

Schooling and variation in the *COMT* gene: the devil is in the details

Daniel Campbell,¹ Johanna Bick,¹ Carolyn M. Yrigollen,² Maria Lee,¹ Antony Joseph,³ Joseph T. Chang,¹ Elena L. Grigorenko,^{1,4,5} and Learning Disabilities Project: Zambia

¹Yale University, New Haven, CT, USA; ²University of California, Davis, CA, USA; ³University of California, Berkeley, CA, USA; ⁴Columbia University, New York, NY, USA; ⁵Moscow State University for Psychology and Education, Moscow, Russia

Background: Schooling is considered one of the major contributors to the development of intelligence within societies and individuals. Genetic variation might modulate the impact of schooling and explain, at least partially, the presence of individual differences in classrooms. **Method:** We studied a sample of 1,502 children (mean age = 11.7 years) from Zambia. Approximately 57% of these children were enrolled in school, and the rest were not. To quantify genetic variation, we investigated a number of common polymorphisms in the catechol-*O*-methyltransferase (*COMT*) gene that controls the production of the protein thought to account for >60% of the dopamine degradation in the prefrontal cortex. **Results:** Haplotype analyses generated results ranging from the presence to absence of significant interactions between a number of *COMT* haplotypes and indicators of schooling (i.e., in- vs. out-of-school and grade completed) in the prediction of nonverbal intelligence, depending on the parameter specification. However, an investigation of the distribution of corresponding *p*-values suggested that these positive results were false. **Conclusions:** Convincing evidence that the variation in the *COMT* gene is associated with individual differences in nonverbal intelligence either directly or through interactions with schooling was not found. *p*-values produced by the method of testing for haplotype effects employed here may be sensitive to parameter settings, invalid under default settings, and should be checked for validity through simulation. **Keywords:** Schooling, nonverbal intelligence, the *COMT* gene, haplotype analysis, haplo.glm, interaction effects.

Introduction

It seems obvious to state that schooling makes a difference in individual developmental and life outcomes, and that it strongly influences who will contribute to society at large and what kind of contributions will be made (Lutz & KC, 2011). In developed (i.e., high-income) countries (The World Bank, 2013a), it is generally taken for granted that formal schooling is closely linked to both individual (Bronfenbrenner, McClelland, Wethington, Moen, & Ceci, 1996) and social (Glaeser, Laibson, & Sacerdote, 2002) capital – that is, it contributes significantly to an individual's worth both to him- or herself and to society at large. Correspondingly, schooling is viewed both as the primary developmental task and as the main accomplishment of such countries' young members prior to entry into adulthood. These societies require schooling and stipulate the minimum number of years their youngsters are expected to spend in school; but these requirements and stipulations are not universal. In postindustrial economies, 96% of their school-aged members are engaged in regular or special education, whereas in emerging economies (less and least developed countries), these percentages are about 85% and 65%, respectively (UNICEF, 2009).

The role of formal schooling in developing (i.e., low- and middle-income) economies is much more tan-

genial to both social (Godoy et al., 2008) and individual (Serpell, 1993) capital; it is not uniformly viewed as either a requirement or as an accomplishment, so parents decide which of their children (if any) should or should not go to school (Brock & Levers, 2007). Primarily due to the uncertainty of the role of formal schooling in developing countries (Grigorenko, Hein, & Reich, in press; Grigorenko et al., 2001; Serpell & Jere-Folotiya, 2008), but also due to other characteristics of these societies (e.g., shortage of funds, societal conflict, and economic instability), many children in low- and middle-income countries are not enrolled in formal education. For example, in 2010, about 132 million children of primary and lower secondary school age were not being schooled formally (UNICEF, 2013); approximately half of these out-of-school children lived in sub-Saharan Africa (UNESCO, 2008).

Yet, the impact of schooling, regardless of the magnitude of its main effect, is not homogeneous; individual differences in classrooms among both students and teachers remain the object of investigation of large subfields in psychology and education. To explain the presence of these differences, chiefly, three different hypotheses have been investigated. First, it has been assumed that these differences are due to the impact of underlying genetic factors that influence either general cognitive (e.g., Plomin et al., 2004) or scholastic (e.g., Martin et al., 2011) performance, although recent large-scale reinvestigations of the previously published data suggest that, if they exist, these effects are of very small

The authors have declared that they have no competing or potential conflicts of interest.

magnitude and cannot, when conceived additively, explain the observed broad range of variability in individual differences in the classroom (Chabris et al., 2012; Rietveld et al., 2013; for intelligence and educational attainment, correspondingly). Second, it has been argued that classroom-based individual differences arise differentially, depending on the quality of instruction, so that genetic influences are pronounced more in higher and less in lower quality classrooms (Taylor, Roehrig, Soden Hensler, Connor, & Schatschneider, 2010). It has also been assumed that genetic factors shape the parameters of brain structure and function (Kwan, Sestan, & Anton, 2012), which, in turn, determine the parameters of basic information processing functions (Rowe et al., 2007); these, in turn, underlie the high (.45–.80) correlations between cognitive and scholastic performance (Luo, Thompson, & Detterman, 2003). It is this third hypothesis that seems to be particularly interesting for an investigation in societies where schooling is not universal, where the impact of schooling can be differential for individuals differentiated by characteristics of basic information processing functions and, in turn, by the parameters of their brain functioning, and, correspondingly, by the variation in the genome which defines these parameters.

To quantify the relevant genetic variation, we genotyped selected markers in the catechol-*O*-methyltransferase (*COMT*) gene (MIM 116790, 22q11). The enzyme produced by this gene is involved in the degradation of catecholamines (dopamine, epinephrine, and norepinephrine) and, through its participation in dopamine turnover, is related to neural functioning. *COMT* is said to account for >60% of the dopamine degradation in the prefrontal cortex (PFC); thus, it plays an important role in regulating dopamine levels in the PFC and, correspondingly, in psychological processes, both lower (e.g., reaction and inspection time) and higher (e.g., metacognitive functions) that engage the PFC (Karoum, Chrapusta, & Egan, 1994).

Two isoforms of *COMT* exist: the soluble isoform (S-*COMT*) and the membrane-bound isoform (MB-*COMT*). These are regulated by different promoters, are of different lengths (221 and 271 amino-acid residues, respectively), and have different functions. Structural DNA variants within this gene and in its near vicinity within noncoding regions have been associated with individual differences in a number of cognitive and affective processes, indicators of brain activity, and neuropsychiatric conditions (Dickinson & Elvevåg, 2009). The pleiotropic nature of the *COMT* gene action has been related to the complex dynamics of cognitive and affective functions (Mier, Kirsch, & Meyer-Lindenberg, 2010; Papaleo et al., 2008; Tunbridge, Harrison, & Weinberger, 2006).

