Available online at www.sciencedirect.com
ScienceDirect
Learning and Individual Differences 18 (2008) 76–88
Learning and Individual Differences
www.elsevier.com/locate/lindif

A twin-family study of general IQ

Marieke van Leeuwen, Stéphanie M. van den Berg *, Dorret I. Boomsma

Department of Biological Psychology, Vrije Universiteit Amsterdam, Van der Boechorststraat 1, 1081 BT, Amsterdam, The Netherlands

Received 5 December 2006; received in revised form 29 March 2007; accepted 6 April 2007

Abstract

In this paper we assess the presence of assortative mating, gene–environment interaction and the heritability of intelligence in childhood using a twin family design with twins, their siblings and parents from 112 families. We evaluate two competing hypotheses about the cause of assortative mating in intelligence: social homogamy and phenotypic assortment, and their implications for the heritability estimate of intelligence. The Raven Progressive Matrices test was used to assess general intelligence (IQ) and a persons IQ was estimated using a Rasch model. There was a substantial correlation between spouses for IQ (r=.33) and resemblance in identical twins was higher than in first-degree relatives (parents and offspring, fraternal twins and siblings). A model assuming phenotypic assortment fitted the data better than a model assuming social homogamy. The main influence on IQ variation was genetic. Controlled for scale unreliability, additive genetic effects accounted for 67% of the population variance. There was no evidence for cultural transmission between generations. The results suggested that an additional 9% of observed IQ test variation was due to gene–environment interaction, with environment being more important in children with a genetic predisposition for low intelligence.

Keywords: Intelligence; Genetics; Assortative mating; Heritability

1. Introduction

Individual differences in intelligence tend to cluster in families (for reviews see e.g. Bouchard, & McGue, 2003, 1981; Boomsma, 1993; Deary, Spinath, & Bates, 2006). The resemblance between relatives can be due to genetic relatedness, environmental similarities, cultural transmission from one generation to the next, social interactions between family members, or a combination of these mechanisms. When one wants to study causes of this resemblance and only first-degree relatives, such as parents and their offspring, or siblings are included in the study design, it is not possible to disentangle shared genetic from shared environmental effects. However, in a twin design such a distinction can be made, because monozygotic (MZ; identical) twins share all, or nearly all of their DNA, while dizygotic twins (DZ; fraternal) share on average 50% of their segregating genes (Boomsma, Busjahn, & Peltonen, 2002; Plomin, DeFries, Craig, & McGuffin, 2002). A larger resemblance of MZ than of DZ twins therefore is suggestive of genetic influences on twin resemblance.

Numerous studies of young and adult twins have explored the etiology of resemblance in intelligence between family members. Twin studies in children estimate the contribution of genetic effects to the variability in intelligence at 25% to 50%. Part of the remaining variance is due to environmental factors shared by children who grow up in the same family (Bartels, Rietveld, Van Baal, & Boomsma, 2002; Plomin et al., 2002; Turkheimer, Haley, Waldron, D'Onofrio, & Gottesman, 2003). Heritability appears to increase with age and the influence of shared environment disappears in early adolescence (see e.g. Bouchard & McGue, 2003; Cherny & Cardon, 1994; Plomin et al., 2002; Plomin and Spinath, 2004; Plomin, Pedersen, Lichtenstein, & McClearn, 1994; Posthuma, De Geus, & Boomsma, 2001; Scarr & Weinberg, 1983).

The classical twin design in which data from MZ pairs are compared to data from DZ pairs relies on several assumptions. It is often assumed that the phenotypes of parents are uncorrelated

* Corresponding author. Tel.: +31 30 253 9782.
E-mail address: S.M.vandenBerg@uu.nl (S.M. van den Berg).

1041-6080/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.lindif.2007.04.006

(i.e., one assumes random mating), that there is no genotype–environment (GE) *interaction*, and no GE *correlation*. GE *interaction* refers to the phenomenon that the influence of a particular genotype may depend on the environment (or vice versa: that the influence of the environment depends on genotype). GE *correlation* refers to the non-random distribution of genotypes over environments and may for instance occur when parents transmit not only their genes but also their environment to their children.

When these assumptions are not met, results from twin studies may be biased. For example, if there is genotype by *common* environment interaction then heritability will be overestimated (Lynch & Walsh, 1998). If there is genotype by *unique* environment interaction then, heritability will be underestimated. If random mating between parents is assumed, while there is non-random mating in the population, this will bias heritability estimates downwards and increases estimates of shared environmental influences. If the classical twin design is extended by including the twins' parents (Fulker, 1982) assortative mating and some forms of GE correlation can be assessed. Additionally, by looking at the association between MZ sum and difference scores, it is possible to detect and estimate GE interaction. In this paper we use such an extended twin design which includes MZ and DZ twins, their siblings and their parents, to study to what extent assortative mating, cultural inheritance and GE interaction and correlation are present for general intelligence (IQ). Data on general IQ were collected in both generations with the Raven Progressive Matrices test.

