

Discontinuity in the genetic and environmental causes of the intellectual disability spectrum

Abraham Reichenberg^{a,b,1}, Martin Cederlöf^c, Andrew McMillan^d, Maciej Trzaskowski^d, Ori Kapara^e, Eyal Fruchter^f, Karen Ginat^f, Michael Davidson^e, Mark Weiser^{e,f}, Henrik Larsson^c, Robert Plomin^d, and Paul Lichtenstein^c

^aDepartment of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY 10029; ^bDepartment of Preventive Medicine, Icahn School of Medicine at Mount Sinai, New York, NY 10029; ^cDepartment of Medical Epidemiology and Biostatistics, Karolinska Institutet, SE-171 77 Stockholm, Sweden; ^dMedical Research Council Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology & Neuroscience, King's College London, SE5 8AF London, United Kingdom; ^eDepartment of Psychiatry, Sheba Medical Center, Tel-Hashomer, 52621 Ramat Gan, Israel; and ^fDepartment of Mental Health, Israel Medical Corps, 01259 Tel-Hashomer, Israel

Edited by Daniel J. Benjamin, University of Southern California, Los Angeles, CA, and accepted by the Editorial Board November 13, 2015 (received for review April 24, 2015)

Intellectual disability (ID) occurs in almost 3% of newborns. Despite substantial research, a fundamental question about its origin and links to intelligence (IQ) still remains. ID has been shown to be inherited and has been accepted as the extreme low of the normal IQ distribution. However, ID displays a complex pattern of inheritance. Previously, noninherited rare mutations were shown to contribute to severe ID risk in individual families, but in the majority of cases causes remain unknown. Common variants associated with ID risk in the population have not been systematically established. Here we evaluate the hypothesis, originally proposed almost 1 century ago, that most ID is caused by the same genetic and environmental influences responsible for the normal distribution of IQ, but that severe ID is not. We studied more than 1,000,000 sibling pairs and 9,000 twin pairs assessed for IQ and for the presence of ID. We evaluated whether genetic and environmental influences at the extremes of the distribution are different from those operating in the normal range. Here we show that factors influencing mild ID (lowest 3% of IQ distribution) were similar to those influencing IQ in the normal range. In contrast, the factors influencing severe ID (lowest 0.5% of IQ distribution) differ from those influencing mild ID or IQ scores in the normal range. Taken together, our results suggest that most severe ID is a distinct condition, qualitatively different from the preponderance of ID, which, in turn, represents the low extreme of the normal distribution of intelligence.

intelligence | twins | heritability | intellectual disability | family study

Intellectual disability (ID), previously known as mental retardation, refers to lifelong cognitive impairment that emerges in childhood. In an increasingly technological world, intellectual disability entails tremendous personal costs for affected individuals and their families, as well as to society in terms of lost intellectual capital. At a purely economic level, the average lifetime additional cost per person with ID is more than \$1 million (1). ID is primarily defined by low cognitive ability, typically IQ scores below 70, which is two SDs below the population mean of 100, resulting in a prevalence of 2–3%. In the rare (<0.5%) and most severe form of ID IQ scores do not exceed 35 (2, 3).

Investigators of intellectual development have most often assumed that ID is the extreme low of the normal IQ distribution. Both cognitive ability and disability are among the most heritable behavioral traits (4). However, a paradox has emerged in genetic research on ID. Exciting advances have been made in identifying noninherited (de novo) mutations as a major source of severe ID (IQs < 35) (5–7), but little progress has been made in identifying genes associated with inherited causes of ID. A resolution to this paradox is that most ID is at the lowest end of the normal distribution of IQ, but severe ID is etiologically distinct, as proposed initially by Lionel Penrose in 1938 (8). This hypothesis has two components: (i) most ID is caused by the same genetic and environmental influences responsible for the normal distribution of IQ and (ii) severe ID is not.

Here, we tested both components of this hypothesis using data from cognitive assessments administered as part of compulsory military service in Sweden 1968–2010 (3 million 18-y-old males) and in Israel 1960–2005 (2.1 million 17-y-old males and females) with 98% male participation in both countries. Diagnoses of severe ID were also available in both countries based on the International Classification of Disease (ICD) (9). From these resources, we identified 400,426 non-twin pairs of brothers and 8,788 male twin pairs from Swedish conscripts and 610,391 sibling pairs of both sexes from Israeli conscripts. From patient records, we identified whether these individuals had a sibling with a diagnosis of severe ID.

From the two cohorts assessed for mandatory military conscription, those who received the lowest possible score (IQ stanine score of 1), and were therefore placed in the bottom 3% (i.e., mild ID) of the distribution, were selected and their siblings compared with the rest of the distribution using IQ as a quantitative trait measure. Importantly, this was repeated for the siblings of individuals diagnosed with severe ID. This study design, combined with DF extremes analysis, a technique named after its developers DeFries and Fulker (10), allows examination of the fundamental issue of the etiologic links between the abnormal and normal. That is, to what extent are the causes of disorders qualitatively different from the influences on normal variation? Specifically, is ID merely the lower tail of the distribution for the same genetic and environmental

Significance

Intellectual disability (ID) is present in almost 3% of children and fundamentally characterized by IQ scores below 70. Genetic research has shown that it is among the most heritable traits, and it has been accepted that ID is the extreme low of the normal IQ distribution. However, we show that, while the genetic and environmental factors influencing mild ID (lowest 3% of IQ distribution) are similar to those influencing IQ in the normal range, factors influencing severe ID (lowest 0.5%) differ from those influencing mild ID or IQ scores in the normal range. Therefore, severe ID is a distinct disorder, qualitatively different from the majority of ID, which in turn represents the low extreme of the normal distribution of intelligence.