The most studied polymorphic variant of the *COMT* gene is a single nucleotide change G/A (also known as the rs4680 single nucleotide polymorphism, SNP), resulting in an amino-acid substitution of valine (Val) with methionine (Met) at codon 108 for S-*COMT* and

codon 158 for MB-*COMT* (Val108/158Met) generating alternative forms of *COMT* with different functional properties (Lachman et al., 1996). Yet, the gene has other numerous polymorphisms that have been extensively studied in isolation, in combination with the rs4680 polymorphism, and in haplotype structures representative of the gene (Witte & Floel, 2012). Of note also is that the variation in the *COMT* gene has been previously associated with individual differences in IQ (Payton, 2009), academic attainment (Enoch, Waheed, Harris, Albaugh, & Goldman, 2009; Yeh, Chang, Hu, Yeh, & Lin, 2009), and numerous PFC-rooted processes (Lundwall, Guo, & Dannemiller, 2012; Stormer, Passow, Biesenack, & Li, 2012).

Here, we investigated the effect of schooling on levels of nonverbal intelligence in a sample of children from a developing country. Our specific hypothesis is that, assuming that schooling is the major causal factor in levels of intelligence around the world, genetic variation (specifically, genetic variation in the *COMT* gene) might modulate the impact of schooling and help explain, at least partially, the presence of individual differences in classrooms. To verify this hypothesis, we ascertained a sample of children with a graded amount of exposure to schooling and assessed them behaviorally and genotypically.

Methods

Study site

Zambia is a lower middle-income country in sub-Saharan Africa with a population of about 13.5 million (The World Bank, 2013b) people (99.5% Africans representing various local tribes), of whom approximately 45% are children younger than 14 years of age. In 2001, the gross school enrollment rate was estimated at 76.9% and the net enrollment rate at 65.1%, with 55.6% of the children at the intake year of age 7 not enrolled in schools (Riddell, 2003). English is the official language, but Zambia's people speak a number of major vernaculars (Bemba, Kaonde, Lozi, Lunda, Luvale, Nyanja, and Tonga) and about 70 other indigenous dialects. This study was carried out in a region of the Eastern Province of Zambia, where the dominant language is Chewa (also known as Nyanja) – a language of the Bantu language family. Although English is the primary language in schools for instruction, it is mandated that children learn to read, upon school entry, in one of the officially designated mother tongues; after 1 year of schooling, they are taught to read in English.

Participants

In Zambia, school enrollment is based on a geographic principle such that a given public school serves a local community (or a number of communities). Using a list of all of the public schools in the Chipata-Chadiza area, we selected, at random,

schools that were either rural or semi-rural/semiurban (i.e., located in or close to Chipata or Chadiza). Knowing the sizes of the schools, we included only those schools that had at least 10 students on their enrollment rosters in each of the grades we worked with (grades 2–6). This was necessary because of the high rate of absenteeism in Zambian government schools. In addition, we attempted to include both rural and semi-rural/semi-urban schools in such proportions that the resulting sample included children from these types of areas in approximately equal portions. As rural schools, especially those in remote areas, tend to be smaller in size, there were more rural schools. After the schools were selected, each school was approached for consent to participate, the procedures of the study were explained, and participant lists were generated. In each school, we attempted to recruit approximately 20% of its students in each grade. To reach this recruitment goal, the initial lists were created for twice as many children, 40% (i.e., to oversample at this stage). These children were selected at random (i.e., selecting a student by his or her number on the school roster), but the number of boys and girls was monitored to be approximately equal; the researchers compiling these lists were blind to the tribal affiliation, family constellation, health condition, or academic achievement of the selected children. When these lists were generated, exclusion criteria were applied based on (a) the presence of known physical or mental disabilities, and (b) the mother tongue being other than Nyanja (i.e., children primarily from the Chewa, Ngoni, Kunda, and other smaller tribes). We did not sample from grade 1 to avoid recruiting children who had not yet experienced a full year of schooling. Children were excluded if they had any known physical or mental disabilities (physical handicap, hearing and/or vision problems, and mental health problems). When the lists were revised, the testing date for a given school was scheduled; the duration of testing at a given school depended on the size of the school. Regardless, the testing of each individual child was completed within 1 day. If the list of participants included children who were absent for the duration of the testing period at a given school, they were not included in the data collection. We compiled lists of out-of-school children by working in parallel with the school and the community (or communities) served by the school. To match the selection procedure, we attempted to recruit children whose ages corresponded to those in the school sample (i.e., ~20% in each age band). Similarly, we monitored the number of boys and girls in the sample to ensure an equal gender distribution or to match the proportion of those absent. All testing was carried out on school grounds. The research team members who administered the tests were unaware of each child's school attendance status (in- or out-of-school). The resulting sample of in-school children included 862

children (49.0% boys, mean age = 11.8 years, $SD = 2.34$). We also worked with local communities and identified children who were not enrolled in school that year ($n = 640$, 53.6% boys; mean age = 11.7 years, $SD = 3.64$); of these children, 539 had had at least one full year of schooling and 101 had never been formally enrolled in school. The majority (932 or 62.1%) of children were from rural areas. Our total sample thus included 1,502 children (50.9% boys, mean age = 11.7 years, $SD = 2.97$).

Materials

To assess cognitive abilities within this sample of Zambian children, we used nonverbal measures of reasoning and memory. Specifically, we chose three subtests of the *Universal Nonverbal Intelligence Test, UNIT* (Bracken & McCallum, 1998). These assessments have been evaluated for use with Nyanja-speaking children (Stemler et al., 2009). The *UNIT* is an individually administered ability test that measures the general intelligence of children/youth aged 5–17 years. We used the Symbolic Memory (SyM), Cube Design (CD), and Spatial Memory (SpM) subtests of the *UNIT* to sample both memory (symbolic and nonsymbolic) and reasoning; a single indicator of nonverbal intelligence (NI) was then constructed. The distribution of NI values in the subgroups of our sample is shown in Figure 1.

To control for the socioeconomic status (SES) of the households of the children, a 17-item questionnaire was administered that gathered information regarding the characteristics of each family's living conditions (family size, structure, and living arrangements), levels of education and earnings of caregivers, sanitary/hygiene conditions (e.g., whether cooking was done inside or outside, whether there was running water, whether there was a latrine and of what kind), and the presence of entertainment (printed materials, radio, and TV) and luxury items (e.g., car). The data from the questionnaire were subjected to a principal-component analysis, where the first component explained 26.4% of the variance and the remaining components explained less than 11% each (i.e., the second principal component explained 10.3%, the third, 8.1%, and so forth). The standardized scores on the first principal component were used in subsequent analyses.