1.1. Spousal resemblance

Spouse and family studies show that spouses resemble each other in IQ scores and traits correlated to IQ, such as educational attainment. Spousal correlations for performance on the Raven Progressive Matrices are around 0.30 (Guttman, 1974; Watkins & Meredith, 1981). For the Wechsler Adult Intelligence Scale (WAIS), spousal correlations are between 0.37 and 0.61 (Mascie-Taylor, 1989; Watson et al., 2004; Williams, 1975). In the Colorado Adoption Project (CAP), the correlation between spouses on the first unrotated principal component derived from a battery of 13 cognitive tests was 0.11, and in the Hawaii Family Study of Cognition (HFSC) this correlation was 0.20 (Phillips, Fulker, Carey, & Nagoshi, 1988). These and other studies clearly show a resemblance in intelligence between spouses. This resemblance, or non-random mating, may be due to marital interaction, phenotypic assortment, or social homogamy.

The hypothesis of marital interaction or convergence states that spousal correlations arise because spouses spend time together. Spouses would tend to become more similar the longer they are together, because they either influence each other or because they share similar experiences. The few studies that tested this hypothesis found no indications of convergence for intelligence (Gilger, 1991; Mascie-Taylor, 1989; Watson et al., 2004).

Under phenotypic assortment it is assumed that spouses choose each other based on observable characteristics (Reynolds, Baker, & Pedersen, 1996), in this case based on intelligence or a trait related to it: individuals would tend to mate with partners with an intelligence level resembling their own. Most models of assortative mating assume phenotypic assortment (Fulker, 1982; Fulker & DeFries, 1983; Rice, Carey, Fulker, & DeFries, 1989; Wadsworth, DeFries, Fulker, & Plomin, 1995). Mascie-Taylor and Vandenberg (1988) tried to estimate the role of personal preference in mate selection by correcting for variables representing proximity such as social class, locality, family size, birth order and type and years of education. After correcting for these variables, there still was a significant correlation between spouses' IQs, suggesting that that this correlation could be ascribed to direct phenotypic assortment.

Social homogamy refers to assortment based on solely environmental similarities. Spousal phenotypes become correlated because spouses meet each other within a particular environment (Reynolds et al., 1996). In the case of intelligence, the social homogamy hypothesis states that people with the same intelligence level live in the same social environment. Within a particular social environment, partners do not choose each other on the basis of intelligence, but since they live in the same environment, they tend to mate with people with a similar IQ. Spousal correlations in the general population occur when social environment is correlated with intelligence.

When there is mate resemblance for intelligence, it may be necessary to include its effects in the genetic model. When resemblance is caused by phenotypic assortment, this induces genetic similarity between parents, which affects the genetic similarity between parents and offspring and among siblings and dizygotic twins. Under random mating, genetic effects are uncorrelated in parents. The correlation between a parental and a child's genotype, and among siblings and in DZ twins is then 1/2. If there is positive phenotypic assortment, these genetic correlations increase and heritability will be underestimated if this effect is not taken into account (Cavalli-Sforza & Bodmer, 1971). When spousal resemblance is purely due to environmental effects that are not correlated with genetic effects, there are no genetic consequences.

1.2. GE interaction

One approach to detect and estimate GE interaction is by looking at the association between MZ intrapair sum (or average) and difference scores (Jinks & Fulker, 1970). Genetic and shared environmental effects add to the similarity of MZ pairs and unique environment to the differences between MZ pairs. When there is a positive correlation between intrapair sum and absolute differences, less intelligent individuals are more similar than more intelligent individuals, and thus more intelligent people are more susceptible to unique environmental influences (Finkel & Pedersen, 2001).

Jinks and Fulker reported intrapair sum/ intrapair difference correlations for IQ of −.10 and −.13, based on data from 19 MZ twin pairs. Jensen (1970) reported a correlation for IQ of −.15 in MZ twins reared apart. And Finkel and Pedersen (2001) reported a correlation of −.11 in MZ twins reared apart and of −.09 in MZ twins reared together. Although these correlations were all non-significant, all correlations were of similar

magnitude and negative, suggesting that the environment might have a greater influence in less intelligent people.

1.3. Parent–offspring resemblance

Including parents in a twin design adds extra information about the origins of individual differences. The resemblance between parents and offspring may reflect genetic transmission, cultural transmission, or both. In the case of genetic transmission, resemblance between parents and offspring is caused by the genes which are transmitted from the parents to their children. In an ordinary family design genetic transmission is confounded with cultural influences of parents on their offspring. Cultural transmission will increase parent–offspring correlations, as well as correlations between siblings and twins who grow up in the same home environment. In the classical twin design, cultural transmission will show up as shared (or common) environmental variance.

Parents may create a particular kind of environment that is correlated with their genotype or their phenotype, for example, bright parents might stimulate their children with schoolwork or provide them with more intelligence-boosting toys. Whenever there is cultural transmission in the presence of genetic transmission, environmental influences become correlated with genetic influences.