Author contributions: A.R., R.P., and P.L. designed research; A.R. performed research; A.R., M.C., A.M., M.T., O.K., M.W., and H.L. analyzed data; M.T., E.F., K.G., M.W., M.D., and H.L. contributed to revisions of the paper; E.F., M.D., and M.W. obtained data, administrative approvals, and revisions of the manuscript; R.P. obtained funding; and A.R., M.C., E.F., K.G., M.D., M.W., H.L., R.P., and P.L. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. D.J.B. is a guest editor invited by the Editorial Board.

Freely available online through the PNAS open access option.

¹To whom correspondence should be addressed. Email: avi.reichenberg@mssm.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1508093112/-DCSupplemental.

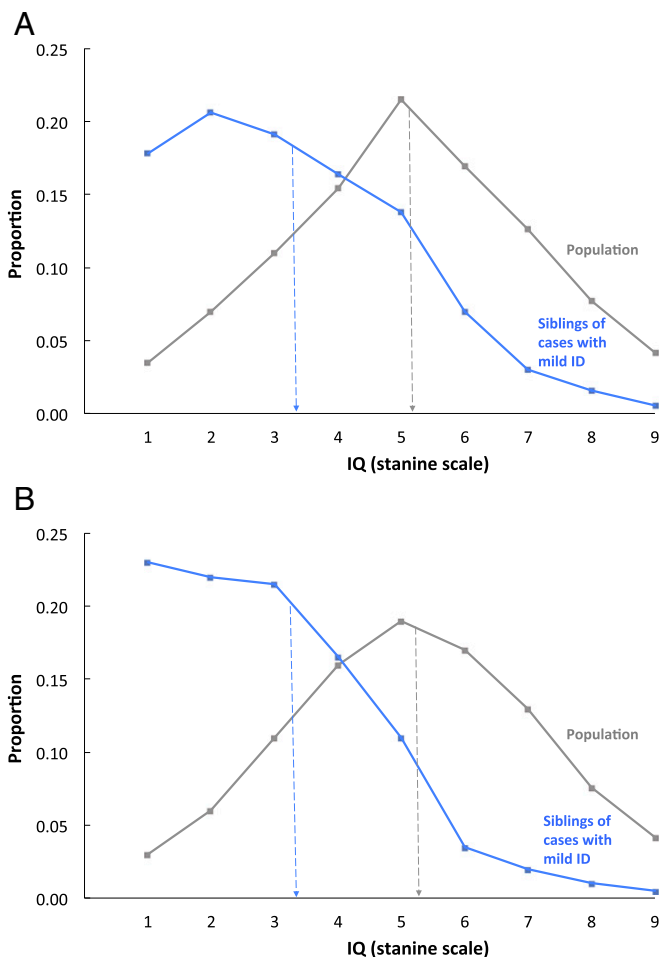


Fig. 1. Mild ID (lowest possible IQ stanine scale score of 1 equals lowest 3% IQ scores) is familial. Siblings of persons with mild ID have significantly and substantially lower IQs than the population. (A) Swedish results for male siblings of persons with IQ stanine score of 1 (mean = 3.31, SD = 1.80, $n = 12,431$ male pairs) and for the entire sibling population (mean = 5.10, SD = 1.95, $n = 382,122$ pairs). (B) Israeli results for male siblings of persons with IQ stanine score of 1 (mean = 3.36, SD = 2.46, $n = 6,800$ male pairs) and for the entire sibling population (mean = 5.49, SD = 1.94, $n = 239,117$ pairs). Note: individuals with severe ID are not included in the stanine IQ score distributions presented.

factors that affect individual differences throughout the normal range of variation in IQ? Indeed, individuals with the lowest IQ could reflect a mixture of those at the lower tail of the population distribution, as well as those whose low performance represents a qualitatively different impairment. This fundamental question has not been directly tested before.

Results

In both countries, results strongly supported the hypothesis that most ID is the lowest end of the normal distribution of IQ, but severe ID is etiologically distinct. As shown in Fig. 1, mild ID (the lowest 3% IQ scores; lowest possible IQ stanine score of 1) is highly familial. Siblings of persons with mild ID have an IQ distribution shifted sharply to the left of the population with a mean (IQ stanine score of 3) halfway between mild ID and the total sibling population average (IQ stanine score of 5).

Application of DF extremes analyses (*Materials and Methods*) to these data showed a significant estimate of 0.43 for Sweden [95% confidence interval (CI): 0.40–0.46] and of 0.47 for Israel (95% CI: 0.41–0.53) for group-differences familiarity for male siblings, indicating that about half of the difference between mild ID and

average IQ is familial in origin. This is called “group-differences familiarity” to distinguish it from the usual estimate of familiarity, which refers to differences among individuals, rather than to mean differences between an extreme group and the population.

Familial effects could be due to genetic or shared environmental factors (i.e., nongenetic factors that make twins similar, such as socioeconomic background or similar in-utero environment). The twin data in Fig. 2 show that this familial effect for mild ID is largely genetic in origin. The results for dizygotic (DZ) co-twins (mean = 2.78) in Fig. 2 are highly similar to those for non-twin siblings in Fig. 1 and much lower than the population [$t(382,243) = 13.19, P < 0.0000001$]. In contrast, monozygotic (MZ) co-twins (mean = 1.96) have much lower IQs than DZ co-twins [$t(200) = 3.88, P = 0.00013$]. In other words, MZ co-twins regress only 0.96 points to the sample mean, whereas same-sex DZ co-twins regress 1.78 points, suggesting that genetics contribute substantially to the mean difference between the probands and the population. Applying the DF extremes analysis (*Materials and Methods*) to these data yields a significant estimate of 0.46 (95% CI: 0.32–0.60) for group-differences heritability, indicating that a large proportion of the mean difference between the probands and the population can be ascribed to genetic factors.

