RESEARCH ARTICLE

What Caused over a Century of Decline in General Intelligence? Testing Predictions from the Genetic Selection and Neurotoxin Hypotheses

Michael A. Woodley of Menie [1,2] · Matthew A. Sarraf [3] · Mateo Peñaherrera-Aguirre [4] · Heitor B. F. Fernandes [4] · David Becker [5]

Abstract

Several converging lines of evidence indicate that general intelligence (g) has declined in Western populations. The causes of these declines are debated. Here, two hypotheses are tested: (1) selection acting against genetic variants that promote g causes the decline and (2) the presence of neurotoxic pollution in the environment causes the decline. A linear mixed model was devised to test (1) and (2), in which the secular decline in a "heritable g" ($g.h$) chronometric factor (comprised of convergent indicators of simple reaction time, working memory, utilization frequencies of high difficulty and also social-intelligence-indicating vocabulary items and per capita macro-innovation rates) was predicted using a neurotoxin chronometric factor (comprised of convergent secular trends among measures of lead, mercury and dioxin + furan pollution, in addition to alcohol consumption) and a polygenic score chronometric factor (comprised of polygenic score means for genetic variants predictive of g, sourced from US and Icelandic age-stratified cohorts). Bivariate correlations revealed that (other than time) only the polygenic score factor was significantly associated with declining $g.h$ ($r = .393$, $p < .05$ vs. .033, ns for the neurotoxin factor). Using a hierarchical linear mixed model approach incorporating 25 year lags between the predictors and $g.h$, time period, operationalized categorically as fifths of a century, accounted for the majority of the variance in the decline in $g.h$ (partial $\eta^2 = .584$, $p < .05$). Net of time period and neurotoxins, changing levels of polygenic scores also significantly predicted variance in the decline in $g.h$ (partial $\eta^2 = .253$, $p < .05$); however, changing levels of neurotoxins did not significantly predict variance in $g.h$ net of time (partial $\eta^2 = .027$ ns). Within-period analysis indicates that the independent significant positive effect of the polygenic score factor on $g.h$ was restricted to the third fifth of a century period ($\beta = .202$, $p < .05$).

Introduction

Sir Francis Galton (1869) made the earliest scientific prediction concerning temporal trends in intelligence, a trait that he thought should be declining across generations owing to the

Matthew A. Sarraf and Mateo Peñaherrera-Aguirre: Joint second authors

✉ Michael A. Woodley of Menie
   Michael.Woodley@vub.ac.be

1 Center Leo Apostel for Interdisciplinary Studies, Vrije Universiteit Brussel, Brussels, Belgium

2 Unz Foundation, Palo Alto, CA, USA

3 University of Rochester, Rochester, NY, USA

4 Department of Psychology, University of Arizona, Tucson, AZ, USA

5 Department of Psychology, Technische Universität Chemnitz, Chemnitz, Germany

relative fertility advantage of those with lower compared to those with higher socioeconomic status (which Galton used as a proxy for intelligence, given the absence of direct measures of the phenotype). Negative associations between measured intelligence (IQ) and fertility were subsequently observed in the 1920s and 1930s, leading psychologists to predict that intelligence, being a heritable trait, would decline by around one to three points per generation due to the action of this "dysgenic" selection (Cattell 1937). Cross-sectional studies conducted in later decades, in which the IQ scores of age-matched cohorts were compared, revealed that IQ, rather than declining, was in fact increasing substantially (Cattell 1950). This anomaly became known as *Cattell's paradox* (since Raymond B. Cattell was the first to draw significant attention to it; Higgins et al. 1962)—the apparent incongruity between observations of (concomitant) rising IQ and "dysgenic" selection. This rapid increase in population-level IQ was subsequently termed the *Flynn effect* (Herrnstein and Murray 1994), after James Flynn, who

demonstrated that these gains are ubiquitous across modernized and modernizing countries and are found on many cognitive ability measures (Flynn 1987; Pietschnig and Voracek 2015). Recent molecular research has confirmed that selection is acting directly against genetic variants associated with cognitive ability and related phenotypes, such as educational attainment, in certain populations (Beauchamp 2016; Conley et al. 2016; Kong et al. 2017; Woodley of Menie et al. 2016). Of particular note is a large study of the population of Iceland, which found evidence that an educational attainment polygenic score (a normally distributed genetic index comprised of multiple single nucleotide polymorphisms which collectively significantly predict phenotypic variance in a trait of interest) has declined over time across cohorts, at a rate consistent with the strength of selection acting against variants captured by this score in this population (Kong et al. 2017). The observation that the frequencies of genes for cognitive ability are declining in Western populations, despite the Flynn effect, adds further significance to Cattell's paradox.

The Co-occurrence Model

A recently proposed solution to Cattell's paradox is the *co-occurrence model* (Woodley and Figueredo 2013), which is based on the observation that genetic selection and environmental influences act on different variance components of intelligence. Far from being a homogeneous phenotype, intelligence is comprised of both general mental ability (*g*) and specific performance components (*s*), i.e., those variances that are specific to certain abilities and tests (Spearman 1904). Utilizing the Method of Correlated Vectors (MCV), a simple statistic that correlates the *g* loadings of a given set of subtests with an associated effect size (such as the magnitude of inbreeding depression effects or the strength of the correlation between test performance and reaction speed) in order to determine the degree to which the *g* loading moderates that effect size, it has been found that the strength of selection acting against intelligence correlates positively with subtest *g* loading ($\rho = .87$, $p < .05$, 95% CI = .87 to .87, $N = 108,040$; Woodley of Menie et al. 2017a), whereas the magnitude of the Flynn effect correlates negatively with *g* loadings ($\rho = -.38$, $p < .05$, 95% CI = $-.39$ to $-.38$, $N = 16,663$; te Nijenhuis and van der Flier 2013). Positive vector correlations are termed *Jensen effects* (after Arthur Jensen, who developed MCV; Rushton 1998), with negative vector correlations correspondingly termed *anti-Jensen effects*. Jensen effects are typically associated with biological and genetic variables, such as subtest heritabilities (Kan et al. 2013; Voronin et al. 2015, Table 3, p. 3), effects of inbreeding depression, and physiological correlates of intelligence, such as nerve conduction velocity and measures of neural metabolic efficiency (Jensen 1998). Anti-Jensen effects, by contrast, are typically associated with influences on intelligence that arise from the environment and culture, such as IQ points gained by lower-ability adoptees from rearing in higher-ability households (te Nijenhuis et al. 2015),

educational interventions (te Nijenhuis et al. 2014), and test practice effects (Lievens et al. 2007; Reeve and Lam 2007; te Nijenhuis et al. 2007). This indicates that selection acts on the highly heritable *g* factor, whereas environmental influences (such as increasing exposure to education and increasing familiarity with cognitive testing) act on the least heritable and most specialized ability variances, thereby producing the Flynn effect. Hence, it is predicted that gains in *s* can co-occur with losses in *g* and that such gains and losses have been co-occurring in at least some modern populations.

The co-occurrence model led to the specific prediction that not all cognitive ability measures will exhibit positive temporal trends. A subset should in fact be tracking the long-term decline in *g* due (in part) to selection. Evidence has been found for this temporal pattern using simple single-parameter measures of endophenotypes considered basic to *g*, such as simple visual and auditory reaction times (Madison et al. 2016; Silverman 2010; Woodley et al. 2013, Woodley et al. 2014; Woodley of Menie et al. 2015b) and color discrimination ability (Woodley of Menie and Fernandes 2016a)—which all track potential declines in *neural efficiency*—and also declining 3D rotational ability (Pietschnig and Gittler 2015), backwards digit span performance (Wongupparaj et al. 2017; Woodley of Menie and Fernandes 2015), and Corsi Blocks performance (Wongupparaj et al. 2017), which all track declines in *working memory*. Single-parameter ratio-scale measures such as these have the potential advantage of measurement invariance with respect to facets of *g* over time (an indicator exhibits measurement invariance if, say, the parameter that it measures at time point A is the same parameter that it measures at time point B). Conventional pencil-and-paper IQ tests by contrast are typically not measurement invariant across cohorts (Fox and Mitchum 2013; Wicherts et al. 2004), consistent with the Flynn effect being associated with increasing ability specialization over time (Woodley and Madison 2013). Other indicators that potentially track declining *g* are those that are relatively insensitive to environmental influences by virtue of being highly heritable (high heritability by definition implies low environmentality, since each variance component is expressed as a percentage of the overall variance explained; Sesardic 2005). Vocabulary exhibits both the highest heritability and the highest *g* loading of any intelligence measure (Kan et al. 2013). The utilization frequencies of the four highest difficulty (thus likely hardest to learn and use appropriately) vocabulary target words from the WORDSUM short-form test appear to have declined across 155 years of written text in Google Ngram viewer (Woodley of Menie et al. 2015a). Consistent with expectations, this frequency decline was predicted by the two-way interaction between the strength of the negative correlation between fertility and item pass rate (capturing selection strength) and item response theory difficulties (both computed using data from the General Social Survey), net of temporal autocorrelation, word age, and population written literacy rates.