In an adoption design, genetic and cultural transmission can be disentangled because then the adopted child's environment is uncorrelated with the intelligence levels of its biological parents. In the CAP study mentioned earlier, IQ data from adopted and non-adopted children were collected at ages 1, 2, 3, 4, 7, 9, 10 and 12 years. When analyzing the IQ data from the adoptive and non-adoptive children up until the age of 12, no significant shared environmental influence was found: all variance could be explained by additive genetic factors and environmental factors that are not shared by children raised in the same family (Bishop, Cherney, & Hewitt, 2003).

The CAP study also collected data on the biological and, if they were adopted, the adoptive parents of these children. Significant genetic transmission for intelligence was found at all ages (Fulker & DeFries, 1983; Humphreys & Davey, 1988; Rice et al., 1989). The CAP data also showed significant cultural transmission from foster parent to offspring but only before the age of 4 years (Fulker & DeFries, 1983; Humphreys & Davey, 1988; Rice et al., 1989). Alarcón, Plomin, Corley, and DeFries (2003) also showed that there was no cultural transmission for specific cognitive abilities at ages 7 and 12: assuming phenotypic assortment, all variance was due to additive genetic effects and random environmental effects. Similar findings were reported by another adoption study (Scarr & Weinberg, 1983) showing that in adolescence, the impact of the family environment on IQ disappears.

1.4. The present study

Up until now only the CAP-study (Alarcón et al., 2003; Fulker & DeFries, 1983; Humphreys & Davey, 1988; Rice et al., 1989) examined the genetic and environmental transmission of intelligence from parents to their children in the presence of spousal resemblance. Other studies using twins sometimes take assortative mating into account when interpreting their results (e.g. Wainwright, Wright, Geffen, Luciano, & Martin, 2005), but do not assess or model assortative mating directly. In the CAP study different measures of IQ were used across generations. Parental intelligence was estimated based on an unstandardised measure of IQ (see above) whereas in the children intelligence was measured, depending on age, using the Bayley Mental Development Index, the Stanford–Binet Intelligence Scale or the Wechsler Intelligence Scale for Children. The measure used in adults resulted in a relatively low spousal correlation when compared to studies using full scale IQ tests (Mascie-Taylor, 1989; Watson et al., 2004; Williams, 1975). There are studies using comparable IQ tests in parents and children but these studies do not report heritability estimates, since the samples studied were not suitable for this purpose (Guttman, 1974; Guttman & Shoham, 1983; Williams, 1975).

In the present study, we collected data on intelligence using Raven's Progressive Matrices (Raven, 1960; Raven, Raven, & Court, 1998), in MZ and DZ twins, one of their siblings and both of their parents. With this design, cultural and genetic transmission can be studied while taking into account spousal resemblance. The inclusion of additional siblings increases the power to detect additive and non-additive genetic effects (Keller, Coventry, Heath, & Martin, 2005; Posthuma and Boomsma, 2000). Raven IQ measures were estimated based on a Rasch model (Rasch, 1966). This way the intelligence measure is not dependent on the particular items that are included in the test and has no a priori distribution.

We expect that additive genetic effects will explain a large part of the individual differences in IQ. We also explore the presence of non-additive genetic influences, or genetic dominance, on IQ variation. Genetic non-additivity has been suggested in studies on inbreeding (Agrawal, Sinha, & Jensen, 1984; Bashi, 1977), reflecting recessive effects of rare alleles that might not contribute much to the variation in the general population. Genetic dominance has at times been suggested in twin and other studies (Chipuer, Rovine, & Plomin, 1990; Fulker, 1979; Jinks & Fulker, 1970). Dominance effects can be masked by assortative mating and cultural transmission in studies with only MZ and DZ twins.

We fitted two models, one assuming phenotypic assortment and one assuming social homogamy to determine which of both model fits the data best. To assess GE interaction, we tested whether there is an association between absolute difference scores in MZ twins (reflecting non-shared environmental effects) and average scores (reflecting familial effects).

2. Materials and methods

2.1. Participants

The study was approved by the Central Committee on Research involving Human Subjects (CCMO). Twins were recruited from the Netherlands Twin Registry (NTR), established by the Department of Biological Psychology at the Vrije

Universiteit (VU) in Amsterdam (Boomsma, Orlebeke, & Van Baal, 1992; Boomsma et al., 2002, 2006). Twin families with an extra sibling between 9 and 14 years were selected from two birth cohorts (1995–1996). Because the twins and siblings also took part in an MRI study, there were several exclusion criteria such as a pacemaker and metal materials in the head. Families with children with a major medical history, psychiatric problems (as reported by the parents), participation in special education, or physical or sensory disabilities were also excluded. A total of 214 families were invited by letter, which was sent out one to two months before the ninth birthday of the twins. Two weeks after receiving the letter, the families were contacted by phone. Of these families 52% (112) agreed to participate. There was no significant difference between the educational level of mothers who did participate and who did not participate in the study ($F(1,195)=.68$, $p=.41$). Of the 112 families, 103 had full siblings who wanted to participate. Parents signed informed consent forms for the children and themselves. Children also signed their own consent forms. Parents were compensated for their travel expenses and children received a present.