The estimate of group shared environment is 0.30 (95% CI: 0.19–0.41) for same-sex twins, suggesting that about one-third of the mean difference between the probands and the population may be due to environmental factors involved in growing up in the same environment. The remaining 24% (95% CI: 0.18–0.29) of the difference is due to error of measurement and nonshared environmental influences.

Twin analyses of individual differences in the entire sample yielded similar results, which suggests that the factors influencing the lowest 3% of IQ scores are similar to those influencing IQ in the normal range. For the entire sample, intraclass correlations were 0.83 for MZ twins ($n = 3,039$ pairs) and 0.57 for same-sex DZ twins ($n = 3,196$ pairs). Model-fitting analysis (*Materials and Methods*) for the unselected sample of twins yields an individual differences heritability estimate of 0.55 (95% CI: 0.39–0.72) for same-sex twins. In other words, about half of the variance in IQ scores throughout the entire distribution can be attributed to genetic factors. Shared environment accounts for 28% (95% CI:

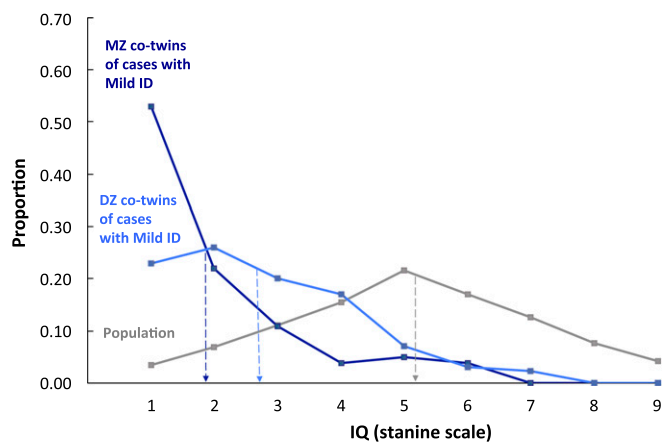


Fig. 2. Mild ID (lowest possible IQ stanine scale score of 1 equals lowest 3% IQ scores) is heritable. MZ co-twins of mild ID cases have significantly and substantially lower IQs (mean = 1.96, SD = 1.37, $n = 79$ pairs) than DZ co-twins (mean = 2.78, SD = 1.52, $n = 123$ pairs), whose IQs are in turn lower than the population (mean = 5.10, SD = 1.95, $n = 382,122$ non-twin pairs). These results are from Sweden only because information on twin zygosity is not available for the Israeli sample. Note: individuals with severe ID are not included in the stanine IQ score distributions presented.

0.13–0.43) of the individual differences in IQ scores, and error of measurement and nonshared environment accounts for an additional 16% (95% CI: 0.13–0.21).

Fig. 3 shows strikingly different results for siblings of individuals with severe ID. In both countries, these individuals had normal IQ scores indistinguishable from the population distribution. Application of DF extremes analyses (*Materials and Methods*) to these data estimated group-differences familiarity to be negligible (upper limit estimate: -0.01 for Sweden and 0.13 for Israel). Comparison with the results above supports the hypothesis that the factors influencing severe ID differ from those influencing mild ID or IQ scores in the normal range.

We examined the effects of sex as a moderator of the group-differences familiarity estimates. The Israeli data (available for both male and females) yielded similar results for male, female, and opposite-sex sibling pairs (Figs. S1 and S2). These findings suggest that sex does not importantly affect estimates of group-differences familiarity. The proportion of men drafted to the military in Sweden fell sharply during the 1990s (following the end of the Cold War). As the draft data may therefore be less representative for recent

cohorts, we conducted a sensitivity analysis stratifying analyses by the year conscription procedure took place (before or after 2000). The results of the group-differences familiarity were similar for those undergoing the conscription procedure before or after the year 2000 (Figs. S3 and S4).

Discussion

In summary, probands with mild ID had siblings with IQ scores that were intermediate between probands and controls. Group-differences familiarity was 43–47% and the group-differences heritability estimate was 46% with a shared environment estimate of 30%. Individual-differences heritability and shared environment estimates were similar: 55% and 28%, respectively. In contrast, probands with severe ID had siblings with normal-average IQ scores, and group-differences familiarity was negligible. To our knowledge this is the first time that significant differences in group-differences familiarity between mild and severe ID have been demonstrated. The results provide the strongest evidence supporting the hypothesis that most ID is caused by the same genetic and environmental influences responsible for the normal distribution of IQ, but that severe ID is not. Using large population-based samples with virtually complete ascertainment protects against bias and enhances generalizability, thus enabling a rigorous test of the hypothesis.

Modern behavioral genetic research and recent molecular genetic studies have shown that common traits are caused by many genes [e.g., educational attainment (4, 11)]. The results presented here also support the position that common disorders are the quantitative extreme of the normal distribution (12). Therefore, similar to IQ, mild ID is caused by many genes of small effect (4). However, the findings of this study highlight a different model for rare, severe disorders. Rare and severe disorders may be etiologically distinct from the rest of the distribution. Results of a recent study on height also support such a hypothesis. Applying both molecular and behavioral genetic methods, it was demonstrated that common genetic variants associated with height in the general population were similarly associated with height at the top and bottom tails of the height distribution. However, height in individuals with extreme short stature (bottom 0.25 percentile) was less determined by common variants (13).

One possible mechanism for noninherited severe ID could involve de novo point mutations. Such mutations can occur in the germ line during embryogenesis or somatically (14). Molecular genetic studies reported associations between rare de novo point mutations and severe ID (5, 15), but more research is needed to prove the causal role of such mutations and what portion of severe ID they account for. Imprinting may also be involved. Imprinting is a form of gene regulation in which gene expression depends on whether the allele was inherited from the male or female parent. When imprinted genes are paternally expressed, the maternal genes are reciprocally silenced and the contrary is true for maternally expressed genes (16). One of the mechanisms for gene silencing is DNA methylation (17). Genetic imprinting has been associated with severe ID (17, 18). Although our understanding of genetic imprinting is nascent, it merits consideration in severe ID.