Similar declines have also been noted in the utilization frequencies across over a century of texts of words connoting high levels of altruism, especially altruism directed towards the group to which the originator of the text belongs (Woodley of Menie et al. 2017a). Altruism may function as a costly signal of *g* (Millet and Dewitte, 2007) and may be an important locus of selection favoring higher population levels of *g* under conditions of inter-group competition; thus, its decline may signal reduced levels of the more *g*-loaded aspects of social intelligence (Woodley of Menie et al. 2017a). Potentially consistent with this, a cross-temporal meta-analysis of performance on a related measure (ability-based emotional intelligence) found statistically significant indications of decline with respect to one performance domain (perceiving emotions) across 14 years (2001-2015) (Pietschnig and Gittler, 2017).

Population-level measures of complex problem-solving ability constitute another class of indicator that seems to track declines in *g*. Such measures include per capita rates of macro-innovation (i.e., conspicuously novel and disruptive innovation, as convergently rated by historians of science and technology; Galton 1869; Huebner 2005; Murray 2003). Since the mid-nineteenth century, it has been found that per capita rates of macro-innovation, and also the rates of eminent individuals responsible for them, have declined (Huebner 2005; Murray 2003), which is consistent with selection against *g* having reduced the numbers of ultra-high ability individuals capable of solving complex scientific and engineering problems (Woodley 2012; Woodley and Figueredo 2013). Collectively, these indicators of long-term declines in *g* have been termed *Woodley effects* (Sarraf 2017[1]).

Recently, it was found that a chronometric factor (i.e., a latent variable comprised of various temporal trends) capturing the decline in "heritable" general intelligence (*g.h*) and comprised of five measures potentially indicative of declining *g* exhibited both high internal consistency among its component indicators and discriminant validity relative to two other chronometric factors, tracking somatic modifications (such as changes in height, cranial fluctuating asymmetry, BMI) and indications of cognitive ability specialization (such as increasing concretization in language use, improvements in short-term memory, increased utilization frequencies of easy-to-learn words) respectively (Woodley of Menie et al. 2017a). It was also found that the vector of the loadings of the *g.h* factor on each of the three psychometric ability indicators used in the construction of the factor (simple visual reaction times, backwards digit span, and high-difficulty vocabulary items) correlated significantly with the vector of the strength of selection acting against each indicator ($r = .97$, $p < .05$, 95% CI = .97 to .97, $N = 15,576$). This is precisely in line with the expectation that the *g.h* chronometric factor is

tracking a phenotypic decline in *g* due to the action of selection, as the psychometric variables that co-vary in time with *g.h* to the greatest degree are the ones under strongest selection.

Neurotoxins

Various researchers have proposed a role for intelligence-depressing neurotoxins in driving secular trends in cognitive measures. Nevin (2000) suggests that declining levels of lead exposure during parts of the twentieth century in the US may have been responsible for both the Flynn effect and changes in the life history characteristics of this population, such as decreasing rates of violence and out-of-wedlock pregnancy. Silverman (2010) claims that increased exposure to lead, chlordane, trichloroethylene and mercury, as byproducts of industrial growth in the West since the nineteenth century, may have slowed simple reaction times, as exposure to these toxins is known to reduce processing speed. ten Tusscher et al. (2014) also contend that dioxin exposure may be associated with reaction speed decline. More recently, Clarke (2015) has proposed the *antiinnatia theory*, which holds that mercury exposure has increased the prevalence of autism, via interference with normal patterns of gene expression, and has promoted the Flynn effect as a byproduct. Another proponent of neurotoxin-based explanations of intelligence change is Barbara Demeneix, whose 2014 book *Losing Our Minds* posits that thyroid hormone disruption provides a bridge linking the effects of neurotoxins with declining cognitive functioning, via such disruption's effects on patterns of gene expression in development. In addition to declining *g* (p.6), Demeneix also posits a larger nomological network of behavioral abnormalities, such as autistic spectrum and attention deficit/hyperactivity disorders, increasing frequencies of which may also stem from the action of endocrine-disrupting neurotoxins (Demeneix, 2014, 2017). Demeneix (2014) discusses the effects of "ever-present neurotoxins" (p.12), such as lead and alcohol, on *g*, in addition to biphenols, mercury, and dioxins along with other more recently manufactured potentially endocrine-disrupting chemical pollutants.

Unlike the co-occurrence model, where predictions are explicit (specifically that indicators of *g.h* should decline due to genetic change, whereas specialized abilities should increase, tracking environmental improvements), it is not at all clear what neurotoxin models are predicting. Different researchers seem to be predicting different things. Nevin (2000) suggests that lead levels have generally declined in industrialized nations during the twentieth century and that this phenomenon is promoting the Flynn effect, whereas Demeneix (2017) argues that *g* might be increasing or decreasing over time—if the former is occurring, the increase is *lower than it would otherwise be* due to the presence of various neurotoxins, but if the latter is occurring, neurotoxins explain at least some significant portion of the decline, and, in

---

[1] The term appears to have been coined in 2015 by Charles Murray. Sarraf (2017, p. 239), however, first defined Woodley effects as "secular trends that plausibly result, in part or in whole, from population-level degradation of the integrity of genetic factors that underlie *g*."

either case, even if lead levels have fallen (pp. 75–91)[2]. Conversely, Silverman (2010) suggests that lead and other industrial byproducts might be increasing over time and that this might be slowing simple reaction speed. Clarke (2015) suggests that increasing exposure to mercury actually boosts intelligence, whereas Demeneix (2014) maintains the opposite. Very little effort has been made to test any of the predictions of variants of this model using secular trend data. Proponents of the neurotoxin model have typically relied upon either "visual correlation" (i.e., juxtaposing different graphs; e.g., Clarke, 2015) or simple assertions that a given neurotoxin could account for a given change in a cognitive indicator, without subjecting the claim to temporal analysis (e.g., Demeneix 2014, 2017; Nevin 2000; Silverman 2010; ten Tusscher et al. 2014).

A key question is whether neurotoxins of one sort or another have their effects on g—or in other words, is the intelligence loss that results from exposure to various neurotoxins associated with the Jensen effect? Establishing that this is the case is important to determining how temporal trends in levels of various neurotoxins might impact different variance components of intelligence. Unfortunately, only limited research has been conducted on this question. A bare-bones meta-analysis found that among Western and Chinese children, the impact of lead toxicity on subtests is largely indiscriminate with respect to their g loadings, being associated with a low-magnitude Jensen effect (N-weighted $\rho = .06$, $p < .05$, 95% CI = .02 to .10, $N = 2041$; Woodley of Menie et al. 2018). The effect of prenatal methylmercury exposure on young adults and young children is associated with a modest-magnitude Jensen effect ($r = .42$, $p < .05$, 95% CI = .37 to .47, $N = 1022$; Debes et al. 2015). This is consistent with the results of a multi-group confirmatory factor analysis conducted by the same group, which found evidence for a small-magnitude moderation effect of g on the group difference (Debes et al. 2016). Reanalysis of published data on the effects of organic mercury exposure in another study (Marques et al. 2016) reporting these data for the five subscales of the Stanford-Binet V test in a sample of Amazonian children yields a weak magnitude anti-Jensen Effect by contrast ($r = -.18$, $p < .05$, 95% CI = $-.28$ to $-.08$, $N = 365$). Alcohol (which, like lead, is an "ever-present neurotoxin") most adversely affects the intelligence of infants exposed to it in utero. Flynn et al. (2014) found that fetal alcohol effects on cognitive abilities, as with lead toxicity, are associated with a low-magnitude Jensen effect ($\rho = .12$, ns, 95% CI = $-.06$ to .29, $N = 125$). Flynn et al. (2014) found that fetal