The 112 families came from all over the Netherlands. Mean age of the twins at time of cognitive assessment was 9.1 years, ranging from 8.9 to 9.5 years. There were 23 MZ male, 23 DZ male, 25 MZ female, 21 DZ female and 20 DZ pairs of opposite sex. Zygosity was based on DNA polymorphisms and questionnaire items. Mean age of the sibs ($N=103$; 59 female) was 11.9 years ranging from 9.9 to 14.9. The mean age of the biological fathers was 43.7 ($N=94$, SD=3.7 years), and of the biological mothers 41.9 ($N=95$, SD=3.4 years).

2.2. Testing procedures

This study collected cognitive, behavioral and hormonal data, pubertal status and structural Magnetic Resonance Imaging (MRI) brain data. Data collection took place on two different days. Cheek swabs, for DNA isolation, were collected at home by parents and children. For cognitive testing, families arrived between nine and eleven o'clock in the morning. Children were tested in separate rooms with a cognitive test battery including the Raven's Standard Progressive Matrices (SPM; Raven, 1960). Parents completed the Raven Advanced Progressive Matrices (APM; Raven et al., 1998). The whole protocol took approximately 5 h, including two short breaks and one longer lunch break.

2.3. Materials

Children were individually tested with the Standard Progressive Matrices (Raven, 1960), which they completed at their own pace after verbal instruction. The test consists of 60 problems divided into five sets of twelve. In each set the first problem is as nearly as possible self-evident. The problems within a set become progressively more difficult. The test is intended to cover the whole range of intellectual development from the time a child is able to grasp the idea of finding a missing piece to complete a pattern, and to be sufficiently long to assess a child's maximum capacity to form comparisons and reason by analogy. The test provides an index of general intelligence. For children retest reliability is .88 (Raven, 1960).

Parents were given the Advanced Progressive Matrices (Raven et al., 1998), since the SPM is too easy for most adults. They received written instructions and made the test at their own pace. The APM is comparable to the SPM with the main difference being the level of difficulty. The APM consists of two sets. The first set contains twelve practice items, to familiarize Ss with the test. The second set consists of 36 items, which are identical in presentation and argument with those in Set I. They only increase in difficulty more steadily and become considerably more complex. Reported retest reliability for adults is .91 (Raven et al., 1998).

2.4. Zygosity determination

In 110 twin pairs, zygosity was determined at the VU Medical Centre with eight highly polymorphic di-, tri- and tetranucleotide genetic markers. The zygosity testing included a multiplex PCR of markers D2S125, D8S1130, D1S1609, D5S816 and a second multiplex reaction of markers 15 ActC, D21S1437, D7S2846, and D10S1423. These two multiplex PCR reactions were performed by the protocol provided in the website of the Marshfield Institute (www.marshmed.org/genetics). Results of the zygosity test were sent to the parents. In the remaining two twin pairs zygosity was based on questionnaire items (Rietveld et al., 2000).

2.5. Statistical analysis

2.5.1. Rasch scores

IQ measures in parents and offspring were estimated based on the Rasch model (Rasch, 1966). In this model, every person is represented by a person parameter θ that reflects that person's ability. Every test item is represented by a difficulty parameter β. The probability that a person j answers item i correctly is parameterized by the logistic function $p\ (Y_{ij}=1)=\Psi\ (\theta_j-\beta_i)$, where θ_j is the person parameter, β_i is the difficulty parameter for that particular item, and $\Psi(x)=\exp(x)/(1+\exp(x))$ see also (Van den Berg, Glas, & Boomsma, 2007). Thus, for example, the probability that person j with ability θ_j answers item i with difficulty level β_i correctly, equals $e^{\theta j-\beta i}/[1+e^{\theta j-\beta i}]$. When $\theta_j-\beta_i=0$, the probability of a correct answer is exactly 50%, as $e^0=1$. When ability dominates the difficulty, $\theta_j>\beta_i$, then the probability is higher than 50%, becoming 100% when ability is infinitely higher than the difficulty. When ability is lower than the difficulty of the item, $\theta_j<\beta_i$, then the probability of a correct answer is lower than 50%, becoming 0% when the ability is infinitely lower than the difficulty. Note that the values for θ and β, the ability of a person and the difficulty of an item, are on the same scale.