Genotyping

We sampled the *COMT* gene through five SNPs (rs737865, rs740603, rs165722, rs4680, and rs165599) attempting to capture variation using information on the amount of linkage disequilibrium (LD) in the *COMT* gene from the Yoruba population in HapMap (<http://hapmap.ncbi.nlm.nih.gov/>) and eight different African populations (Mukherjee et al., 2010). The SNPs were selected based on their reported associations with different cognitive and

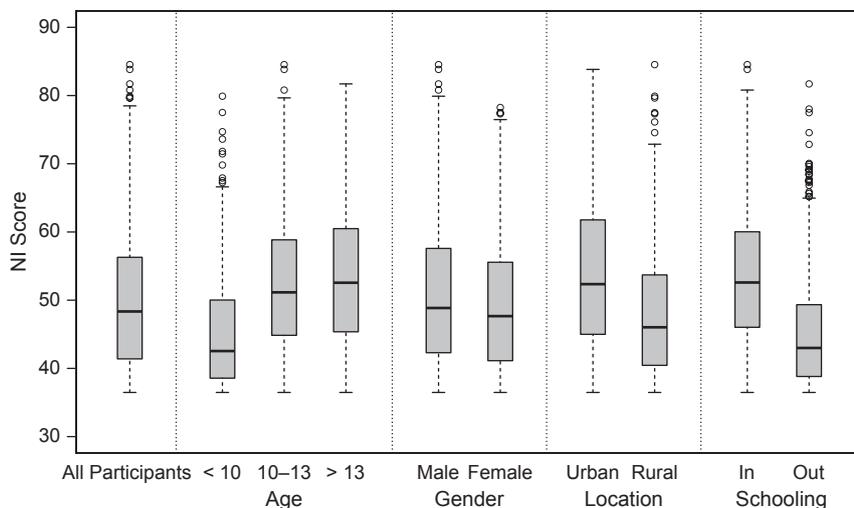


Figure 1 NI scores in the study samples, stratified by age, gender (Male or Female), location (Rural or Urban), and amount of schooling (In = currently enrolled; Out = currently not enrolled). NI scores are presented using T-scores (mean = 50 and *SD* = 10)

neuropsychiatric outcomes and derived allele frequencies in the studied African populations (<http://alfred.med.yale.edu/>). Supplemental online material presents corresponding allele/genotype (Table S1) and haplotype (Table S2) frequencies, and the amount of LD in the gene in this sample (Table S3).

Statistical analyses

For each of the five SNPs, a linear regression model to predict NI on the basis of demographic, schooling, and genetic (single marker) variables was fit using R software (R Development Core Team, 2011). The number of copies of the ancestral allele at each SNP was tabulated for each participant and included as genetic variables in the regression models; each model included the main effects of all variables and interaction effects for the SNP genotype and schooling status. Both single-marker and haplotype analyses were performed. To evaluate the predictive power of single markers, we used multiple regression analyses. To evaluate the predictive power of haplotypes, we used a haplotype regression method (Lake et al., 2003; Schaid, Rowland, Tines, Jacobson, & Poland, 2002), as implemented in the R haplo.stats library (http://mayoresearch.mayo.edu/mayo/research/schaid_lab/software.cfm). This method assigns weights to the various haplotype pairs that are consistent with an individual's genotypic data and allows for the recursive estimation of haplotype frequencies and effects. Inferences about haplotype effects are made through a weighted linear regression in the 'M' step of this EM-type algorithm. In these analyses, we investigated all possible combinations of two, three, four, and five marker haplotypes. Wald statistics were computed to test the significance of haplotype effects (both main and interaction), and (unadjusted) proportions of variance explained by all haplotype variables were calculated to measure the effect size of the genetic contribution in predicting NI.

One seemingly technical parameter setting in the haplo.stats library involves specifying a threshold frequency such that each haplotype with an estimated frequency above this threshold is included as a separate category for the haplotype, while haplotypes with estimated frequencies below the threshold are combined into a single category. The default setting for this threshold in the haplo.stats library is 1%. However, simulations on datasets in which no haplotype effect is present (obtained by permuting either the NI values, to simulate no relationship of NI to any variable, or by permuting the genotype values, to simulate no relationship between NI and genetic effects but maintaining the relationship with environment) suggest that such a small haplotype frequency threshold can yield very inflated rates of Type I error for haplotype effects, perhaps through overfitting to the data by inclusion of multiple low-frequency, low-signal haplotypes. As the haplotype frequency threshold increases, however, the number of haplotype effects included in the model shrinks, and potentially biologically interesting haplotypes are excluded from analysis. As a compromise between these competing considerations, we attempted to select the smallest haplotype frequency threshold that was consistent with a uniform distribution of *p*-values (tested by a Kolmogorov–Smirnov test) on datasets simulated under the null hypothesis of no haplotype effect. Haplotype frequency thresholds of 10% were sufficient to achieve uniformity for all combinations of SNPs when NI was permuted.

Results

We fit regression equations predicting the level of NI based on the child's demographic characteristics (age, sex, rural vs. urban, family SES) and schooling status (currently out of school and grade completed). These variables explained a substantial proportion of the variance in NI and its components (see

Table 1). As anticipated based on the data presented in Figure 1, there were strong and consistent effects of age (older children performed better than younger children), location (rural vs. urban, with rural children performing worse than urban children), SES (children from families with higher level of SES obtaining higher NI scores), and somewhat weaker and inconsistent effects of gender (girls tending to perform worse than boys). Predictably, there were also strong and consistent effects of schooling status so that being currently enrolled in school and being in a higher grade were predictive of higher NI scores. Although fluctuating when obtained from different regression equations, all these effects were significant and substantial – they explained approximately 38% of the variance in NI.

To investigate the degree of association between variation in the SNPs and the level of NI, regression analyses were carried out in which all of the main effects and interaction effects between particular genetic variants and schooling status were included. When genetic variants of the *COMT* gene were considered individually, the evidence supporting the association between *COMT* variation and levels of NI was inconsistent. There was no association between the main effects of any individual SNP and NI scores. However, when regressions predicting the subtest scores (SyM, CD, and SpM) were considered, we found the main effect for SNP rs165722 to be significant for predicting the SpM score. SNPs rs737865 and rs165722 showed significant interactions with schooling in predicting NI and one of the subtest scores each (SyM for rs737865 and SpM for rs165722). We interpreted this pattern of findings as suggestive of an association between genetic variation in the *COMT* gene and individual differences in NI, but this association could not be attributed to a particular SNP.