cocaine exposure, by contrast, is associated with a low-magnitude anti-Jensen effect ($\rho = -.23$, $p < .05$, 95% CI = $-.35$ to $-.10$, $N = 215$). Finally, Calderón-Garcidueñs et al. (2008) reported data on the negative effects of atmospheric pollution (which causes neuroinflamation) on cognitive ability using the WISC-R for a sample of 55 individuals sourced from Mexico City, compared with a control group of 18 sourced from the countryside. Metzen (2012) reanalyzed these data utilizing MCV, finding a low-magnitude anti-Jensen effect ($r = -.17$, ns, 95% CI = $-.39$ to .06, $N = 73$). Aggregating across all of these vector correlations reveals a low-magnitude Jensen effect ($\rho = .11$, $p < .05$, 95% CI = .08 to .14, $N = 3841$).

These results suggest that increased exposure to various neurotoxins might reduce g; however, it is unclear whether the results of these vector correlation analyses are entirely meaningful, given that in many of these studies, the healthy control groups (against whom the exposure groups are compared) are likely to have higher g, simply as a consequence of the tendency of lower levels of g to predispose individuals to poorer quality life outcomes in many respects (see Herrnstein and Murray 1994). Thus, the comparison of group means on subtests is likely to be contaminated with a pre-existing g difference between comparison groups. This will in turn confound to some degree differences in s stemming from the direct action of neurotoxins on narrow cognitive abilities, potentially yielding zero-magnitude vector correlations.

The assumption that neurotoxic chemicals do not have substantial effects on g in infants and children - or, at the very least, effects that persist through to adulthood - is bolstered by the observation that g is only very weakly associated with measures of developmental instability such as fluctuating asymmetry (Banks et al. 2010), indicating that the trait is strongly canalized against environmental factors that would disturb its development (Woodley of Menie and Fernandes 2016b). g also has a very high narrow-sense heritability in adulthood (.54–.88, Bouchard Jr 2004, p. 150; .86, when g is directly modeled as a latent variable, Panizzon et al. 2014), which appears to be relatively stable across time, social groups, and countries (Bates et al. 2016; Figlio et al. 2017; Plomin 2002, Rushton 1989, Sundet et al. 1988, cf. Tucker-Drob and Bates 2016). An implication of these findings is that, in adults, g has low *global* trait modifiability (Sesardic 2005) and is correspondingly resistant to factors that would substantially alter its heritability in the course of development (such as the sort of gene-environment interactions on which Demeneix's (2014) model relies heavily; Sesardic (2005) notes that despite decades searching for gene-environment interactions on g, no unambiguous examples have been found—apparent environmental effects on children typically fade out in adulthood as additivity rises). The preponderance of the data does not therefore appear supportive of a non-trivial role for neurotoxins in causing secular declines in g among the adult populations of Western nations.

---

[2] Demeneix's (2014, 2017) somewhat confused qualitative inferences might result in part from an apparent lack of familiarity with the literature on secular trends in intelligence. In discussing the Flynn effect (Demeneix 2017, pp. 75–91), for example, she wrongly maintains that Woodley et al.'s (2013) finding of a decline in g is at odds with studies revealing Flynn effects. However, Woodley et al. (2013, 2014) made clear that the Flynn effect and dysgenesis on g can co-occur as they involve different phenotypic variance components of intelligence.

Directly Testing the Genetic and Neurotoxin Causation Models

Although the weight of evidence reviewed thus far indicates that the neurotoxin causation model is considerably weaker than the genetic causation model, a direct test of both models is possible utilizing temporal analysis. A chronometric *g.h* factor comprised of various converging indicators of declining *g.h* has already been described and validated in a previous temporal analysis (Woodley of Menie et al. 2017a). The existence of data on secular trends in both polygenic scores associated with cognitive ability and neurotoxins allow for a potentially very direct test of the influence of genetic and neurotoxic factors on the chronometric *g.h* factor. Based on the foregoing considerations, it is expected that the decline in the *g.h* factor will be driven predominantly by the genetic predictor and not by the neurotoxin predictor.

Methods

Variables

Three chronometric factors were computed, capturing declining *g.h*, polygenic scores comprised of variants predictive of educational attainment and *g* (the two measures share around 60% linkage-pruned genetic variance; Okbay et al. 2016) and neurotoxins respectively. In an effort to ensure high spatial contiguity, the data are mostly sourced from populations or (in the case of some of the neurotoxin measures) physical environments in the US and UK, the two exceptions being one of the polygenic score trends, which came from the population of Iceland (where the pattern of selection against educational attainment mirrors that found in the US and UK), and the dioxin trend, which came from sediments sourced from a lake in Europe. Missing data among these chronometric factors were handled via multivariate imputation (McKnight et al. 2007). All data utilized in the present analysis are publicly available on Dryad (https://doi.org/10.5061/dryad.nb301).

Chronometric "Heritable *g*" Factor

This chronometric factor captures the common temporal trend variance among five indicators of declining *g.h* and was computed using unit-weighted factor scoring which simply involves standardizing the indicators and creating an average—with the indicator-average correlations serving as factor loadings (Gorsuch 1983). This "course-grained" factor analysis is recommended when either indicator or case numbers are low and it furthermore avoids problems associated with sample specificity of factor scoring coefficients generated by standard errors of inconsistent magnitudes across differently sized samples (Gorsuch 1983). The *g.h*-decline indicators include

(i) US and UK simple visual reaction time means corrected for methods variance (sourced from Woodley et al. 2014); (ii) US backwards digit span means (sourced from Woodley of Menie and Fernandes 2015); (iii) a common factor comprised of the utilization frequencies in Google Ngram Viewer of four high-difficulty (as determined via item response theory) WORDSUM words (sourced from Woodley of Menie et al. 2015a; see Woodley of Menie et al. 2017a, p.78, Table 9 for information on the unit-weighted factor loadings for each item); (iv) a common factor among the utilization frequencies of 10 altruism-indicating words employed by Charles Darwin (Darwin 1871; see Woodley of Menie et al. 2017a, p.78, Table 8 for information on unit-weighted factor loadings of each item) in his description of the phenomenon of social selection. And, finally, (v) declining per capita rates of macro-innovation sourced from the US and UK and weighted by the population sizes of these countries (from Bunch and Hellemans 2004, as utilized in the analysis of Huebner 2005). Therefore, the chronometric *g.h* factor captures declining *g.h* as measured by processing speed, working memory, vocabulary ability, social intelligence and complex problem-solving ability. The data were available for the years 1850 to 2008, and the properties of the chronometric *g.h* factor are fully described by Woodley of Menie et al. (2017a; the indicator-level data are available in full in the Nexus 200 supplementary data appendix).

Chronometric Educational Attainment Polygenic Score Factor

The educational attainment polygenic scores were sourced from two populations. The first was computed using a framework set of approximately 620,000 high-quality single nucleotide polymorphisms covering the entire genome in a large sample ($N = $ 129,808 individuals) representative of the entire population of Iceland (Kong et al. 2017). Kong et al. (2017) break their sample out into eight lots of birth year bins (spanning 1916 to 1986) and fit a curve to the means. The data were extracted directly from their Figure 2 (p. E729). The second polygenic score trend was obtained from Domingue et al. (2017), who compute an educational attainment polygenic score, drawing from a set of 1.7 M single nucleotide polymorphisms, for a combined sex sample of 8845 non-Hispanic Whites, sourced from the US Health and Retirement Study. The data are available for 37 years, spanning the 1919 to 1955 birth cohorts (these being the cohorts for which sampling was highest). The data were extracted directly from Figure 5 in the *bioArxiv* version of their paper (p. 27) at yearly intervals. The presence of selection acting directly against these variants in this particular sample has been confirmed in earlier studies by Conley et al. (2016) and Beauchamp (2016), with the latter finding that selection against these variants should be reducing attained years of education by around 1.5 months per generation, which equates to loss in *g* of 1.06 IQ points (Woodley of Menie et al. 2017a). The two secular trends are

Table 1 Unit-weighted loadings of the polygenic score chronometric factor onto each polygenic score trend along with year range, observation number and source

Polygenic score indicators	Polygenic score factor loadings	Range (years)	Number of observations	Source
Icelandic polygenic scores	.951**	1916-1986	8	Kong et al. (2017)
US polygenic scores	.989**	1919-1955	37	Domingue et al. (2017)

**p < .001

here concatenated into a single unit-weighted polygenic score chronometric factor, the loadings of which are presented in Table 1. The loadings are extremely high magnitude, indicating that the factor is well specified. The data span the years 1916 to 1986, encompassing 42 observations.