These hypothesized mechanisms for severe ID are genomic, although not inherited. It is important to keep in mind, however, that environmental factors could also cause severe ID. Several environmental factors have been associated with severe ID. Multiple studies documented hypoxia around birth as potentially causal for severe ID, and the association was confirmed in a recent meta-analysis (19). Maternal infections during pregnancy (e.g., rubella) have been associated with severe ID in the offspring (20), as well as exposure to environmental toxins (e.g., lead) during pregnancy or after birth (21–23). Another group of potential environmental risks comprises conditions in pregnancy and birth (e.g., diabetes, preterm birth, and growth restriction) (24). These have multiple potential etiologies, and their causal role in severe ID requires further research. Conditions in pregnancy and birth have also been associated with IQ and

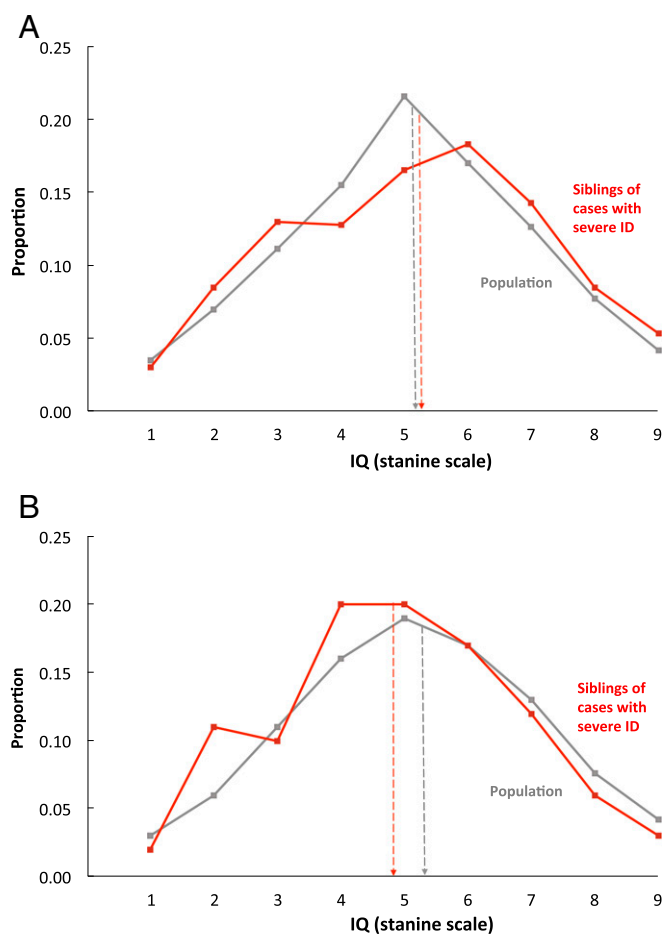


Fig. 3. Severe ID is not familial and thus not affected by the same genetic factors as mild ID or IQ in the normal range. Siblings of persons with severe ID have IQs indistinguishable from the rest of the population. (A) Swedish results: distribution of IQ scores for individuals who have a sibling diagnosed as having severe ID (mean = 5.17, SD = 2.06, $n = 400$ pairs) and the entire sibling population distribution (mean = 5.10, SD = 1.95, $n = 381,122$ pairs). (B) Israeli results: distribution of IQ scores for individuals who have a sibling diagnosed as having severe ID (mean = 4.90, SD = 2.02, $n = 297$ pairs) and the entire sibling population distribution (mean = 5.49, SD = 1.94, $n = 239,117$ pairs). Note: individuals with severe ID are not included in the stanine IQ score distributions presented.

mild ID (25). When considering environmental factors, timing of exposure could be of critical importance. There are periods in brain development during which the nervous system is especially sensitive to certain environmental exposures (26, 27). Detrimental effects of environmental factors may therefore depend on timing and/or dose of exposure. For example, maternal nutrition before or early in pregnancy, but not later in pregnancy was associated with language development (28). Maternal viral infections can interfere with neurodevelopment and the earlier in pregnancy they occur, the more severe their effect. Adverse consequences of maternal rubella were present when infection occurred early, but not in mid-pregnancy (29). In *SI Materials and Methods* and *Table S1* we explored this hypothesis and examined the relationship between birth variables representing environmental events of early and late gestation and mild and severe ID. Although we observed a stronger relationship between the birth variables and severe ID than that between the birth variables and mild ID, the relationship was particularly strong for adverse events happening around birth (*Table S1*), providing preliminary evidence for the potential importance of timing of environmental effects in causing severe ID.

Further research is needed to test the hypotheses about potential mechanisms of severe ID. Here we identified several future directions for research, but interpretation of our data has some limitations. First, we were not able to use the twin data to examine severe ID because the number of twin pairs with severe ID and information on zygosity was very low. Replication in other cohorts is important, particularly in samples including sufficient numbers of twins with severe ID. Larger samples could also help test the specific hypotheses about the role of de novo mutations and environmental factors in severe ID. A germ line de novo mutation is identical for MZ twins, but not for DZ twins or siblings. Therefore, the de novo hypothesis does not negate high MZ twin concordance for severe ID.

Second, data on diagnosis of ID and data on IQ scores come from different sources. The former is based on diagnoses in childhood by health services, whereas the latter is derived from the IQ testing by the military. However, the diagnosis of ID has been shown to be reliable and is carried out by well-trained diagnostic teams.