Table 1 Predicting levels of NI by children's age, gender, location, and schooling status

Variable	<i>B</i>	<i>SE</i>	β	<i>p</i> -value
Constant	39.15	2.46		<.001
Age	0.40	0.11	.12	<.001
Gender (Girls)	-1.02	0.42	-.05	.016
Type of residence (Rural)	-2.20	0.51	-.10	<.001
SES	1.35	0.13	.26	<.001
Schooling				
Out of school (Yes)	1.37	2.35	.01	.56
Grade completed	2.39	0.20	.38	<.001
$R^2 = .377, F(5,1415) = 142.7 (p < 2.2e-16)$				

With SyM, CD, and SpM as dependent variables, barring the Gender variable, all other variables exert similar influence. Gender does not have a significant effect when predicting SyM (*p*-value .63), and with CD and SpM, it has a similar negative effect as above (*p*-values .00071 and .061, respectively). R^2 for these were .204, .285 and .341, respectively.

Accordingly, we investigated the possibility that the variation in the *COMT* gene may be associated with the variation in NI via the gene's haplotypes. The level of LD in the coding region of the *COMT* gene was shown to be relatively low in Africa (Mukherjee et al., 2010). Pairwise LD was estimated by r^2 and D' (Table S3). Our findings concur with those in the literature, indicating, in general, low levels of LD in the *COMT* gene in this Nyanja-speaking sample from Zambia. Yet, three SNPs, rs165722, rs4680, and rs165599, demonstrated the presence of LD pairwise. Table 2 presents the results of the haplotype regression models with statistically significant results of the Wald test using the default haplotype frequency threshold of 1%, after correcting for multiple comparisons using a Bonferroni correction with 52 tests (26 sets of markers for both main and main + interaction effects). The proportion of the variance explained by *COMT* haplotypes averaged at around 6.6% and reached as high as 10.7%. However, *p*-values obtained for these marker combinations from simulated datasets in which NI was permuted were distinctly nonuniform and extremely biased toward zero, suggesting that the reported *p*-values were likely to indicate false positives.

When the haplotype frequency threshold was increased to 10%, which was sufficient to mitigate the nonuniformity of *p*-values when NI was permuted, the proportions of variance explained were reduced and the *p*-values were increased considerably. Two sets of SNPs still achieved, after an appropriate Bonferroni correction, similar levels of statistical significance using this setting: rs740603-rs165722-rs4680 and rs727865-rs740603-rs165722-rs165599. Graphical representations of the interaction between grade completed and the three most common haplotypes for rs737865-rs740603-rs165722-rs165599 are shown in Figure 2 (a,b). However, the interaction term *p*-values obtained by permutation for these two combinations of markers were still nonuniform (Kolmogorov–Smirnov tests *p*-values were $p = 2.96E-5$ and $p < 2.2E-16$, respectively) and biased toward zero when genotypes were permuted instead of NI. Permuting genotypes seems more appropriate than permuting NI, as permuting genotypes gives data simulated from the null hypothesis while preserving the relationship between NI and environmental variables. Using a haplotype frequency threshold of 15% yielded no statistically significant combinations of SNPs after the Bonferroni correction. Histograms of *p*-values from permutations of NI and genotype values are shown in Figure 3.

In brief, these results initially suggested that, after accounting for the presence and duration of schooling and such demographic variables as age, gender, location (urban vs. rural), and SES, a sizeable amount of variance in NI can be associated with genetic variation, in particular, variation in the *COMT* gene. The interactions between haplotypes

Table 2 Genetic association analyses for indicators of nonverbal intelligence (NI) and *COMT* haplotypes

Haplotype	Haplotype frequency threshold of 1% (default)			Haplotype frequency threshold of 10%			Haplotype frequency threshold of 15%		
	Nominal <i>p</i> -value		% Variance Explained by Haplotypes	Nominal <i>p</i> -value		% Variance Explained by Haplotypes	Nominal <i>p</i> -value		% Variance Explained by Haplotypes
	Main & Inter	Inter		Main & Inter	Inter		Main & Inter	Inter	
1,2,3	1.22E-07	.000118	5.94%	.029	.022	1.90%	.029	.022	1.90%
1,2,5	.000638	.00715	3.87%	.877	.892	0.20%	.693	.698	0.08%
1,3,5	<1E-15	.00085	3.32%	.009	.005	1.07%	.009	.005	1.07%
2,3,4	1.81E-08	6.25E-06	5.59%	.00057	.0036	3.06%	.596	.370	0.32%
2,3,5	2.41E-12	6.59E-07	7.05%	.004	.026	2.86%	.693	.698	1.30%
1,2,3,4	4.03E-12	7.05E-09	8.25%	.017	.005	1.84%	.692	.556	0.26%
1,2,3,5	<1E-15	6.46E-09	10.73%	1.05E-6	1.32E-4	4.19%	.613	.518	0.18%
1,3,4,5	8.43E-06	7.85E-05	3.05%	.129	.057	0.89%	.042	.250	0.56%
2,3,4,5	4.88E-15	3.08E-05	9.40%	.904	.827	0.16%	.641	.511	0.13%
1,2,3,4,5	9.81E-12	9.22E-05	9.06%	.761	.674	0.11%	.761	.674	0.11%

SNPs: 1 = rs737865, 2 = rs740603, 3 = rs165722, 4 = rs4680, 5 = rs165599. For each set of SNP loci and haplotype frequency threshold, we tested two null hypotheses: (a) haplotypes have no main effects and no interactive effects with either schooling variable (Out of School or Grade Completed) (Main & Inter), and (b) no interactions between haplotypes and either schooling variable (allowing for main effects of haplotypes) (Inter). For each combination of loci, we calculated a proportion of variability explained by the corresponding haplotypes by making use of two regressions, the first being a standard multiple regression using only the nongenetic covariates and the second being the weighted least squares regression done in the last (converged) step of the haplotype regression EM algorithm. The explained proportion of variability was found by dividing the difference between the variabilities of the fitted values of the two regressions by the total variability in the NI values. Marker combinations reported in this table have nominal *p*-values under $.05/2 \approx .00096$ using the default haplotype frequency threshold of 1%. Marker combinations shown in bold font also have nominal *p*-values under $.05/52 \approx .00096$ using a haplotype frequency threshold of 10%, which limits the degree of overfitting allowed by the model. When a haplotype frequency threshold of 15% is used, no marker combinations achieved statistical significance at the $.05/52 \approx .00096$ level.

and years of schooling seemed particularly noteworthy, as seen by the extremely small *p*-values obtained for the interaction effects. However, simulations in which the relationship between NI and haplotypes are removed *also* yield extremely small *p*-values, suggesting that the reported *p*-values may not reflect valid statistical significance of haplotype effects and instead may be a result of inflated Type I error. The risk of Type I error can be reduced somewhat by increasing the haplotype frequency threshold above the default setting, but doing so might preclude the possibility of detecting any effects of biologically meaningful haplotypes.