It must be noted that only the trends with respect to the 1940 and more recent birth cohorts can be said to mostly capture the effects of genetic selection, as longevity and g are positively correlated; hence, the older samples will have higher polygenic scores in part due to a survival bias—indeed, longevity and educational attainment/g are *genetically*, as well as phenotypically, correlated (Domingue et al. 2017; Kong et al. 2017). Nevertheless, it has been found that the secular decline persists even after explicitly controlling for survival bias (Kong et al. 2017); thus, the decline in the means of the polygenic scores utilized in the present analysis can be said to be due to a mixture of both selection and survival bias.

Chronometric Neurotoxin Factor

Neurotoxin indicators were selected on the basis that they satisfy Demeneix's (2014) criterion of being "ever present." This is taken to imply that the substances in question have been present in the environment for over a century (i.e., since the Industrial Revolution of the nineteenth century at the least). The "ever present" criterion is significant, as indications of declining g have been noted since the nineteenth century (Woodley of Menie et al. 2017a), thus for neurotoxic substances to have been primarily responsible for these trends, they must have been present in the environment at least since then. Four neurotoxins were found to qualify as "ever present" (based on the aforesaid definition): (i) mercury pollution, measured as concentrations (ng/l) using ice cores obtained from the Fremont Glacier, Wyoming (data available from 1879 to 1994, from Clarke 2015, extracted from Figure 1, p. 47, originally from Schuster et al. 2002); (ii) lead pollution. Measured based on exposure (in tons per thousand of the US population for both gasoline and white lead; data available from 1876 to 1986, from Nevin 2000, extracted from Figure 12, p. 21);

PCDD/PCDF (dioxin and furan) pollution, estimated based on toxic equivalent sediment concentration (ng/kg) from Lake Constance, Europe (data available from 1895 to 1993, from Hagenmeier and Walczok 1996, extracted from Figure 2, p.103; data on sediment core B was used, as pollutant levels were sampled for more time points), and finally alcohol consumption, which was obtained for the years 1899 to 1995 for the UK (measured in liters of ethanol consumed per capita, sourced from portmangroup.org.uk) and for the years 1876 to 1995 for the US (measured in gallons consumed per capita for all beverages, sourced from Vinepair.com). Both UK and US alcohol consumption data were combined into a common unit-weighted alcohol factor (the temporal correlation between both measures was 0.26, p=0.055, N=55 years).

In order to conserve Brunswik (1952) symmetry, the neurotoxin trends were matched to the same level of chronometric aggregation as the g.h factor (five indicators) and the polygenic score factor (two indicators). There is furthermore a reasonable expectation of temporal covariance among these indicators as their time trends should be jointly influenced by industrialization (which will have driven up environmental levels of lead, mercury and dioxin, and also alcohol production). Therefore, the four neurotoxin trends were standardized and loaded onto a unit-weighted chronometric neurotoxin factor. The loadings are listed in Table 2—consistent with expectations, all are positive in sign and all load significantly onto the factor, indicating that the factor is reasonably well specified. The data span the years 1876 to 1995, encompassing 120 observations.

Results

Model Specifications and Output

Two analyses were conducted. The first involved simply establishing a matrix of temporal bivariate correlations among time and the chronometric factors. The second analysis involved a more detailed linear mixed model, in which the polygenic score and neurotoxin chronometric factors were used to predict the g.h factor, lagged by one "standard" (25 year) generation (so, for example, 1916 polygenic scores were aligned with 1941 g.h etc.). If neurotoxins have their biggest negative effects on g in childhood or infancy (or even in utero), the suppressing effects should manifest at a later point in time, subsequent to exposure. The polygenic score means were collected from different birth cohorts; thus, the impact of changes in these means on population g.h will also only manifest in subsequent decades. To avoid any temporal confounds, the continuous variable of time was polytomized into fifths of a century 20 year periods and used as a nesting variable containing toxicity and polygenic scores independently.
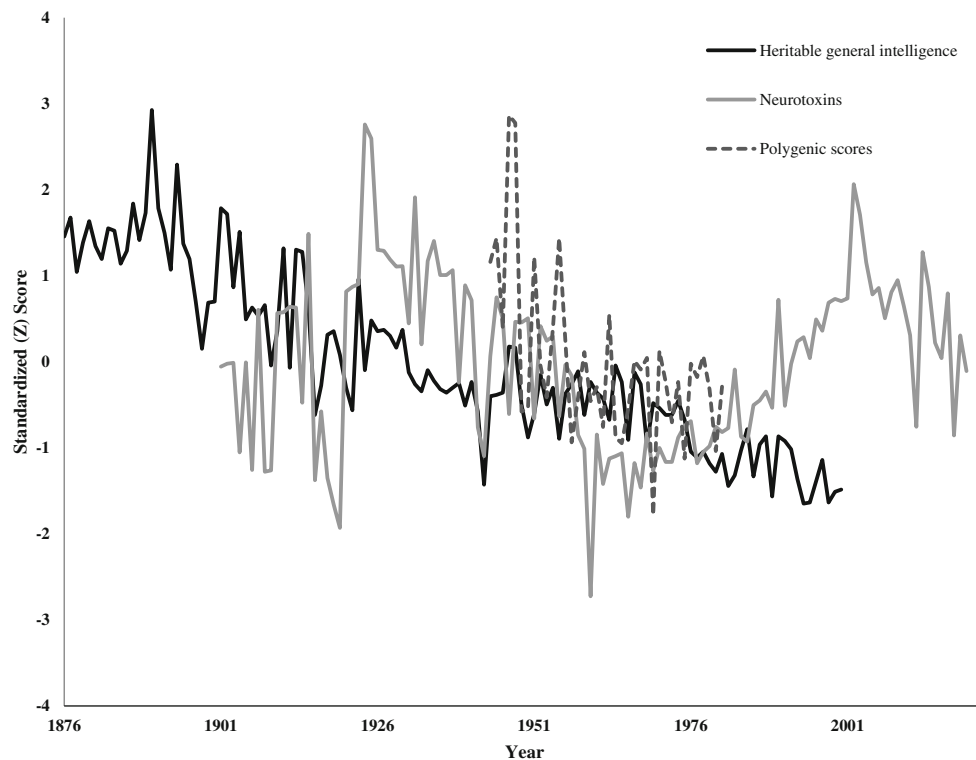
Fig. 1 Temporal trends for the neurotoxin and polygenic score chronometric factors along with g.h (with the predictors lagged by 25 years)

All data were analyzed via a hierarchical linear mixed model. Consequently, to predict *g.h.*, the time period variable was entered first followed by neurotoxins nested within fifths of a century periods and polygenic scores nested within periods. To determine whether autoregressive effects were present among the data, the covariance parameter estimate was computed using a first-order heterogeneous autoregressive structure (ARH(1)). The results of this analysis indicated no significant effect of autoregression upon the time trends analyzed. Toxicity and polygenic scores were standardized prior to the analysis. Partial η^2 were computed for each variable in the model.

Table 3 presents the bivariate temporal correlation matrix among time, the polygenic score, neurotoxin and *g.h* chronometric factors.

