship between 5-HTTLPR genotypes and political participation? In some respects, it leaves it in the same place as the great majority of CGA studies: promising initial results followed by replications providing occasional hints at support but enough failures to raise serious concerns. Doubtless, the lack of stronger confirmation is attributable to many of the complexities that Charney and English delineate.

Ultimately we see problems on both sides of this debate. Fowler and Dawes are obviously aware of the many complications conspiring to make it difficult to be confident of the effects of allelic variation on complex phenotypes. Perhaps they should have refrained from emphasizing that the results are "clear" in showing that two genes are "significantly associated with voter turnout" (2008: 587–88). At the same time, Charney and English's implication that allelic associations across the board should not be attempted because of these complications goes too far. The number of published CGA studies is increasing geometrically, and the great majority of them are being done by geneticists and genetic psychologists who are well aware of the challenges (the number conducted solely by political scientists can be counted on one hand). Focusing on 5-HTTLPR alone, a Google search generates well over 1,000 hits. Not all constitute original studies, but a significant percentage reference such studies, and most are testing for possible connections with complex phenotypes. Of course, just because geneticists and others are conducting ever increasing numbers of CGA studies does not mean this is the optimal path or that the cautions of Charney and English should be ignored. To the contrary, we believe their warnings are valuable—as long as they are not paralyzing.

The solution to the complexity and inconsistency of CGA results is more and better research. Given the embryonic stage of this research in political science it would be unwise to heed calls to abandon the effort before it gets underway. Yet it is prudent to conduct and report the results of CGA studies with the utmost care. This is the direction in which the genetics community has been headed for quite some time. Indeed, the concerns expressed by Charney and English echo previously published official guidelines in the genetics community. After noting that "making sense of rapidly evolving evidence on genetic associations is crucial to making genuine advances in human genomics," the baseline recommendation of the relevant report (known as STREGA) is not that CGA studies be discontinued but rather that the quality of the research design and the "transparency of reporting" be enhanced (Little et al. 2009).

Noting that "the literature on candidate gene associations is full of reports that have not stood up to rigorous replication," the editorial policy of the journal *Behavior Genetics*, for example, is now that

"authors conduct a direct replication analysis prior to publication" (Hewitt 2012: 1). We are pleased to learn that the *APSR* has recently adopted a similar policy. In addition, authors hoping to publish in *Behavior Genetics* are encouraged to "pay particular attention to appropriate corrections to significance criteria for multiple testing" (Hewitt 2012: 1). It is clear that geneticists' response to the problems of CGA studies is to demand replication, multiple test corrections, transparency in reporting, and strong theory. With regard to the last criterion, unlike approaches such as genome-wide association studies (GWAS) that are discovery driven, CGA studies are theory driven so it is essential that there be strong a priori reasons for the specific expectations, particularly when an interaction is involved. On this point, note that including an environmental variable on the right-hand side of the equation, even as an interaction, raises potentially thorny causal direction issues (does church attendance affect political participation, or does political participation affect church attendance?).

Though demanding that CGA studies be theoretically based, replicated, and corrected for multiple hypothesis tests makes sense, advocating such policies raises the larger issue of the appropriateness of singling out one particular area of study for special treatment. CGA studies are not the only instances in which scholars cherry-pick those hypotheses that "work" from scores that were tested. In fact, we know that such practices are common among political scientists. What is the justification for applying Bonferroni corrections only to hypothesis testing that involves candidate genes? Similarly, the need for replication does not apply exclusively to CGA studies. A movement is afoot in psychology in which independent scholars are attempting to replicate a large number of previously published studies, and initial results suggest that failure to replicate is hardly confined to CGA studies (Carpenter 2012). Cherished political science findings may not be much more stable than those in psychology (Manzi 2012). For that matter, Charney and English's criticisms of phenotypic specification apply not just to Fowler and Dawes' study but also to all voting behavior studies that use survey-based variables, and such studies are not hard to find in political science. In short, many of the general concerns quite properly raised by Charney and English with regard to CGA studies could profitably be extended to much of the research conducted by political scientists, to the benefit of all.

REFERENCES

Bertolino, Alessandro, Giampiero Arciero, Valeria Rubino, Valeria Latorre, Mariapia De Candia, Viridiana Mazzola, Giuseppe Blasi, et al. 2005. "Variation of Human Amygdala Response during Threatening Stimuli as a Function of 5'HTTLPR Genotype and Personality Style." *Biological Psychiatry* 57 (12): 1517–25.

Buckholtz, Joshua W., and Andreas Meyer-Lindenberg. 2008. "MAOA and the Neurogenetic Architecture of Human Aggression." *Trends in Neurosciences* 31 (3): 120–29.

Carpenter, Siri. 2012. "Psychology's Bold Initiative." *Science* 335 (6076): 1558–61.