Discussion

Although schooling is one of the strongest environmental effects, few, if any, studies have considered whether and how the effect of schooling is differentiated for different genotypes or haplotypes. Such an inquiry appears promising as an attempt to understand why, under the powerful influence of schooling as a main effect, there is still a tremendous amount of individual differences among children in the same classroom. To our knowledge, this is one of the first inquiries to examine such differentiation. To perform such an inquiry, we utilized a large out-of-school cohort of children, in which a number of polymorphic sites of the *COMT* gene were genotyped. Three commentaries can be drawn from these analyses.

First, in accordance with the larger body of literature that investigates the effects of schooling on IQ in

both high- (Stelzl, Merz, Ehlers, & Remer, 1995) and low-income (Jukes & Grigorenko, 2010) countries, these data show a direct association between current school enrollment, duration of schooling, and level of NI. The natural experiment/observational study captured here, in which a substantial number of children were not, at least during the academic year of the study, enrolled in school, is particularly noteworthy. Such a situation is impossible to observe in high income countries. Of interest also is the observation that, on average, children who have had inconsistent schooling or no schooling demonstrate similar levels of NI. Furthermore, these children show substantially lower NI compared with children in school, notwithstanding the presence of high performers in both groups of out-of-school children. Perhaps this observation points to the importance of continuous schooling; a haphazard education, interrupted by periods of being out of school, appears to be less effective in supporting intellectual development than continuous enrollment.

Second, there appears to be no convincing statistically significant genetic effects on NI present in our sample, whether modeled in a straightforward linear regression or in a more sophisticated haplotype regression. Although the *COMT* gene seemed to be a plausible candidate for being one of the carriers of genetic effects, assumed to be relevant to understanding the presence of individual differences in response to intervention (Conti & Heckman, 2010), in this study, no statistically convincing evidence for the relevance of the variation in *COMT* to these

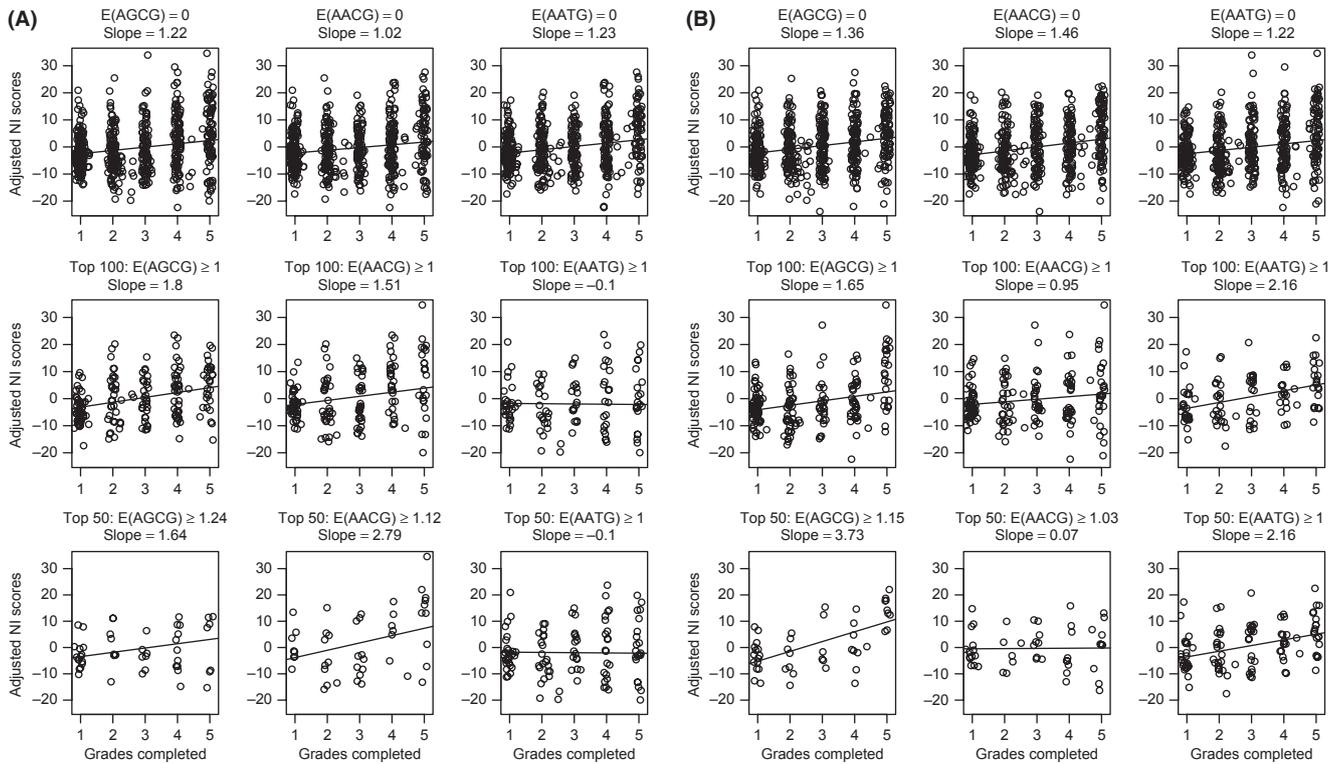


Figure 2 (a) The change in NI per year of schooling depending on the number of copies of the three most common *COMT* haplotypes for the set of four markers rs737865-rs740603-rs165722-rs165599. The adjusted NI scores are plotted against grades completed for children currently in school. In the first column of three plots, which pertains to the ancestral haplotype AGCG, the topmost plot shows only children having no copies of the AGCG haplotype, and the middle and bottom plots show children whose expected numbers of AGCG haplotypes are in the top 100 and the top 50, respectively, among all children. The quantity E(AGCG) is the expected number of AGCG haplotypes in an individual, defined as the weighted sum of possible haplotype counts (0, 1, or 2) times the corresponding probabilities fit by the model. The adjusted NI scores are obtained by extracting residuals from a regression of NI on the four variables age, sex, rural versus urban, and SES. The values for grades completed were jittered in the plots by adding small random perturbations to minimize overplotting and enhance the visibility of the points. The plots suggest that a component of NI is an interaction between the haplotypes and the number of grades completed, in which the rate of growth of NI scores with grades completed is increasing in the number of AACG haplotypes, decreasing in the number of AATG haplotypes, but roughly constant in the number of AGCG alleles (the ancestral allele). (b) Analogous interaction plots for the change in NI per year of schooling depending on the number of copies of the same three *COMT* haplotypes, for the case where genotype values have been permuted. A Wald test for the interaction effect of haplotypes and grade completed yielded a *p*-value of .065. The fact that changes in slope within columns are observable in the case where the interaction has been removed by permutation (as opposed to no change in slope with regard to haplotype dosage) suggests that the patterns observable in Figure 2(a) may be due to chance as well

differences was mounted. Of note is that, given the homogeneity of our sample with regard to the structure of the general population from which it was drawn, the distribution of social-economic factors, and the sameness of schooling in this geographic region, this sample was rather advantageous for detecting these effects. Thus, even though, given consistently high heritability estimates for all schooling-related phenotypes, there is a great deal of interest in understanding the etiology of these differences, anchoring them in specific genes might turn out to be a difficult task. It may be that the starting point in such analyses should consider not single, but multiple genes and their interactions (Zuk, Hechter, Sunyaev, & Lander, 2013).