Among the bivariate temporal correlations, time and polygenic scores both significantly predict the decline in *g.h.*

Neurotoxins do not predict the decline in *g.h*; neither are they predicted by time. There appears to be a significant positive association between neurotoxins and polygenic scores (neurotoxin levels are higher when polygenic score levels are higher). There is no theoretical reason to expect such an association, which may simply reflect chance variation.

The nested (within period) and non-nested (between period) hierarchical linear mixed model analyses are presented in Tables 4 and 5. For the nested analysis, the intercepts were modeled as random effects. All analyses were conducted using SAS 9.3; for the second set of analyses, the command *proc MIXED* was used.

The nested hierarchical LMM reveals no significant differences in the *g.h* means when the means for the third and fourth fifth of a century periods are compared with the reference period (fifth fifth of a century , $\beta = 0$). The first and second fifth of a century periods

Table 2 Unit-weighted loadings of the neurotoxin chronometric factor onto each neurotoxin indicator along with year range, observation number and source

Neurotoxin indicators	Neurotoxin factor loadings	Range (years)	Number of observations	Source
Lead	.314*	1876-1986	111	Nevin (2000)
Mercury	.871**	1879-1994	50	Schuster et al. (2002)
Alcohol consumption (US+UK)	.888**	1876-1995	106	Portmangroup.org.uk (UK), Vinepair.com (US)
Dioxin+Furan	.602*	1895-1993	14	Hagenmeier and Walczok (1996)

**$p <.001; * $p <.01

Table 3 Bivariate correlation matrix among time, the *g.h*, polygenic score and neurotoxin chronometric factors utilizing the full range of years for which each factor was available

	g.h	Time	Neurotoxins	Polygenic score
g.h	–	− .89**	.033	.393*
Time		–	.004	− .577**
Neurotoxins			–	.319*
Polygenic score				–

***p < .0001, **p < .01, *p < .05

were associated with too much "missingness" among the covariates; therefore they were excluded from the analysis.

Among the predictors, only the decline in polygenic scores in the third fifth of a century period significantly predicted the decline in *g.h*; however, neurotoxins have no significant within period effects on *g.h*.

The hierarchical LMM reveals that only time (operationalized as fifth of a century periods) and polygenic scores independently predict significant amounts of the variance in declining *g.h*.

The temporal trends of the three factors (with 25 year lagging) are graphed in Fig. 1.

Discussion

The results of both of the analyses are consistent with the prediction, as the polygenic score factor is the only one of the two predictors (other than time period) to significantly predict variance in declining *g.h*. The relationships are furthermore not confounded with autoregressive effects (as established via the 0 value for the ARH(1) step in the multi-level model). Only polygenic scores in the third fifth of a century period had significant positive effects as predictors of *g.h* decline; neurotoxins had no significant within-period effects and predicted a non-significant amount of the overall variance.

While the results of the linear mixed model analysis are consistent with our prediction, the amount of variance in the *g.h* decline explained overall by the polygenic score chronometric factor is modest (25.3%). There are several potential reasons for this. First, the most densely sampled of our polygenic score data in time (from the US cohort) came from the smaller of the two cohorts, which necessarily introduces measurement error into the point estimations of mean polygenic score by birth year. This is likely largely responsible for the high year-on-year heterogeneity observed among these data points. Conversely, the Icelandic sample (which had > 14 times the number of individuals) was much more poorly sampled across time (eight measurement occasions in total),

Table 4 Nested hierarchical linear mixed model examining the effect of the standardized neurotoxins (ZT) and standardized polygenic score (ZPGS) factors upon *g.h* nested within time periods (operationalized as fifths of a century)

Multilevel Linear Model					
		Model Fit			
		AIC	BIC	-2RLL	
		13.8	-8.2	-8.2	
		Solution for Random Effects			
Covariance Parameter	Estimate	Standard Error	Z-value	Significance	
ARH (1)	0	.	.	.	
Residual	.048	.011	4.47	<.0001	
		Solution for Fixed Effects			
Effect	Estimate (*β*)	Standard Error	Denominator DF	t-value	Significance
Intercept	-.697	.454	40	-1.53	.133
Third fifth 20th Century	.59	.093	40	.2	.842
Fourth fifth 20th Century	.231	.469	40	.49	.624
Fifth fifth 20th Century	0	.	.	.	.
ZT (Third fifth 20th Century)	.248	.133	40	1.86	.07
ZT (Fourth fifth 20th Century)	.014	.091	40	.15	.879
ZT (Fifth fifth 20th Century)	.375	.457	40	.82	.417
ZPGS (Third fifth 20th Century)	.202	.058	40	3.48	.001
ZPGS (Fourth fifth 20th Century)	-.08	.091	40	-.89	.38
ZPGS (Fifth fifth 20th Century)	.211	.265	40	.80	.43

Table 5 Hierarchical linear mixed model examining the effect of time period (operationalized as fifths of a century), polygenic score, and neurotoxin factors upon *g.h*. The partial η^2 values were estimated via a type I sum of squares GLM

Effect	Numerator *DF*	Denominator *DF*	*F* value	Significance	Partial η^2
Fifths of a century	2	40	25.72	< .0001	.584
ZT (fifths of a century)	3	40	.18	.906	.018
ZPGS (fifths of a century)	3	40	3.5	.027	.253

entailing that relatively high levels of "missingness" were present in those data. Polygenic scores are themselves relatively crude and suffer from the "missing heritability" problem (i.e., they predict much less variance in the traits of interest that is known to be under genetic control based on the results of classical behavior genetic studies). This means that a great many of the variants used in constructing these scores may be "false positives" for the trait of interest and that many additional variants that are under selection are likely to be absent from these scores. This potentially reduces their validity substantially. Second, measurement error is also present among the indicators comprising the *g.h* chronometric factor (although these are relatively much more densely sampled across time for the most part), in addition to which a case could be made that the factor is contaminated with variance stemming from additional sources, which is consistent with the observation from Woodley of Menie et al. (2017a) that two of the five indicators (simple visual reaction time and the social intelligence measure) did not exhibit the higher affinity for the parent chronometric *g.h* factor (relative to the latent Nexus factor common to all three chronometric factors) that the other three measures did. One of the indicators (declining per capita macro-innovation rates), while showing high affinity for its parent chronometric factor, is clearly not exclusively a product of declining *g* either and could plausibly be driven in part by social, economic and scientific factors operating in tandem with selection (such as efficiency ceilings, low hanging fruit e.g. Huebner, 2005). These "contaminants" will necessarily diminish the validity of the factor, introducing additional measurement error. Third, there are other genetic and demographic factors that will contribute to the decline in *g*, whose contributions will not be detected utilizing the polygenic score chronometric factor employed here. These include the effects of replacement migration involving lower-ability individuals immigrating to (predominantly Western) countries populated by higher-ability individuals (this has been found to be a significant predictor of the *anti-Flynn effect* or the tendency for tests that once showed secular gains to show declines instead, especially when the aggregate *g* loading of the test is high; Woodley of Menie et al. 2017b). Another factor contributing marginally to the decline in *g* might be the accumulation of deleterious mutations resulting from the relaxation of purifying selection on industrialized populations (Woodley of Menie and Fernandes 2016b). It is expected that the effects of these

additional factors will be captured by the effects of time period, which corresponds to all time-varying factors that contribute to the decline in *g.h* above and beyond the predictors.

As was mentioned in the "Methods" section, the polygenic score declines are not purely capturing genetic selection acting against these variants but are also capturing survival bias or mortality selection. This additional effect will potentiate the apparent cross-sectional decline in the polygenic scores among older cohorts (the effects of survival bias become exponentially more pronounced as a function of cohort age; Kong et al. 2017); thus, it might be the case that the significant effect of polygenic scores on *g.h* in the third fifth of a century period (which would contain the oldest cohorts) is confounded with this survival bias. This could be countered with the observation that the secular decline in polygenic scores persists, even when explicitly controlled for survival bias in the case of very large *N* samples (such as the Icelandic cohort utilized in Kong et al. 2017). Furthermore, although non-significant, the effect of polygenic scores in the fifth of a century period was also positive in sign, which is suggestive of the presence of the effect among the youngest cohorts also. The use of other very large *N* samples containing cognitive polygenic score data (such as UK BioBank, *N* > 500,000) and controlling explicitly for survival bias can allow for this issue to be comprehensively addressed in future research.