The rationale for the Rasch model can be presented by analogy to the success of an athletic hurdle jumper: some people jump higher than others do. For each jumper there might exist a hurdle with a certain height where only 50% of the attempts is successful. If the hurdle's height increases, the probability of a successful jump decreases whereas it increases when the

hurdle's height decreases. If a hurdle is very low, the probability of a successful jump approaches one; when the hurdle is very high, the probability of a successful jump approaches zero. In the Rasch model, a person's ability is defined as the difficulty level where the probability of a correct answer (or jump) is 50%. The model assumes local independence. This means that the probability of a success is entirely explained by the θ and β parameter: given θ and β, the probability of a correct answer is not dependent on whether other items are answered correctly or whether other people answered the same item correctly. This assumption is for example also used in the common factor model, where only one factor explains all correlations among the indicator variables (Spearman, 1927). Thus, an assumption in the Rasch model is unidimensionality of ability.

The Rasch model has a number of nice properties. The most important is the property of invariant comparison or separability of person and item parameters: the comparison between two persons is independent of the particular measurement instrument and other persons being measured at the same time. The estimated difference in ability measures between two persons is the same regardless whether we use all items from a test or any possible subset of the items (if all items measure the same ability). This is for example not true when we merely use the number of correct answers. Similarly, the estimated difference in difficulty level between two items is the same regardless which people are used to measure the difficulty of the items. It does not matter whether we take 20 persons with an ability of 80 and 30 persons with an ability of 100, or we take 50 people with ability scores uniformly distributed between 70 and 100. This is related to a second property of Rasch scales: the estimation of ability and difficulty needs no assumption about their distribution. There is for example no need for a constraint on the distribution of the ability parameters, such as a normal distribution. The distribution is an empirical question. If the Rasch model fits the data, then the estimates of the θ parameters can be regarded as interval level measures of ability on the logit scale and one can check whether on that scale, the distribution of the ability parameters is normal. This is not true for sum scores: the distribution is a direct consequence of the difficulty levels of the items in the test.

An important point is that the ability measures based on the Rasch model are estimates, just as a sum score is an estimate of the true score in classical test theory. A Rasch estimate for ability is more reliable when the test contains many items with difficulty levels comparable to the true ability score. Therefore, and in contrast to classical test theory, the reliability of an ability measure may vary across the scale. For more on Rasch modeling, see Smith and Smith (2004) and Bond and Fox (2001).

Studies have shown that the Raven is largely unidimensional (Rost & Gebert, 1980), but there are also indications that the Raven test might be multidimensional (Lynn, Allik, & Irwing, 2004; Van der Ven & Ellis, 2000; Vigneau & Bors, 2005). Multidimensionality is often noticed in tests with items varying widely in difficulty. Linear factor models then usually show several factors, one for each difficulty level, a phenomenon generally attributed to non-linearity (Gibson, 1959). In the case of the Raven, the dimensions are highly correlated. Lynn, Allik, and Irwing (2004) showed for the Standard Raven that all three factors they found loaded highly onto one second-order factor. The correlations between the three factors and the second-order factor were .95, .80 and .90. Thus, the use of a unidimensional Rasch model leads to only very limited bias. There are also indications that the Standard Raven is biased across gender (Abad, Colom, Rebollo, & Escorial, 2004; Mackintosh & Bennett, 2005) and that there are sex by age interactions (Lynn et al., 2004). Despite these indications of suboptimal fit of the Rasch model, imperfect scaling is to be preferred over no scaling at all. The bias due to multiple highly intercorrelated factors is negligible.

The Rasch based intelligence scores were estimated using the Gibbs sampler as implemented in the BUGS software (http://www.mrc-bsu.cam.ac.uk/bugs) by taking the mean of each individual's posterior distribution. The estimation procedure used no assumptions regarding the distribution of the intelligence scores or item difficulties. Extreme scores (like no item correct or all items correct) are inestimable in the Rasch model. Therefore, individuals who had extreme scores were assigned a value half a logit higher than the second highest scoring individuals.

2.5.2. Extended twin design

In the classical twin study, the relative influence of variation in genes and environment is estimated by comparing MZ and DZ correlations, or covariances. The more similar MZ twins are relative to DZ twins, the more variability in phenotype is caused by genetic variability. When DZ twins resemble each other and are as alike as MZ twins, the resemblance between twins is caused by shared environment, and therefore it can be concluded that part of the variability in intelligence is caused by variability in shared environment. A distinction can be made between variation caused by additive genetic effects (A; caused by the additive effects of alleles at multiple loci), dominance genetic effects (D; non-additive effects of alleles), and environmental effects (E). Environmental effects might be correlated in offspring since they share potentially important environmental factors such as SES. The covariance of E is often denoted as the shared or common environmental variance component (C). The assumption is that MZ twins have the same DNA sequence and therefore A and D are perfectly correlated in MZ twins. DZ twins and siblings share on average half of their segregating genes, therefore the genetic correlation between their additive genetic values (A) is 1/2 (this correlation is higher in the presence of phenotypic assortment). The genetic correlation between the dominance deviations (D) is 1/4. Formally, stated as a random effects model the phenotypes of twins and siblings are modeled as:

$$P_{\text{sibling1}} = h^*A_1 + d^*D_1 + e^*E_1,$$

and

$$P_{\text{sibling2}} = h^*A_2 + d^*D_2 + e^*E_2,$$

where A, D and E are standardized to have unit variance. $\text{Corr}(A_1\ A_2)=1$ for MZ twins, $\text{Corr}(A_1\ A_2)=1/2$ for DZ twins

and full siblings; Corr(D_1 D_2)=1 for MZ twins, Corr(D_1 D_2) =1/4 for DZ twins and siblings and Corr(E_1 E_2) is to be estimated. The variance in P due to A, D and E is given by the square of h, d and e, respectively, so that Var(P)=$h^2+d^2+e^2$. The variance attributable to Common environment (environment shared by siblings from the same family) is obtained as: Var(C)=Corr(E_1 E_2)*e^2. Note that e^2 also contains variance due to measurement error. When only data from twins and siblings reared together are available, it is only possible to estimate Corr(E_1 E_2) under the assumption that d is zero or any other specified value, and vice versa, since a model including free parameters for both Corr(E_1 E_2) and d is not identified.

For a parent, we have

$$P_{\text{parent}} = h^*A_p + d^*D_p + e^*E_p.$$

In the absence of assortative mating, the expectation for Corr (A_p A_1)=Corr(A_p A_2) is 1/2. When there is no cultural transmission, or any other shared environment between parents and offspring, the expectation for Corr(E_p E_1)=Corr(E_p E_2) equals 0. Regardless of cultural transmission and assortative mating, the expectation for Corr(D_p D_1)=Corr(D_p D_2) is 0. When data from both twins and parents are available, the effects of cultural transmission and genetic dominance can be estimated at the same time.

Two different approaches were evaluated to model assortment between the parents of twins. The first model assumed that spousal resemblance was due to phenotypic assortment. The second model assumed that spousal correlation was caused by social homogamy. Fig. 1 shows a path diagram of the model assuming phenotypic assortment. It is based on Fulker (1982) with the addition of dominance genetic variance. The phenotypes of the parents and one child are represented by IQ_f, IQ_m and IQ_o (father, mother and offspring). Variability in intelligence is caused by variation in A, E and D, and these are represented as latent factors in the model and have unit variance. The factor loadings on the latent factors are represented by, h (for A), e (for E), and d (for D). Parents pass their genes to their children, which is represented by arrows going from A of the parents to A of the child, with the factor loading 1/2. In the children, part of the genetic variance is explained by transmission from the parents. The remaining residual additive genetic variance represents the variance that results from recombination. Because dominance effects are not transmitted from parents to offspring there are no paths going from the parental Ds to the child's D (Cavalli-Sforza & Bodmer, 1971).

The Greek letters on the left of the diagram in Fig. 1 represent the correlations induced by phenotypic assortment. Coefficient γ represents the genotypic correlation between the parents, ε the environmental correlation between the parents, and δ represents the correlation of the environment of one parent with the genotype of the other parent. There is no dominance correlation between the parents, since in the case of polygenic inheritance this correlation is negligible (Cavalli-Sforza & Bodmer, 1971). All three correlations are induced by phenotypic assortment that can be represented as a parameter μ equal to the spousal correlation. This spousal correlation can be drawn as a co-path (Cloninger, 1980) between the phenotypes