Third, because the IQ data used were collected for military placement, some individuals may have an incentive to underperform, thus leading to measurement errors that depend on true IQ. Moreover, intentional underperforming may have a shared environmental component, potentially leading to biases. However, the tests used by the military have been extensively validated against standard measures of IQ (30–32) and measures of scholastic ability (33). Furthermore, the assessments by the military are broad and include more than IQ testing alone (34, 35). The whole procedure is aimed at detecting disease, as well as identifying persons suitable for higher military training and prestigious military roles (officer training, service in prestigious combat and Air force units, etc.). Scoring high on the IQ tests is necessary to be eligible for higher military training. Thus, there is a strong incentive to perform well on the IQ tests. Furthermore, military service is mandatory: duration of service does not vary as a function of IQ and low IQ is not a sufficient reason for exclusion from military service. Thus, underperforming is not sufficient for exemption from service. Examining our data does not support the presence of a substantial group of underperformers. If such a group existed, then one would expect to see either a left shift in the entire IQ distribution or a small “bump” on the left side of the IQ distribution, reflecting the effect of a large group of underperformers. Inspection of figures in this article does not provide such evidence. Finally, for intentional underperforming to introduce a substantial bias to the results of this study, underperformance should be conditional on the exposure, i.e., mild or severe ID. Thus, siblings of individuals with mild ID should be motivated to underperform and siblings of individuals with severe ID to overperform. Siblings would have to do this consistently to bias the results across such large samples. Taken together, we strongly believe that intentional underperforming could not substantially bias our results.

We conclude that the most severe form of ID is a distinct disorder, the etiology of which is qualitatively different from mild ID. Despite the well-known single-gene causes of severe ID, most severe ID is caused by idiosyncratic factors not related to inherited influences on intelligence. In contrast, mild ID represents the low genetic extreme of the normal distribution of intelligence, which supports the proposition that common disorders represent the extremes of quantitative traits (12). Thus, most ID is likely to be caused by the same genetic factors responsible for the rest of the distribution of intelligence so genes found to be associated with intelligence will also be associated with ID other than severe ID.

Materials and Methods

This study was approved by the regional ethics committee at the Karolinska Institutet, Stockholm, and by the ethics committee of the Sheba Medical Center, Tel Hashomer, Israel. Because the study used existing databases and analytic datasets were deidentified, it received a waiver of informed consent from human subjects.

Samples.

Swedish cohorts. Data came through linkage of population-based registers in Sweden that use unique personal identification numbers assigned to each Swedish citizen at birth or upon arrival in Sweden as immigrants, enabling accurate linkage.

A total of 3,025,168 men who were eligible for conscription at age 18 to the Swedish military between January 1968 and December 2012 were identified through the Multi-Generation Register (MGR), which includes information on all individuals in Sweden born from January 1932 onward who were alive and residing in Sweden on January 1960 or later (36). Conscription is mandatory for men. Only men with a severe handicap or a chronic disease were exempt (35, 37).

The conscription procedure consists of a series of tests of physical and mental health status and personality and intellectual capacity. The IQ test consists of the following four subtests—verbal comprehension; spatial ability; problem solving, reasoning, and mathematical and general knowledge; and technical and physical problem solving (35, 37)—and was validated against standard measures of intelligence (30). The IQ score was standardized against the entire population to follow a Gaussian distribution between 1 and 9 with a mean of 5 (SD = 2). We obtained data on IQ at age 18 from the conscription database.

We identified 598 men within the cohort who were exempt from conscription with at least one diagnosis of severe intellectual disability [as defined by code 313–314 in the *International Classification of Diseases*, 8th Revision (ICD-8); code 318.1–318.2 in ICD-9; and code F72–F73 in ICD-10) in the National Patient Register in Sweden. The National Patient Register contains diagnostic information on all those with psychiatric hospitalizations since 1973 and outpatient diagnoses since 2001 (for detailed description of reporting procedures to the register, see ref. 38). All infants and preschool children are regularly seen at well-child care clinics and undergo routine medical and developmental screening. In Sweden, all children aged 4 undergo routine general health screening, which includes mandatory developmental assessment (motor, language, cognitive, and social development) conducted by a nurse and pediatrician. Children with any suspected developmental disorder (including intellectual disability) are referred for further assessment by a specialized team in a child psychiatry unit or habilitation service. During the study period, diagnoses were made by diagnostic teams with a psychiatrist, clinical psychologist, and speech pathologist or occupational therapist, depending on clinical manifestations. For a diagnosis of intellectual disability the evaluation is made by a psychologist according to standardized tests with high reliability.

Among eligible men included in the MGR, male full siblings were identified using family relationship information recorded in the registry. To minimize potential differences in the family environment, pairs of male siblings born closest to each other were included. Each family contributed only one sibling pair to the analysis. In total, 381,622 male–male sibling pairs were identified in this manner. A total of 400 sibling pairs were discordant for intellectual disability (i.e., had one sibling with a diagnosis of severe intellectual disability and one without).

Twin pairs were identified as two male siblings recorded as being born on the same day. Zygosity was identified by linkage to the Swedish Twin Registry (39, 40). A full description of zygosity assignments was previously published (40). In total, 3,039 MZ male twin pairs and 3,196 DZ male twin pairs were identified in this manner.

Israeli cohorts. Data came from the Israeli military Draft Board Registry (DBR), which includes information on the unselected Israeli Jewish population of adolescents aged 17 y. The Draft Board mandatory assessment determines intellectual, medical, and psychiatric eligibility for compulsory military service. The population assessed by the draft board therefore includes

individuals who would be eligible for military service, as well as those who will be exempt owing to medical, psychiatric, or social reasons.