It might also be claimed that the polygenic score data supporting the "dysgenic" theory of decreases in *g.h* only appear to do so because they cover a relatively short span of years (compared to the neurotoxin factor)—in other words, what appears to be evidence for the "dysgenic" theory may be an instance of capitalizing on chance. The likelihood of this is low, however, in light of the fact that there are data indicative of selection against *g*, captured in terms of negative correlations between fertility and both IQ (Lynn 1996; Lynn and van Court 2004) and educational attainment (Skirbekk 2008), dating back to the early to mid-nineteenth century. This indicates a potentially constant downward trend in polygenic scores for a century and a half, at least in parts of the West.

Despite limitations, our analysis appears to be supportive of the central prediction of the co-occurrence model, i.e., that declining cognitive polygenic scores potentially resulting from selection have actually had real-world impacts on phenotypic indicators of *g*, rather than having their "dysgenic" effects attenuated by environmental improvements (via, e.g., establishment of the sorts of gene-environment interactions that might have

driven down the penetrance of these *g*-related variants, allowing "phenotypic" intelligence and educational attainment to rise—which represents an alternative solution to Cattell's paradox (Beauchamp 2016; Lynn 1996; cf. Woodley of Menie 2016). These results also constitute the first comprehensive test of neurotoxin theories of the decline in *g*. Since the temporal trends of a reasonable subsample of "ever-present" neurotoxic substances do not predict declining *g* across a reasonable span of years at the level of direct correlation or when competed with polygenic scores for a more restricted subset of years, neurotoxin caustion theories can in all likelihood be ruled out as accurate explanations of the so-called Woodley effect.

Some proponents of neurotoxin models of *g* (or more broadly intelligence) decline will doubtless be skeptical of the results of our study. Demeneix (2017, p. 88) suggests that Woodley et al.'s (2013) "dysgenic" argument denies the relevance to *g* decline of neurotoxic influences, (partly) in light of the fact that the lead levels of modernized environments have fallen in recent decades—for this constitutes a failure to appreciate the rising prevalence of other neurotoxins and their possible deleterious effects. She also raises the familiar but long-dead (Sesardic 2005) corpse of "biological determinism and genetic determinism"[3] (p. 87) to discredit the "dysgenic" theory.

These concerns would surely apply in equal measure to the present study but are nonetheless without merit, reflecting lack of familiarity with the germane science. As indicated earlier, Demeneix's (2014, 2017) theory does not explicitly offer testable predictions and furthermore appears to be constructed in such a way as to guard against disconfirmation, thus exhibiting the key feature of a "degenerating" research program (Lakatos 1970). Provided that Demeneix's (2017) theory is explicitly said to be compatible with *both* negative or positive (and so presumably even stable) temporal trajectories of population-level intelligence, it is unclear how the theory could be tested at all or what precisely it adds to explanations of time trends in intelligence. Should it be noted that changes in lead exposure (coupled with the other neurotoxins considered here) lack a meaningful statistical relationship with *g* (or more broadly speaking intelligence), Demeneix and similarly inclined researchers can (and have) simply invoke(d) some of the other innumerable putative neurotoxins—while shirking the obligation to offer predictions or subject their claims to appropriate testing.[4] But since the neurotoxins considered here appear to be unrelated to levels of *g.h* over time, despite having well-established intelligence-depressing effects, why suppose that other neurotoxins are to blame? Demeneix might update her theory to contend that neurotoxins have their adverse impact primarily or exclusively on narrow abilities (*s*), which are generally more sensitive to environmental

insult; but distinguishing such an impact from the countervailing factors driving the Flynn effect (in order to demonstrate that the former actually exists) would require substantial statistical analysis; purely or largely qualitative theorizing simply would not do.

There are further reasons to doubt that neurotoxins have any significant bearing on changes in *g*, reasons which connect with Demeneix's (2017) allegation of "biological . . . and [sic] genetic determinism." In the first place, and as reviewed in the "Introduction", the weight of empirical evidence leaves little doubt that *g* is under far stronger genetic than environmental control. Both Flynn et al. (2014) and Gottfredson (2005) make it clear that environmental insults and deprivations have little effect on *g*, and even in the worst cases (e.g., severe malnutrition and infectious disease burdens not known in the West for decades), losses of intelligence are rarely permanent (Gottfredson 2005, p.316). Moreover, efforts to *raise* intelligence, and certainly *g* (Jensen 1998), via direct environmental intervention in childhood are almost never successful (Bailey et al. 2017; Protzko 2015; te Nijenhuis et al. 2014). The politicized objection of "genetic determinism" does nothing to vitiate these findings.

Furthermore, one must ask why the environmentalist obfuscations[5] that Demeneix (2017, pp. 87–88) employs in her critique of the allegedly "genetic determinist" "dysgenic" theory do not apply with equal force to her own model. Given the degree to which *g* is under genetic control, it is entirely unsurprising that decreased population-level enrichment for *g*-related genetic variants stemming from selection has caused *g* to decline. Conversely, substantial *environmental* influences on *g* would be highly surprising in that *no* environmental factor has been shown to unambiguously or persistently affect *g* in adults, congruent with other evidence of the trait's low global modifiability. If, then, genes should *not* be expected to have rather direct effects on *g* levels (because of the supposedly inextricable contributions of genetic, epigenetic, and environmental factors to phenotypic development), such that a "dysgenic" theory of the sort offered here would be illegitimate, why should it be anticipated that environmental factors, which at most only very weakly affect *g*, *do* have such effects? What allows Demeneix (2014, 2017) and others to straightforwardly infer from increasing levels of neurotoxin exposure a secular loss, whether relative or absolute, in intelligence if the interactions among relevant causal factors are simply too complex to disentangle? As a general rule, *environmental determinists* about intelligence, are almost never aware that their endeavors to complicate the relationship between genes and

---

[3] Without distinguishing the two, giving the charge a pleonastic quality.
[4] This is perfectly consistent with environmental determinist stratagems used to trivialize or eliminate the role of genes in generating differences in intelligence (see Brand 1999).

[5] One example of these can be seen in her effort to explain IQ differences among individuals as a function of "modulat[ed] gene expression" related to "prenatal placental blood supply and nutrition" (Demeneix 2017, p. 87).

intelligence would create problems no less, or even more, serious for their favored theoretical models compared to hereditarian ones.

Finally, it needs to be stated that our findings as they pertain to temporal trends in neurotoxins and *g.h* should not be taken to indicate that environmental neurotoxin exposure is cognitively harmless. As Flynn et al. (2014) note, even though factors such as prenatal alcohol and cocaine exposure do not have strong effects on *g*, they still suppress performance on intelligence tests and reduce quality of life for the afflicted. Furthermore, our findings do not rule out the possibility that neurotoxin exposure may have influenced other, potentially more environmentally sensitive, aspects of human psychology. The results of Nevin's (2000) temporal analysis of the temporal trend in per capita lead exposure and out-of-wedlock pregnancy along with various forms of criminality, indicating an effect of the former on the latter, are quite compelling and may provide fertile grounds for exploring the impacts of changing levels of environmental neurotoxins on human behavioral ecology.

References

Bailey, D., Duncan, G. J., Odgers, C. L., & Yu, W. (2017). Persistence and fadeout in the impacts of child and adolescent interventions. *Journal of Research on Educational Effectiveness, 10*, 7–39.

Banks, G. C., Batchelor, J. H., & McDaniel, M. A. (2010). Smarter people are (a bit) more symmetrical: a meta-analysis of the relationship between intelligence and fluctuating asymmetry. *Intelligence, 38*, 393–401.

Bates, T., Hansell, N. K., Martin, N. G., & Wright, M. J. (2016). When does socioeconomic status (SES) moderate the heritability of IQ?: no evidence for *g* x SES interaction for IQ in a representative sample 1,176 Australian adolescent twin pairs. *Intelligence, 56*, 10–15.

Beauchamp, J. P. (2016). Genetic evidence for natural selection in humans in the contemporary United States. *Proceedings of the National Academy of Sciences USA, 113*, 7774–7779.