Caspi, Avshalom, Ahmad R. Hariri, Andrew Holmes, Rudolf Uher, and Terrie E. Moffitt. 2010. "Genetic Sensitivity to the Environment: The Case of the Serotonin Transporter Gene and Its Implications for Studying Complex Diseases and Traits." *American Journal of Psychiatry* 167 (5): 509–27.

Caspi, Avshalom, Joseph McClay, Terrie E. Moffitt, Jonathan Mill, Judy Martin, Ian W. Craig, Alan Taylor, and Richie Poulton. 2002. "Role of Genotype in the Cycle of Violence in Maltreated Children." *Science* 297 (5582): 851–54.

Caspi, Avshalom, Karen Sugdon, Terrie E. Moffitt, Alan Taylor, Ian W. Craig, HonaLee Harrington, Joseph McClay, et al. 2003. "Influence of Life Stress on Depression: Moderation by a Polymorphism in the 5-HTT Gene." *Science* 301 (5631): 386–89.

Charney, Evan, and William English. 2012. "Candidate Genes and Political Behavior." *American Political Science Review* 106 (1): 1–34.

Fowler, James H. 2006. "Altruism and Turnout." *Journal of Politics* 68 (3): 674–83.

Fowler, James H., and Christopher T. Dawes. 2008. "Two Genes Predict Voter Turnout." *Journal of Politics* 70 (3): 579–94.

Fowler, James H., and Cindy D. Kam. 2007. "Beyond the Self: Social Identity, Altruism, and Political Participation." *Journal of Politics* 69 (3): 813–27.

Goldstein, David B. 2009. "Common Genetic Variation and Human Traits." *New England Journal of Medicine* 360 (17): 1696–98.

Hariri, Ahmad R., Emily M. Drabant, and Daniel R. Weinberger. 2006. "Imaging Genetics: Perspectives from Studies of Genetically Driven Variation in Serotonin Function and Corticolimbic Affective Processing." *Biological Psychiatry* 59 (10): 888–97.

Hewitt, John K. 2012. "Editorial Policy on Candidate Gene Association and Candidate Gene-by-environment Interaction Studies of Complex Traits." *Behavior Genetics* 42 (1): 1–2.

Hu, Xianzhang, Gabor Oroszi, Jeffrey Chun, Tom L. Smith, David Goldman, and Marc A. Schuckit. 2005. "An Expanded Evaluation of the Relationship of Four Alleles to the Level of Response to Alcohol and the Alcoholism Risk." *Alcohol Clinical Experimental Research* 29 (1): 8–16.

Jablonka, Eva, and Gal Raz. 2009. "Transgenerational Epigenetic Inheritance: Prevalence, Mechanisms, and Implications for the Study of Heredity and Evolution." *Quarterly Review of Biology* 84 (2): 131–76.

Little, Julian, Julian P. T. Higgins, John P. A. Ioannidis, David Moher, France Gagnon, Erik von Elm, Muin J. Khoury, Barbara Cohen, George Davey-Smith, et al. 2009. "STrengthening the REporting of Genetic Association Studies (STREGA)—An Extension of the STROBE Statement." *European Journal of Clinical Investigation* 39 (4): 247–66.

Luco, Reini F., Qun Pan, Kaoru Tominaga, Benjamin J. Blencowe, Olivia M. Pereira-Smith, and Tom Misteli. 2010. "Regulation of Alternative Splicing by Histone Modification." *Science* 327 (5968): 996–1000.

Manzi, Jim. 2012. *Uncontrolled*. New York: Perseus.

Mefford, Heather C., Andrew J. Sharp, Carl Baker, Andy Itsara, Zhaoshi Jiang, et al. 2008. "Recurrent Rearrangements of Chromosome 1q21.1 and Variable Pediatric Phenotypes." *New England Journal of Medicine* 359 (16): 1685–99.

Meyer-Lindenberg, Andreas, Joshua W. Buckholtz, Bhaskar Kolachana, Ahmad R. Hariri, Lukas Pezawas, Giuseppe Blasi, Ashley Wabnitz, et al. 2006. "Neural Mechanisms of Genetic Risk for Impulsivity and Violence in Humans." *Proceedings of the National Academy of Sciences USA* 103 (16): 6269–74.

Poliseno, Laura, Leonardo Salmena, Jiangwen Zhang, Brett Carver, Willliam J. Haveman, and Pier P. Pandolfi. 2010. "A Coding-independent Function of Gene and Pseudogene mRNAs Regulates Tumor Biology." *Nature* 465 (7301): 1033–38.

Settle, Jaime E., Christopher T. Dawes, Nicholas A. Christakis, and James H. Fowler. 2010. "Friendships Moderate an Association between a Dopamine Gene Variant and Political Ideology." *Journal of Politics* 72 (4): 1189–98.

Wendland, J. R., B.J. Martin, M.R. Kruse, K-P Lesch, and D.L. Murphy. 2006. "Simultaneous Genotyping of Four functional Loci of Human SLC6A4, with a Reappraisal of 5-HTTLPR and rs25531." *Molecular Psychiatry* 11: 1–3.