A total of 2,178,842 men and women who were assessed by the Israeli military Draft Board between January 1960 and December 2005 were identified through the DBR. IQ tests consist of the following four subtests: an Otis-type verbal intelligence test; a verbal reasoning test; a nonverbal reasoning test similar to the Raven's Progressive Matrices; and a mathematical knowledge test (31, 34, 41). The tests are progressive, beginning with relatively simple questions and becoming more difficult. The tests are group-administered and time-limited. All of the scores are based on the number of correct answers. The IQ score is standardized against population norms and follows a Gaussian distribution between 1 and 9 with a mean of 5.4 ($SD = 1.95$). The sum of the scores for the four tests forms a validated measure of general intelligence (IQ) (31, 34). The correlation between the draft board scale and the Wechsler IQ scale was 0.79 (32).

During the time covered by this study the draft board assigned diagnoses based on the *International Classification of Diseases*, 8th Revision (ICD-8), 9th Revision (ICD-9), and 10th Revision (ICD-10). Psychiatric diagnoses are assigned by a board-certified psychiatrist experienced in treating adolescents. The standard procedure for psychiatric diagnosis includes a face-to-face assessment as described in detail elsewhere (42). For individuals with developmental disabilities (including intellectual disability) the standard procedure is modified. At age 17 y, their medical status is reported in detail to the draft board by government agencies and other organizations responsible for their care and protection. Such reports include the current diagnosis according to contemporary criteria. The original childhood diagnosis and subsequent clinical history up to age 17 y are also commonly reported. The draft board generally assigns a diagnosis based on review of these materials rather than a face-to-face assessment (43).

Among eligible men included in the DBR, full siblings were identified using family relationship information recorded in the registry (44). This information is based on the unique personal identification number assigned to each Israeli citizen at birth or upon immigration. To minimize potential differences in the family environment, pairs of male siblings born closest to each other were included. Each family contributed only one sibling pair to the analysis. In total, 246,214 male–male sibling pairs were identified in this manner. A total of 297 sibling pairs were discordant for intellectual disability (i.e., had one sibling with a diagnosis of severe intellectual disability and one without).

A total of 120,973 female–female sibling pairs and 250,301 opposite-sex sibling pairs were similarly identified using the family relationship information. The proportion of female–female pairs is lower because orthodox Jewish women (about 25% of women) are exempt from military induction and data on females assessed by the Draft Board is maintained for only 25 y.

Statistical Analysis.

Descriptive statistics. Descriptive statistics, including distribution of IQ scores and means and SDs, were computed. Differences between unaffected cosiblings and control siblings were compared using Analysis of Covariance models adjusting for year of birth. A sibling's individual-differences familiarity was computed using an intraclass correlation coefficient. *P* values were set at 0.05 (two-tailed). **Individual and group familiarity.** Quantitative genetics offers an approach to the issue of the relation between the etiology of the mean of selected groups (such as low IQ) and the etiology of individual differences throughout the population. This approach, developed by DeFries and colleagues for samples of twins and adapted for the study of siblings (10, 45), can be used to investigate the extent to which low cognitive abilities are the etiological extreme of the normal distribution of cognitive ability (for a summary see Box S1).

The conceptual framework for the approach is illustrated in Fig. S5 in relation to siblings and twins [for a detailed description see Plomin (46)]. Low-ability probands will fall toward the extreme of the distribution. The mean of the sibling of the proband will regress to the unselected population mean to the extent that familial factors are unimportant to the etiology of the disorder. In contrast, if familial factors are important, the mean of the siblings of the probands will be lower than the population mean. The use of quantitative measures enables estimation of the magnitude of familial resemblance. The extent to which the mean difference between the siblings and the population is similar to the difference between the probands and the population provides a quantitative estimate of the extent to which the mean difference between the probands and the population is due to familial factors (46). This estimate is called "group familiarity."

The key is whether the magnitude of group familiarity for the low-ability group is different from traditional estimates of familiarity of individual differences in the population ("individual-differences familiarity"). If group familiarity for the low-ability group differs from individual-differences familiarity, this suggests that low ability is etiologically different from individual differences in the population. In contrast, if group familiarity for the low-ability group is similar to individual-differences familiarity, this suggests that low ability may represent the low end of the normal distribution of genetic and environmental

influences on individual difference in ability (47). The latter (group familiarity equals individual familiarity) is not as clearly interpretable as the former (group familiarity does not equal individual familiarity). When group familiarity differs from individual-differences familiarity, there seems to be no explanation other than that the condition is etiologically different from variation for the rest of the distribution. In contrast, finding that group familiarity equals individual-differences familiarity does not prove that the etiology of the conditions and the normal distribution are the same. The condition and the distribution may yield the same magnitude of familiarity even though the factors underlying the familiarity differ (46).

Although the IQ of probands with severe ID is not reported, we conservatively estimated the upper limit of group familiarity by assuming an average IQ in probands with severe ID to be equal to a stanine score of 1.

DF extremes analysis. Twins can be used to approach the issue of the relation between the etiology of the mean of selected groups (such as low IQ) and the etiology of individual differences throughout the population (i.e., if the resemblance in familiarity in such sibling analyses is due to shared heredity or environment). The DF extreme analysis uses twins (10) to investigate the extent to which low-cognitive abilities are etiologically the extreme of the normal distribution of cognitive ability. The DF analysis uses the quantitative parameter group heritability (h^2_g), which represents the extent to which a mean difference on a quantitative measure between a selected group and the unselected population is heritable (i.e., is due to genetic differences) (46). The key is whether the magnitude of group heritability for the low-ability group is different from traditional estimates of heritability of individual differences in the population (individual heritability).

DF analysis examines differential regression to the mean, for example, for identical (MZ) and fraternal (DZ) twins. IQ scores of co-twins of low-ability probands are expected to regress toward the mean of the unselected population (47). However, to the extent that low ability is due to genetic factors, the regression to the mean will be smaller for the co-twins from MZ twin pairs than for DZ co-twin pairs (illustrated in Fig. S5). The reason for this is because the amount of regression toward the mean is a function of the magnitude of the similarity between the two variables (here the proband and co-twin IQ scores). The higher the similarity, the less regression to the mean. If a trait is genetically influenced, the similarity for MZ twins will be greater than for DZ twins. Hence, DZ co-twins will regress further toward the mean of the unselected population.