Bouchard Jr., T. J. (2004). Genetic influence on human psychological traits. *Current Directions in Psychological Science, 13*, 148–151.

Brand, C. (1999). IQ heritability—the emptiness of modern environmentalism. A review of: Robert J. Sternberg & Elena Grigorenko (eds.) (1997). *Intelligence, Heredity and Environment.* Cambridge, UK: Cambridge University press. *Personality and Individual Differences, 26*, 767–774.

Brunswik, E. (1952). *The conceptual framework of psychology (international encyclopedia of unified science, volume 1, number 10)*. Chicago, IL: The University of Chicago Press.

Bunch, B., & Hellemans, A. (2004). *The history of science and technology: a browser's guide to the great discoveries, inventions, and the people who made them from the dawn of time to today*. New York, NY: Houghton Mifflin Company.

Calderón-Garcidueñs, L., Mora-Tiscareño, A., Onitveros, E., Gómez-Garza, G., Barragán-Mejía, G., Broadway, J., … Engle, R.W. (2008). Air pollution, cognitive deficits and brain abnormalities: a pilot study with children and dogs. *Brain and Cognition, 68*, 117–127.

Cattell, R. B. (1950). The fate of national intelligence: test of a thirteen-year prediction. *The Eugenics Review, 42*, 136–148.

Cattell, R. B. (1937). *The fight for our national intelligence*. London, UK: P. S. King & Son, Ltd..

Clarke, R. P. (2015). Rising-falling mercury pollution causing the rising-falling IQ of the Lynn-Flynn effect, as predicted by the antiinnatia theory of autism and IQ. *Personality and Individual Differences, 82*, 46–51.

Conley, D., Laidley, T., Belsky, D. W., Fletcher, J. M., Boardman, J. D., & Domingue, B. W. (2016). Assortative mating and differential fertility by phenotype and genotype across the 20th century. *Proceedings of the National Academy of Sciences, USA, 113*, 6647–6652.

Darwin, C. (1871). *The descent of man, and selection in relation to sex*. London, UK: D. Appleton & Co..

Debes, F., Ludvig, A., Budtz-Jørgensen, E., Weihe, P., & Grandjean, P. (2015). The effects of methylmercury on general intelligence in children and young adults. Oral presentation given at the 16th Annual Meeting of the International Society for Intelligence Research. Albuquerque, New Mexico, USA, September.

Debes, F., Weihe, P., & Grandjean, P. (2016). Cognitive deficit at age 22 years associated with prenatal exposure to methylmercury. *Cortex, 74*, 358–369.

Demeneix, B. (2014). *Losing our minds: how environmental pollution impairs human intelligence and mental health*. Oxford, UK: Oxford University Press.

Demeneix, B. (2017). *Toxic cocktail: how chemical pollution is poisoning our brains*. New York, NY: Oxford University Press.

Domingue, B., Belsky, D. W., Harrati, A., Conley, D., Weir, D., & Boardman, J. (2017). Mortality selection in a genetic sample and implications for association studies. *International Journal of Epidemiology, 46*, 1285–1294. (BioArxiv: https://doi.org/10.1101/049635).

Figlio, D. N., Freese, J., Karbownik, K., & Roth, J. (2017). Socioeconomic status and genetic influences on cognitive development. *Proceedings of the National Academy of Sciences, USA, 114*, 13441–13446.

Flynn, J. R. (1987). Massive IQ gains in 14 nations: what IQ tests really measure. *Psychological Bulletin, 101*, 171–191.

Flynn, J. R., te Nijenhuis, J., & Metzen, D. (2014). The *g* beyond Spearman's *g*: Flynn's paradoxes resolved using four exploratory meta-analyses. *Intelligence, 44*, 1–10.

Fox, M. C., & Mitchum, A. L. (2013). A knowledge based theory of rising scores on "culture-free" tests. *Journal of Experimental Psychology: General, 142*, 979–1000.

Galton, F. (1869). *Hereditary genius*. London, UK: MacMillan.

Gorsuch, R. L. (1983). *Factor analysis*. Hillside, NJ: L. Earlbaum Associates.

Gottfredson, L. S. (2005). What if the hereditarian hypothesis is true? *Psychology, Public Policy, and Law, 11*, 311–319.

Hagenmeier, H., & Walczok, M. (1996). Time trends in levels, patterns and profiles for PCDD/PCDF in sediment cores of Lake Constance. *Organohalogen Compounds, 28*, 101–104.

Herrnstein, R. J., & Murray, C. (1994). *The bell curve: intelligence and class structure in American life*. New York, NY: Free Press Paperbacks.

Higgins, J. V., Reed, E. W., & Reed, S. C. (1962). Intelligence and family size: a paradox resolved. *Eugenics Quarterly, 9*, 84–90.

Huebner, J. (2005). A possible declining trend for worldwide innovation. *Technological Forecasting and Social Change, 72*, 980–986.

Jensen, A. R. (1998). *The g factor: the science of mental ability*. Westport, CT: Praeger.

Kan, K. J., Wicherts, J. M., Dolan, C. V., & van der Maas, H. L. J. (2013). On the nature and nurture of intelligence and specific cognitive abilities: the more heritable, the more culture dependent. *Psychological Science, 24*, 2420–2428.

Kong, A., Banks, E., Poplin, R., Garimella, K. V., Maguire, J. R., et al. (2017). Selection against variants in the genome associated with educational attainment. *Proceedings of the National Academy of Sciences, USA, 114*, E727–E732.

Lakatos, I. (1970). Falsificationism and the methodology of scientific research programmes. In I. Lakatos & A. Musgrave (Eds.), *Criticism and the growth of knowledge* (pp. 91–196). Cambridge, UK: Cambridge University Press.

Lievens, F., Reeve, C. L., & Heggestad, E. D. (2007). An examination of psychometric bias due to retesting in selection settings. *Journal of Applied Psychology, 92*, 1672–1682.

Lynn, R. (1996). *Dysgenics: genetic deterioration in modern populations*. Westport, CT: Praeger.

Lynn, R., & van Court, M. (2004). New evidence of dysgenic fertility for intelligence in the United States. *Intelligence, 32*, 193–201.

Madison, G., Woodley of Menie, M. A., & Sänger, J. (2016). Possible secular slowing auditory simple reaction time in Sweden (1959-1985). *Frontiers in Human Neuroscience, 10*, 407.

Marques, R. C., Bernardi, J. V. E., Cunha, M. P. L., & Dórea, J. G. (2016). Impact of organic mercury exposure and home delivery on neurodevelopment of Amazonian children. *International Journal of Hygiene and Environmental Health, 219*, 498–502.

McKnight, P. E., McKnight, K. M., Sidani, S., & Figueredo, A. J. (2007). *Missing data: a gentle introduction*. New York, NY: Guildford Publications.

Metzen, D. (2012). *The causes of group differences in intelligence studied using the method of correlated vectors and psychometric meta-analysis*. Unpublished Masters Thesis, University of Amsterdam.

Millet, K., & Dewitte, S. (2007). Altruistic behaviour as a costly signal of general intelligence. *Journal of Research in Personality, 41*, 316–326.

Murray, C. (2003). *Human accomplishment: the pursuit of excellence in the arts and sciences, 800 BC to 1950*. New York, NY: Harper Collins.

Nevin, R. (2000). How lead exposure relates to temporal changes in IQ, violent crime, and unwed pregnancy. *Environmental Research Section A, 83*, 1–22.

Okbay, A., Beauchamp, J.P., Fontana, M.A., Lee, J.J., Pers, T.H., Rietveld, C.A., et al. (2016). Genome-wide association study identifies 74 loci associated with educational attainment. *Nature, 533*, 539–542.

Panizzon, M. S., Vuoksimaa, E., Spoon, K. M., Jacobson, K. C., Lyons, M. J., et al. (2014). Genetic and environmental influences on general cognitive ability: is *g* a valid latent construct? *Intelligence, 43*, 65–76.

Pietschnig, J., & Gittler, G. (2015). A reversal of the Flynn effect for spatial perception in German speaking countries: Evidence from a crosstemporal IRT-based meta-analysis (1977–2014). *Intelligence, 53*, 145–153.