Third, the use of a statistical software package for haplotype regression “out of the box” and employing default settings initially yielded statistical results that appeared to discover many haplotype effects with incredibly small *p*-values. However, further investigation revealed that the haplo.stats package, with its

default settings for haplotype frequency threshold, reported similarly small *p*-values on permuted data that eliminated the systematic relationships between NI and genetic variables, suggesting that the originally reported *p*-values did not, in fact, indicate a strong relationship. When nondefault settings were used to alleviate the downward bias of the reported *p*-values, the promisingly small *p*-values disappeared, leading us to conclude that the seemingly strong relationships between haplotypes and NI were not actually present. However, increasing the haplotype frequency threshold comes at a cost. Effects of haplotypes whose frequencies are below the specified threshold are averaged together by the haplotype regression model into a “rare” haplotype effect, precluding the model from capturing individual haplotype effects. As the threshold is increased, more haplotypes are lumped together into this rare category until only the most commonly occurring haplotypes remain. It may be, however, that a common haplotype will convey substantial risk or benefit; if it did, the

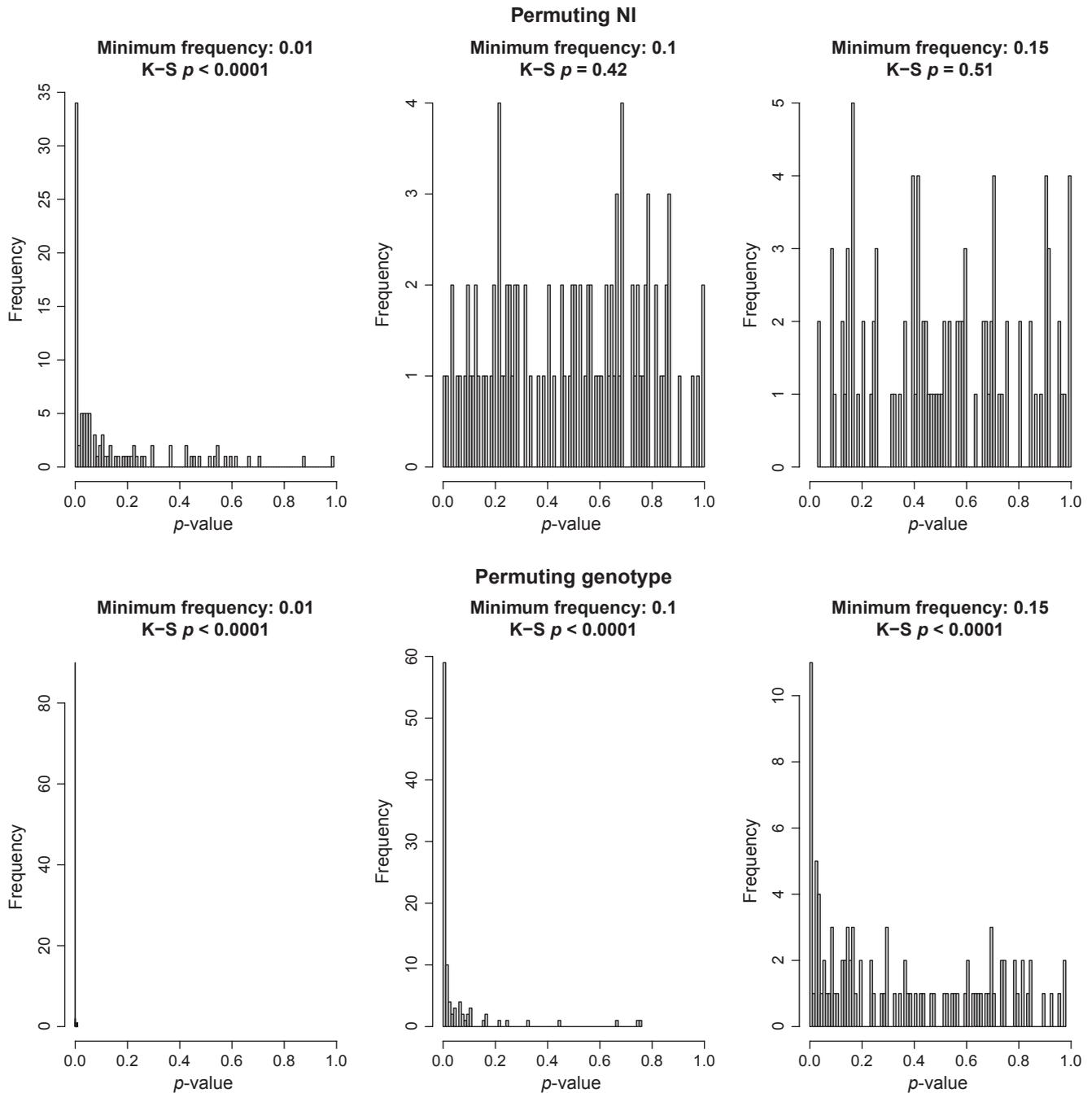


Figure 3 Histograms of simulated p -values for significance of the interaction terms in the haplotype regression model under the null hypothesis of no association between NI and haplotypes, obtained by permuting the NI values (top row) or permuting the genotype values (bottom row), for the four markers rs737865-rs740603-rs165722-rs165599. Settings for the haplotype frequency threshold allowed by the model are 1% (left), 10% (middle), and 15% (right). Also shown are p -values from a Kolmogorov–Smirnov test comparing each set of p -values to a Uniform distribution. Note that when NI is permuted, a haplotype frequency threshold of 10% is sufficient to yield uniformly distributed p -values under the null hypothesis of no association, but when genotype is permuted, even 15% is insufficient to yield uniformly distributed p -values

haplotype should have become either highly prevalent in the population or extremely rare through natural selection. If the rare haplotypes are the ones most likely to have large genetic effects, but are exactly the haplotypes we are unable to estimate separately using a high haplotype frequency threshold, then adopting a high threshold helps to preserve the statistical validity of the model (uniformly distributing p -values under the null hypothesis), but impairs its ability to detect important biological effects. The fact that the

reported p -values varied so widely in scale (by orders of magnitude, in fact) when the haplotype frequency threshold was altered speaks not of any specific limitation of the particular statistical package employed, but rather of a general observation that not every statistical model with given settings is appropriate to every dataset. For pairs of SNPs, for instance, haplotype regression did not report excessively small p -values even under the default settings. It was only when applying haplotype regression to sets

of three, four, and five SNPs, where many more rare haplotypes are likely and there is much greater flexibility for the model to overfit to the large number of haplotypes, that inflated Type I errors were observed. A thorough understanding of the statistical models and methods being applied is essential to drawing correct inferences about relationships in the data; without questioning the results obtained here under default settings and exploring them further, we would have drawn very different conclusions regarding the effect of *COMT* haplotypes on NI in this dataset.