The DF analysis employs a multiple regression model that explicitly tests the differential regression to the mean for MZ and DZ co-twins and provides an estimate of group heritability (47). The basic DF model is represented as the regression $C = B_1P + B_2R + A$, in which the co-twin's IQ score (*C*) is predicted from the proband's IQ score (*P*) and the coefficient of relatedness (*R*), which is 1.0 for MZ and 0.5 for DZ pairs. Because the proband mean is transformed to a mean of 1.0 and the unselected population to a mean of 0.0, the mean of the co-twin score for MZ and DZ twins estimates their group-differences familiarity (total familial similarity). The regression weight B_2 estimates group-differences heritability, the differential regression to the population mean for MZ and DZ twins. Group shared environment or twin resemblance not explained by genetic factors can be estimated by subtracting group-differences heritability from MZ group-differences familiarity (10, 48).

Individual-differences model fitting for twins. We applied the standard maximum-likelihood model-fitting analysis for the classical twin design to variance/covariance matrices for same-sex twins as described by Neale and Cardon (49). The ACE model, which estimates parameters for additive genetic variance (*A*), common or shared environment (*C*), and environmental influences that are not shared (*E*), assumes that genetic effects are additive and that MZ and DZ twins experience equally similar environments.

Although the twin method is robust (50), it has its limitations. Some of these limitations may lead to underestimates rather than overestimates of heritability because, if true, they make MZ twins less similar than they would otherwise have been. For example, it has been argued that the atypical gestation of MZ twins causes increased rates of disorders (51), although other studies indicate that this is not the case (52). Other problems could go in either direction. For example, two-thirds of MZ twins share the same chorion (outer fetal sack), which can lead to shared influences such as infection, shared vasculature, and other anomalies of sharing a crowded chorion (50). Some problems might inflate heritability estimates, most notably the possibility that MZ twins share more similar postnatal environments than DZ twins, although it seems that this is not usually the cause of their greater phenotypic similarity, but rather the consequence of their genetic identity (48).

ACKNOWLEDGMENTS. R.P. is supported by a Medical Research Council Research Professorship Award (G19/2) and a European Research Council Advanced Investigator Award (295366). A.R. was supported, in part, by

the National Institute for Health Research Specialist Biomedical Research Centre for Mental Health Award to the South London and Maudsley

National Health Service Foundation Trust and the Institute of Psychiatry, Psychology, and Neuroscience, King's College London.