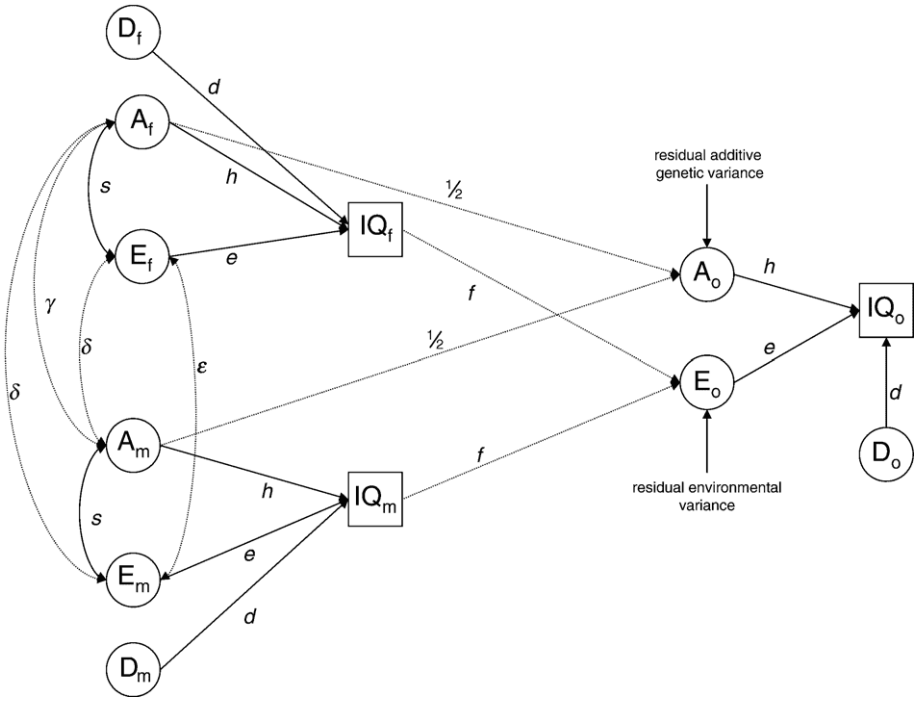

Fig. 1. Path model for the spouse and parent–offspring correlations under the assumption of phenotypic assortment with γ representing genotypic correlation between parents, ε environmental correlation between parents, δ correlation between environment of one parent with genotype of other parent, f cultural transmission, s genotype environment correlation, f father, m mother, o offspring, A additive genetic value (h its factor loading), E environmental value (e its factor loading), D dominance variation (d its factor loading). Twins and sibling are not drawn in this figure for clarity reasons; however they mirror the drawn components (the relationship between twins are drawn in Fig. 2).

of the parents instead of the paths which are represented by the Greek letters. Parameter μ can be written as a function of γ, ε, or δ (cf Fulker, 1982):

$$\mu = \gamma/(h+se)^2 = \varepsilon/(e+hs)^2 = \delta/(e+hs)(h+es).$$

Cultural transmission is represented by f, the regression of the child's environment on the parents' phenotypes. If f is not equal to 0, genotype and environment in the offspring generation become correlated (GE correlation). It is assumed that the system is at equilibrium, thus stable over generations, and therefore genotype and environment are correlated to the same extent in the parents as in the offspring. This GE correlation, s, is represented by the double-headed arrow between A and E of the parents. The correlation is implied in the offspring generation and at equilibrium equals $(1+\mu)fh/[1-(1+\mu)fe]$ (Eaves, Eysenck, & Martin, 1989). The residual environmental variance represents environmental effects not transmitted by the parents.

Fig. 2 represents the effects of phenotypic assortment on the twin and sibling correlations. The phenotypes of the children are indicated by IQ_{T1}, and IQ_{T2} (oldest twin and youngest twin). The sibling data are omitted from the figure for clarity, but the expectations for twin-sib resemblance are the same as for DZ twin resemblance. Variation in intelligence is caused by variation in A, D and E, and the factor loadings for these variance components are represented by h, d, and e. Since dominance variation is not transmitted from the parents to their offspring, spousal resemblance does not influence correlations between dominance deviations in siblings and DZ twins (Cavalli-Sforza & Bodmer, 1971). Since A is transmitted from parents to their offspring, mate resemblance influences the twin and sibling correlations. MZ twins share the same DNA regardless whether phenotypic assortment takes place or not; therefore the genetic correlation in A between MZ twins stays 1. For DZ twins and sibs the correlation in A depends on the genotypic correlation between the parents. On average DZ twins share half their DNA, but the correlation between the genotypic values changes as a function of the genotypic correlation between the parents, γ.

Table 1
Expected correlations between family members based on two genetic models

Correlation	Expectation
Phenotypic assortment	
MZ	$h^2+e^2\ (2f^2(1+\mu)+\beta)+d^2+2hse$
DZ / siblings	$1/2\ h^2\ (1+\gamma)+e^2\ (2f^2(1+\mu)+\beta)+1/4d^2+2hse$
Parent–child	$1/2\ h\ (h+se)\ (1+\mu)+ef\ (1+\mu)$
Spouse	μ
Social homogamy	
MZ	$h^2+d^2+c^2+e^2$
DZ / siblings	$1/2\ h^2+1/4\ d^2+c^2+e^2$
Parent–child	$1/2\ h^2+c^2+e^2$
Spouse	c^2

The environmental correlation among offspring as the result of cultural transmission depends not only on f, but also on the phenotypic correlation between the parents, μ. Var(C), the environmental variance in the classical twin model that is shared by offspring, is now represented as the variance in the phenotype due to cultural transmission, $2\ e^2\ f^2\ (1+\mu)$, plus $e^2\ \beta$, the residual environmental variance shared by siblings (see Fig. 2, cf. Boomsma & Molenaar, 1987). GE correlation is represented by parameter s, both within and across twins.