Conclusions

Genetic variation in *COMT* did not appear to affect nonverbal intelligence directly or modulate the effect of schooling on nonverbal intelligence. This study contributes to a growing literature on the complexity of the mechanics of covariation in the genome (sampled here by a single gene) and the environment (sampled here by a single, but multifaceted, environmental factor – schooling). Studies of both single cohorts, as exemplified here, and, perhaps, multiple cohorts with differential exposures to formal education, should be carried out in the hopes of understanding the molecular bases of individual differences in achievement as outcomes of schooling. A special note of caution should be made with regard to the utilization of off-the-shelf software without careful scrutiny of the results for possible biases and departures from statistical validity.

Supporting information

Additional Supporting Information may be found in the online version of this article:

Table S1. Sample Allele and Genotype Frequencies.

Table S2. Sample Haplotype Frequencies.

Table S3. Pairwise Levels of LD (r^2 and D).

Acknowledgements

The preparation of this article was supported by funds from the US National Institutes of Health, NIH [awards TW006764 and DC007665 (PI: Grigorenko), HD048830 (PI: Pugh), HD052120 (PI: Wagner), and T32MH18268 (PI: Leckman, T32 fellows Campbell and Bick)], and Autism Speaks #7614 (PI: Campbell). Grantees undertaking such projects are encouraged to freely express their professional judgment. This article, therefore, does not necessarily represent the position or policies of the NIH or of Autism Speaks, and no official endorsement should be inferred. We are grateful to Ms. Mei Tan for her editorial assistance. **Learning Disabilities Project:** Zambia University of Zambia, Zambia: Florence Chamvu, Jacqueline Jere-Folotiya, Bestern Kaani, Kalima Kalima, Sophie Kasonde-N'gandu, Aidan Mambwe, Robert Serpell; Macha Research Trust / Malaria Institute at Macha, Zambia: Philip E. Thuma; Emmanuel College, USA: Jonna Kwiatkowski; Stanford University, USA: Hilary Chart; Tufts University, USA: Linda Jarvin; Yale University, USA: Lesley Hart, Tina Newman, Sara Sparrow (in memoriam); Wesleyan University, USA: Steven E. Stemler.

Correspondence

Elena L. Grigorenko, Child Study Center, Department of Psychology, Department of Epidemiology & Public Health, Yale University, 230 South Frontage Road, New Haven, CT 06519-1124, USA;

Email: Elena.grigorenko@yale.edu

Key points

- Individual differences in nonverbal intelligence are associated with demographic variables and the effect of schooling.
- Children from rural areas and of lower SES performed worse than their urban, higher SES peers.
- Children who are out of school even only for 1 year, as a group, perform worse than the children who are enrolled in school.
- Genetic variation in the *COMT* gene does not appear to modulate the effect of schooling on nonverbal intelligence.

References

- Bracken, B.A., & McCallum, R.S. (1998). *The Universal Nonverbal Intelligence Test (UNIT)*. Itasca, IL: Riverside.
- Brock, C., & Levers, L.Z. (Eds.) (2007). *Aspects of education in the Middle East and North Africa*. Oxford, UK: Symposium Books.
- Bronfenbrenner, U., McClelland, P., Wethington, E., Moen, P., & Ceci, S.J. (1996). *The state of Americans*. New York: Free Press.
- Chabris, C.F., Hebert, B.M., Benjamin, D.J., Beauchamp, J., Cesarini, D., van der Loos, M.,... & Laibson, D. (2012). Most reported genetic associations with general intelligence are probably false positives. *Psychological Science*, 23, 1314–1323.
- Conti, G., & Heckman, J.J. (2010). Understanding the early origins of the education–health gradient: A framework that can also be applied to analyze gene–environment interactions. *Perspectives on Psychological Science*, 5, 585–605.
- Dickinson, D., & Ellevåg, B. (2009). Genes, cognition and brain through a *COMT* lens. *Neuroscience*, 164, 72–87.
- Enoch, M., Waheed, J., Harris, C., Albaugh, B., & Goldman, D. (2009). *COMT* Val158Met and cognition: Main effects and interaction with educational attainment. *Genes, Brain and Behavior*, 8, 36–42.
- Glaeser, E.L., Laibson, D., & Sacerdote, B. (2002). An economic approach to social capital. *The Economic Journal*, 112, F437–F458.