- Centers for Disease Control and Prevention (CDC) (2004) Economic costs associated with mental retardation, cerebral palsy, hearing loss, and vision impairment: United States, 2003. *MMWR Morb Mortal Wkly Rep* 53(3):57–59.
- American Psychiatric Association and American Psychiatric Association. Task Force on DSM-IV (1994) Diagnostic and Statistical Manual of Mental Disorders: DSM-IV (American Psychiatric Association, Washington, DC), 4th Ed, p 886.
- Roeleveld N, Zielhuis GA, Gabreëls F (1997) The prevalence of mental retardation: A critical review of recent literature. *Dev Med Child Neurol* 39(2):125–132.
- Plomin R, DeFries JC, Knopik VS, Neiderhiser JM (2012) *Behavioral Genetics* (Worth Publishers, New York), 6th Ed.
- de Ligt J, et al. (2012) Diagnostic exome sequencing in persons with severe intellectual disability. *N Engl J Med* 367(20):1921–1929.
- Ellison JW, Rosenfeld JA, Shaffer LG (2013) Genetic basis of intellectual disability. *Annu Rev Med* 64:441–450.
- Veltman JA, Brunner HG (2012) De novo mutations in human genetic disease. *Nat Rev Genet* 13(8):565–575.
- Penrose LS (1938) A Clinical and Genetic Study of 1280 Cases of Mental Defect (H. M. Stationery, London), p 159.
- World Health Organization (1992) International Statistical Classification of Diseases and Related Health Problems (World Health Organization, Geneva), 10th Ed.
- DeFries JC, Fulker DW (1988) Multiple regression analysis of twin data: Etiology of deviant scores versus individual differences. *Acta Genet Med Gemellol (Roma)* 37(3-4): 205–216.
- Rietveld CA, et al.; LifeLines Cohort Study (2013) GWAS of 126,559 individuals identifies genetic variants associated with educational attainment. *Science* 340(6139): 1467–1471.
- Plomin R, Haworth CM, Davis OS (2009) Common disorders are quantitative traits. *Nat Rev Genet* 10(12):872–878.
- Chan Y, et al. (2011) Common variants show predicted polygenic effects on height in the tails of the distribution, except in extremely short individuals. *PLoS Genet* 7(12): e1002439.
- Vadlamudi L, et al. (2010) Timing of de novo mutagenesis: A twin study of sodium-channel mutations. *N Engl J Med* 363(14):1335–1340.
- Hamdan FF, et al.; Synapse to Disease Group (2009) Mutations in SYNGAP1 in autosomal nonsyndromic mental retardation. *N Engl J Med* 360(6):599–605.
- Isles AR, Wilkinson LS (2000) Imprinted genes, cognition and behaviour. *Trends Cogn Sci* 4(8):309–318.
- Reichenberg A, Mill J, MacCabe JH (2009) Epigenetics, genomic mutations and cognitive function. *Cogn Neuropsychiatry* 14(4-5):377–390.
- Nicholls RD, Knoll JH, Butler MG, Karam S, Lalonde M (1989) Genetic imprinting suggested by maternal heterodisomy in nondeletion Prader-Willi syndrome. *Nature* 342(6247):281–285.
- Modabbernia A, Mollon J, Boffetta P, Reichenberg A (2015) Impaired gas exchange at birth and risk of intellectual disability and autism: A meta-analysis. *J Autism Dev Disord*, in press.
- Rantakallio P, von Wendt L (1985) Risk factors for mental retardation. *Arch Dis Child* 60(10):946–952.
- McDermott S, et al. (2014) Are different soil metals near the homes of pregnant women associated with mild and severe intellectual disability in children? *Dev Med Child Neurol* 56(9):888–897.
- Beattie AD, et al. (1975) Role of chronic low-level lead exposure in the aetiology of mental retardation. *Lancet* 1(7907):589–592.
- World Health Organization (2010) *Childhood Lead Poisoning* (World Health Organization, Geneva).
- Cans C, et al. (1999) Aetiological findings and associated factors in children with severe mental retardation. *Dev Med Child Neurol* 41(4):233–239.
- Sørensen HT, et al. (1997) Birth weight and cognitive function in young adult life: Historical cohort study. *BMJ* 315(7105):401–403.
- Mayberry RI, Lock E, Kazmi H (2002) Linguistic ability and early language exposure. *Nature* 417(6884):38.
- Brainard MS, Knudsen EI (1998) Sensitive periods for visual calibration of the auditory space map in the barn owl optic tectum. *J Neurosci* 18(10):3929–3942.
- Roth C, et al. (2011) Folic acid supplements in pregnancy and severe language delay in children. *JAMA* 306(14):1566–1573.
- Miller E, Craddock-Watson JE, Pollock TM (1982) Consequences of confirmed maternal rubella at successive stages of pregnancy. *Lancet* 2(8302):781–784.
- Carlstedt B, Mårdberg B (1993) Construct validity of the Swedish Enlistment Battery. *Scand J Psychol* 34:353–362.
- Reichenberg A, et al. (2005) Elaboration on premorbid intellectual performance in schizophrenia: Premorbid intellectual decline and risk for schizophrenia. *Arch Gen Psychiatry* 62(12):1297–1304.
- Brill N (2005) Recall bias in schizophrenia. PhD thesis (Bar-Ilan University, Ramat-Gan, Israel).
- Carlstedt B, Gustafsson JE (2005) Construct validation of the Swedish Scholastic Aptitude Test by means of the Swedish Enlistment Battery. *Scand J Psychol* 46(1):31–42.
- Davidson M, et al. (1999) Behavioral and intellectual markers for schizophrenia in apparently healthy male adolescents. *Am J Psychiatry* 156(9):1328–1335.
- David AS, Malmberg A, Brandt L, Allebeck P, Lewis G (1997) IQ and risk for schizophrenia: A population-based cohort study. *Psychol Med* 27(6):1311–1323.
- Ekbohm A (2011) The Swedish multi-generation register. *Methods Mol Biol* 675: 215–220.
- Nyberg J, et al. (2014) Cardiovascular and cognitive fitness at age 18 and risk of early-onset dementia. *Brain* 137(Pt 5):1514–1523.
- Sandin S, Nygren KG, Iliadou A, Hultman CM, Reichenberg A (2013) Autism and mental retardation among offspring born after in vitro fertilization. *JAMA* 310(1): 75–84.
- Lichtenstein P, et al. (2006) The Swedish Twin Registry in the third millennium: An update. *Twin Res Hum Genet* 9(6):875–882.
- Silventoinen K, Magnusson PK, Tynelius P, Kaprio J, Rasmussen F (2008) Heritability of body size and muscle strength in young adulthood: A study of one million Swedish men. *Genet Epidemiol* 32(4):341–349.
- Rabinowitz J, et al. (2000) Cognitive and behavioural functioning in men with schizophrenia both before and shortly after first admission to hospital. Cross-sectional analysis. *Br J Psychiatry* 177:26–32.
- Weiser M, et al. (2001) Association between nonpsychotic psychiatric diagnoses in adolescent males and subsequent onset of schizophrenia. *Arch Gen Psychiatry* 58(10): 959–964.
- Reichenberg A, et al. (2006) Advancing paternal age and autism. *Arch Gen Psychiatry* 63(9):1026–1032.
- Weiser M, et al. (2008) Subtle cognitive dysfunction in nonaffected siblings of individuals affected by nonpsychotic disorders. *Biol Psychiatry* 63(6):602–608.
- DeFries JC, Fulker DW (1985) Multiple regression analysis of twin data. *Behav Genet* 15(5):467–473.
- Plomin R (1991) Genetic risk and psychosocial disorders: Links between the normal and abnormal. *Biological Risk Factors for Psychosocial Disorders*, eds Rutter M, Caser P (Cambridge University Press, Cambridge, UK), pp 101–138.
- Saudino KS, Plomin R, Pedersen NL, McClearn GE (1994) The etiology of high and low cognitive ability during the second half of the life span. *Intelligence* 19(3):359–371.
- Dale PS, et al. (1998) Genetic influence on language delay in two-year-old children. *Nat Neurosci* 1(4):324–328.
- Neale MC, Cardon LR (1992) *Methodology for Genetic Studies of Twins and Families* (Kluwer Academic Publishers, Dordrecht, The Netherlands).
- Martin N, Boomsma D, Machin G (1997) A twin-pronged attack on complex traits. *Nat Genet* 17(4):387–392.
- Phillips DI (1993) Twin studies in medical research: Can they tell us whether diseases are genetically determined? *Lancet* 341(8851):1008–1009.
- Christensen K, Vaupel JW, Holm NV, Yashin AI (1995) Mortality among twins after age 6: Fetal origins hypothesis versus twin method. *BMJ* 310(6977):432–436.
- Plomin R, Kovas Y (2005) Generalist genes and learning disabilities. *Psychol Bull* 131(4):592–617.