Pietschnig, J., & Voracek, M. (2015). One century of global IQ gains: a formal meta-analysis of the Flynn effect (1909–2013). *Perspectives on Psychological Science, 10*, 282–306.

Pietschnig, J., & Gittler, G. (2017). Is ability-based emotional intelligence impervious to the Flynn effect? A cross-temporal meta-analysis (2001-2015). *Intelligence, 61*, 37-45.

Plomin, R. (2002). Quantitative trait loci and general cognitive ability. In J. Benjamin, R. P. Ebstein, & R. H. Belmaker (Eds.), *Molecular genetics and human personality* (pp. 211–230). Washington, DC: American Psychiatric Publishing, Inc..

Portman Group (n.d.). Retrieved September 16, 2017, from http://www.portmangroup.org.uk/

Protzko, J. (2015). The environment in raising early intelligence: a meta-analysis of the fadeout effect. *Intelligence, 53*, 202–210.

Reeve, C. L., & Lam, H. (2007). The relation between practice effects, test-taker characteristics and degree of *g*-saturation. *International Journal of Testing, 7*, 225–242.

Rushton, J. P. (1998). The "Jensen effect" and the "Spearman–Jensen hypothesis" of black–white IQ differences. *Intelligence, 26*, 217–225.

Rushton, J. P. (1989). The generalizability of genetic estimates. *Personality and Individual Differences, 10*, 985–989.

Sarraf, M. (2017). Review of historical variability in heritable general intelligence: its evolutionary origins and sociocultural consequences. *Personality and Individual Differences, 109*, 238–241.

Schuster, P. F., Krabbenhoft, D. P., Nafz, D. L., Cecil, L. D., Olson, M. L., Dewild, J. F., et al. (2002). Atmospheric mercury deposition during the last 270 years: a glacial ice core record of natural and anthropogenic sources. *Environmental Science and Technology, 36*, 2303–2310.

Sesardic, N. (2005). *Making sense of heritability*. Cambridge, UK: Cambridge University Press.

Silverman, I. W. (2010). Simple reaction time: it is not what it used to be. *American Journal of Psychology, 123*, 39–50.

Skirbekk, V. (2008). Fertility trends by social status. *Demographic Research, 18*, 145–180.

Spearman, C. (1904). General intelligence, objectively determined and measured. *American Journal of Psychology, 15*, 201–293.

Sundet, J. M., Tambs, K., Magnus, P., & Berg, K. (1988). On the question of secular trends in the heritability of intelligence test scores: a study of Norwegian twins. *Intelligence, 12*, 47–59.

te Nijenhuis, J., & van der Flier, H. (2013). Is the Flynn effect on *g*?: a metaanalysis. *Intelligence, 41*, 802–807.

te Nijenhuis, J., Jongeneel-Grimen, B., & Kirkegaard, E. O. W. (2014). Are Headstart gains on the *g* factor? A meta-analysis. *Intelligence, 46*, 209–215.

te Nijenhuis, J., Jongeneel-Grimen, B., & Armstrong, E. L. (2015). Are adoption gains on the *g* factor? A meta-analysis. *Personality and Individual Differences, 73*, 56–60.

te Nijenhuis, J., van Vianen, A. E. M., & van der Flier, H. (2007). Sore gains on *g*-loaded tests: no *g*. *Intelligence, 35*, 283–300.

ten Tusscher, G.W., Leijs, M., Boer, L.D., Legler, J., Olie, K., Spekreijse, H., … Koppe, J. (2014). Neurodevelopmental retardation, as assessed clinically and with magnetoencephalography and electro-encephalography, associated with perinatal dioxin exposure. *Science of the Total Environment, 491-492*, 235–239.

Tucker-Drob, E. M., & Bates, T. C. (2016). Large cross-national differences in gene x socioeconomic status interaction on intelligence. *Psychological Science, 27*, 138–149.

VinePair. (n.d.). Retrieved September 15, 2017, from https://www.vinepair.com.

Voronin, I., te Nijenhuis, J., & Malykh, S. B. (2015). The correlation between *g* loadings and heritability in Russia. *Journal of Biosocial Science, 28*, 1–11.

Wicherts, J. M., Dolan, C. V., Oosterveld, P., Hessen, D. J., van Baal, G. C. M., Boomsma, D. I., et al. (2004). Are intelligence tests measurement invariant over time? Investigating the nature of the Flynn effect. *Intelligence, 32*, 509–537.

Wongupparaj, P., Wongupparaj, R., Kumari, & Morris, R. G. (2017). The Flynn effect for verbal and visuospatial short-term and working memory: a cross-temporal meta-analysis. *Intelligence, 64*, 71–80.

Woodley, M. A. (2012). The social and scientific temporal correlates of genotypic intelligence and the Flynn effect. *Intelligence, 40*, 189–204.

Woodley, M. A., & Figueredo, A. J. (2013). *Historical variability in heritable general intelligence: its evolutionary origins and socio-cultural consequences*. Buckingham, UK: Buckingham University Press.

Woodley, M. A., & Madison, G. (2013). Establishing an association between the Flynn effect and ability differentiation. *Personality and Individual Differences, 55*, 387–390.

Woodley, M. A., te Nijenhuis, J., & Murphy, R. (2013). Were the Victorians cleverer than us? The decline in general intelligence

estimated from a meta-analysis of the slowing of simple reaction time. *Intelligence, 41*, 843–850.

Woodley, M. A., te Nijenhuis, J., & Murphy, R. (2014). Is there a dysgenic secular trend towards slowing simple reaction time? Responding to a quartet of critical commentaries. *Intelligence, 46*, 131–147.

Woodley of Menie, M. A. (2016). Consideration of cognitive variance components potentially solves Beauchamp's paradox. *Proceedings of the National Academy of Sciences, USA, 113*, E5780–E5781.

Woodley of Menie, M. A., & Fernandes, H. B. F. (2015). Do opposing secular trends on backwards and forwards digit span evidence the co-occurrence model? A comment on Gignac (2015). *Intelligence, 50*, 125–130.

Woodley of Menie, M. A., & Fernandes, H. B. F. (2016a). Showing their true colours: possible secular declines and a Jensen effect on color acuity—more evidence for the weaker variant of Spearman's other hypothesis. *Personality and Individual Differences, 88*, 280–284.

Woodley of Menie, M. A., & Fernandes, H. B. F. (2016b). The secular decline in general intelligence from decreasing developmental stability: theoretical and empirical considerations. *Personality and Individual Differences, 92*, 194–199.

Woodley of Menie, M. A., Fernandes, H. B. F., Figueredo, A. J., & Meisenberg, G. (2015a). By their words ye shall know them: evidence of genetic selection against general intelligence and concurrent environmental enrichment in vocabulary usage since the mid 19th century. *Frontiers in Psychology, 6*, 361.

Woodley of Menie, M. A., Figueredo, A. J., Sarraf, M. A., Hertler, S., Fernandes, H. B. F., & Peñaherrera Aguirre, M. (2017a). *The rhythm of the west: a biohistory of the modern era, AD 1600 to the present. Journal of Social Political and Economic Studies, monograph series, no. 37.* Washington, DC: Scott Townsend Press.

Woodley of Menie, M. A., Peñaherrera-Aguirre, M., Fernandes, H. B. F., & Figueredo, A. J. (2017b). What causes the anti-Flynn effect? A data synthesis and analysis of predictors. *Evolutionary Behavioral Sciences.* https://doi.org/10.1037/ebs0000106.

Woodley of Menie, M. A., Schwartz, J. A., & Beaver, K. M. (2016). How cognitive genetic factors influence fertility outcomes: a meditational SEM analysis. *Twins Research and Human Genetics, 19*, 628–637.

Woodley of Menie, M. A., te Nijenhuis, J., & Murphy, R. (2015b). The Victorians were still faster than us. Commentary: Factors influencing the latency of simple reaction times. *Frontiers in Human Neuroscience, 9*, 452.

Woodley of Menie, M.A., te Nijenhuis, J.,Shibaev, V., Li,M., & Smit, J. (2018). Are the effects of lead exposure linked to the *g* factor? A meta-analysis. Working Paper.