- Godoy, R., Seyfried, C., Reyes-Garcia, V., Huanca, T., Leonard, W.R., McDade, T., ... & Vadez, V. (2008). Schooling's contribution to social capital: Study from a native Amazonian society in Bolivia. In J. Elliott & E.L. Grigorenko (Eds.), *Western psychological and educational theory in diverse contexts* (pp. 137–163). New York: Routledge/Taylor & Francis Group.
- Grigorenko, E.L., Geissler, P.W., Prince, R., Okatcha, F., Nokes, C., Kenny, D.A., ... & Sternberg, R.J. (2001). The organization of Luo conceptions of intelligence: A study of implicit theories in a Kenyan village. *International Journal of Behavior Development, 25*, 367–378.
- Grigorenko, E.L., Hein, S., & Reich, J. (in press). Cultural manifestation of intelligence in formal and informal learning environments. In L. Jensen (Ed.), *The Oxford handbook of human development and culture: An interdisciplinary perspective*. New York: Oxford University Press.
- Jukes, M.C.H., & Grigorenko, E.L. (2010). Assessment of cognitive abilities in multiethnic countries: The case of the Wolof and Mandinka in the Gambia. *British Journal of Educational Psychology, 80*, 77–97.
- Karoum, F., Chrapusta, S.J., & Egan, M.F. (1994). 3-Methoxytyramine is the major metabolite of released dopamine in the rat frontal cortex: Reassessment of the effects of antipsychotics on the dynamics of dopamine release and metabolism in the frontal cortex, nucleus accumbens, and striatum by a simple two pool model. *Journal of Neurochemistry, 63*, 972–979.
- Kwan, K.Y., Sestan, N., & Anton, E.S. (2012). Transcriptional co-regulation of neuronal migration and laminar identity in the neocortex. *Development, 139*, 1535–1546.
- Lachman, H.M., Papolos, D.F., Saito, T., Yu, Y.M., Szumlanski, C.L., & Weinshilboum, R.M. (1996). Human catechol-O-methyltransferase pharmacogenetics: Description of a functional polymorphism and its potential application to neuropsychiatric disorders. *Pharmacogenetics, 6*, 243–250.
- Lake, S.L., Lyon, H., Tantisira, K., Silverman, E.K., Weiss, S.T., Laird, N.M., & Schaid, D.J. (2003). Estimation and tests of haplotype-environment interaction when linkage phase is ambiguous. *Human Heredity, 55*, 56–65.
- Lundwall, R.A., Guo, D.C., & Dannemiller, J.L. (2012). Exogenous visual orienting is associated with specific neurotransmitter genetic markers: A population-based genetic association study. *PLoS One, 7*, e30731.
- Luo, D., Thompson, L.A., & Detterman, D.K. (2003). The causal factor underlying the correlation between psychometric g and scholastic performance. *Intelligence, 31*, 67–83.
- Lutz, W., & KC, S. (2011). Global human capital: Integrating education and population. *Science, 333*, 587–592.
- Martin, N.W., Medland, S.E., Verweij, K.J., Lee, S.H., Nyholt, D.R., Madden, P.A., ... & Martin, N.G. (2011). Educational attainment: A genome wide association study in 9538 Australians. *PLoS One, 6*, e20128.
- Mier, D., Kirsch, P., & Meyer-Lindenberg, A. (2010). Neural substrates of pleiotropic action of genetic variation in COMT: A meta-analysis. *Molecular Psychiatry, 15*, 918–927.
- Mukherjee, N., Kidd, K.K., Pakstis, A.J., Speed, W.C., Li, H., Tarnok, Z., ... & Kidd, J.R. (2010). The complex global pattern of genetic variation and linkage disequilibrium at catechol-O-methyltransferase. *Molecular Psychiatry, 15*, 216–225.
- Papaleo, F., Crawley, J.N., Song, J., Lipska, B.K., Pickel, J., Weinberger, D.R., & Chen, J. (2008). Genetic dissection of the role of catechol-O-methyltransferase in cognition and stress reactivity in mice. *Journal of Neuroscience, 28*, 8709–8723.
- Payton, A. (2009). The impact of genetic research on our understanding of normal cognitive ageing: 1995 to 2009. *Neuropsychology Review, 19*, 451–477.
- Plomin, R., Turic, D., Hill, L., Stephens, M., Williams, J., Owen, M., & O'Donovan, M. (2004). A functional polymorphism in the succinate-semialdehyde dehydrogenase (aldehyde dehydrogenase 5 family, member A1) gene is associated with cognitive ability. *Molecular Psychiatry, 9*, 582–586.
- R Development Core Team (2011). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Riddell, A. (2003). *The introduction of free primary education in sub-Saharan Africa*. France: UNESCO Paris.
- Rietveld, C. A., Medland, S.E., Derringer, J., Yang, J., Esko, T., Martin, N. W., ... & Koellinger, P.D. (2013). A GWAS of 126,559 individuals identifies genetic variants associated with educational attainment. *Science, 340*, 1467–1471.
- Rowe, D.L., Cooper, N.J., Liddell, B.J., Clark, C.R., Gordon, E., & Williams, L.M. (2007). Brain structure and function correlates of general and social cognition. *Journal of Integrative Neuroscience, 6*, 35–74.
- Schaid, D.J., Rowland, C.M., Tines, D.E., Jacobson, R.M., & Poland, G.A. (2002). Score tests for association between traits and haplotypes when linkage phase is ambiguous. *American Journal of Human Genetics, 70*, 425–434.
- Serpell, R. (1993). *The significance of schooling: Life-journeys in an African society*. New York: Cambridge University Press.
- Serpell, R., & Jere-Folotiya, J. (2008). Developmental assessment, cultural context, gender, and schooling in Zambia. *International Journal of Psychology, 43*, 1–2.
- Stelzl, I., Merz, F., Ehlers, T., & Remer, H. (1995). The effect of schooling on the development of fluid and crystallized intelligence: A quasi-experimental study. *Intelligence, 21*, 279–296.
- Stemler, S.E., Chamvu, F., Chart, H., Jarvin, L., Jere, J., Hart, L., ... & Grigorenko, E.L. (2009). Assessing competencies in reading and mathematics in Zambian children. In E.L. Grigorenko (Ed.), *Multicultural psychoeducational assessment* (pp. 157–185). New York: Springer.
- Stormer, V.S., Passow, S., Biesenack, J., & Li, S.C. (2012). Dopaminergic and cholinergic modulations of visual-spatial attention and working memory: Insights from molecular genetic research and implications for adult cognitive development. *Developmental Psychology, 48*, 875–889.
- Taylor, J., Roehrig, A.D., Soden Hensler, B., Connor, C.M., & Schatschneider, C. (2010). Teacher quality moderates the genetic effects on early reading. *Science, 328*, 512–514.
- The World Bank. (2013a). How we classify countries. Available from: <http://data.worldbank.org/about/country-classifications> [last accessed 30 April 2013].
- The World Bank. (2013b). Zambia. Available from: <http://data.worldbank.org/country/zambia> [last accessed 30 April 2013].
- Tunbridge, E.M., Harrison, P.J., & Weinberger, D.R. (2006). Catechol-o-methyltransferase, cognition, and psychosis: Val158Met and beyond. *Biological Psychiatry, 60*, 141–151.
- UNESCO (2008). *Overcoming inequality: Why governance matters*. Paris, France: The United Nations Educational, Scientific and Cultural Organization.
- UNICEF. (2009). Children out of school. Available from: http://www.childinfo.org/education_outofschool.php [last accessed 30 April 2013].
- UNICEF. (2013). Out-of-school children initiative. Basic education and gender equality. Available from: http://www.unicef.org/education/bege_61659.html [last accessed 30 April 2013].
- Witte, A.V., & Floel, A. (2012). Effects of COMT polymorphisms on brain function and behavior in health and disease. *Brain Research Bulletin, 88*, 418–428.
- Yeh, T.K., Chang, C.Y., Hu, C.Y., Yeh, T.C., & Lin, M.Y. (2009). Association of catechol-O-methyltransferase (COMT) polymorphism and academic achievement in a Chinese cohort. *Brain & Cognition, 71*, 300–305.
- Zuk, O., Hechter, E., Sunyaev, S.R., & Lander, E.S. (2013). The mystery of missing heritability: Genetic interactions create phantom heritability. *Proceedings of the National Academy of Sciences, 109*, 1193–1198.

Accepted for publication: 22 May 2013

Published online: 19 August 2013