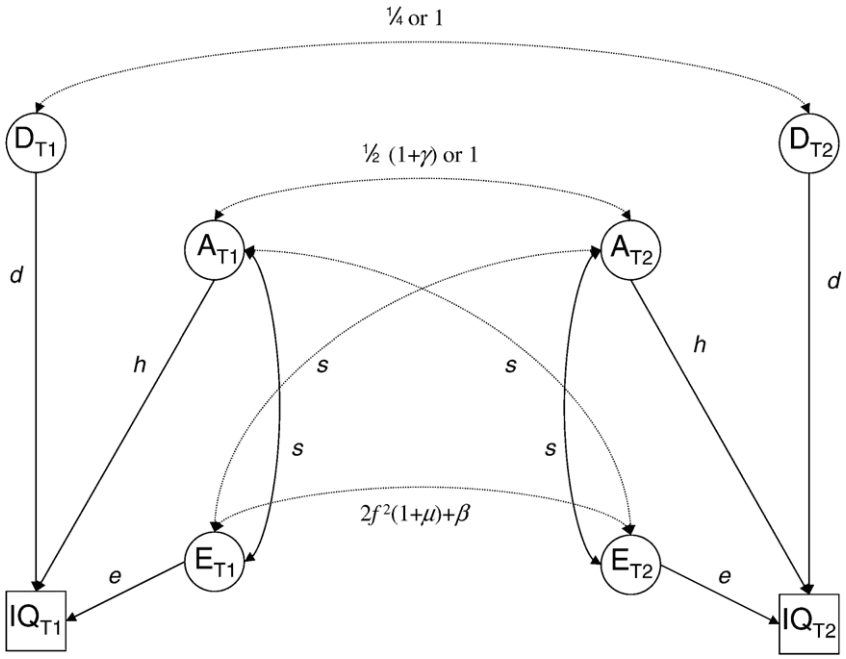

Fig. 2. Path model for the twin correlations under the assumption of phenotypic assortment with γ representing genotypic correlation between parents, μ spousal correlation, β residual environmental covariance not explained by cultural transmission, f cultural transmission, s genotype environment correlation, A additive genetic value (h its factor loading), E environmental value (e its factor loading), D dominance variation (d its factor loading).T1 oldest twin, T2 youngest twin. The sibling is not drawn in this figure for clarity reasons; however the relationship between twins and sibling is similar to the relationship between dizygotic twins.

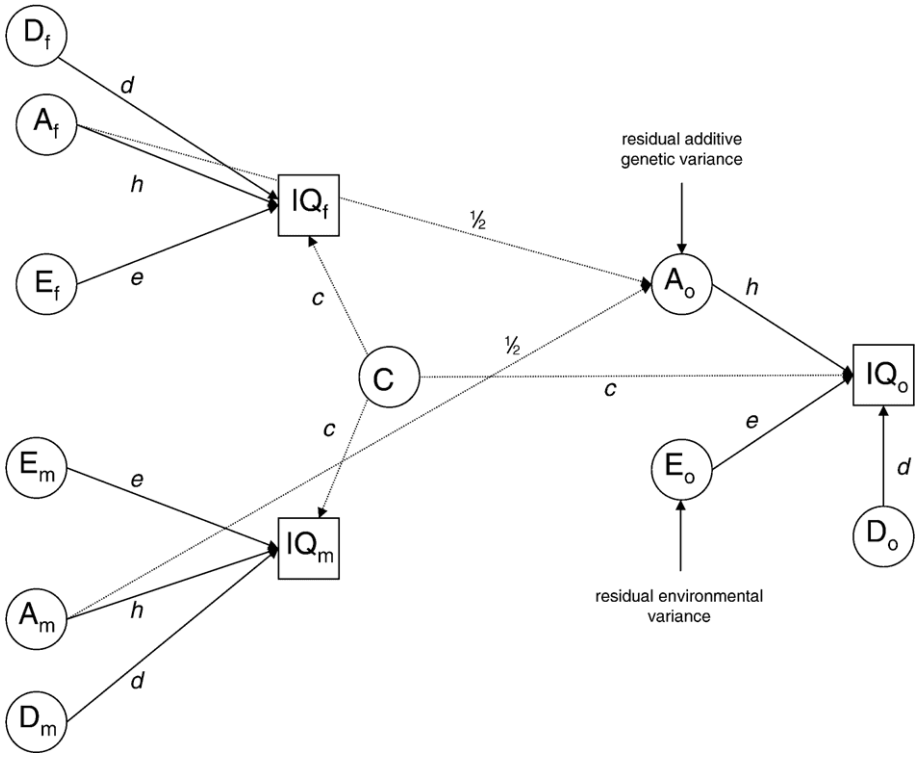

Fig. 3. Path model for the spouse and parent–offspring correlations under the assumption of social homogamy, with f representing father, m mother, o offspring, *A* additive genetic value (*h* its factor loading), *E* unique environmental value (*e* its factor loading), *D* dominance variation (*d* its factor loading), *C* common environmental value (*c* its factor loading). Twins and sibling are not drawn in this figure for clarity reasons; however they mirror the drawn components (the relationship between twins and sibling are drawn in Fig. 4).

Table 1 presents the derived expected correlations between family members in this model. Estimation of both Var(*D*) and β is not possible.

Fig. 3 presents an alternative model that assumes that spousal correlation is due to social homogamy. Here, phenotypic resemblance in IQ in parents is only accounted for by a common environmental effect, *C*, that is uncorrelated with genotype *A*. This environmental effect is assumed to be the same in their children with an equal influence on the phenotype, *c*. Since the child's environment does not depend on the phenotypes of the parents, there is no GE correlation. Fig. 4 gives the implications of these assumptions for the resemblance between twins and

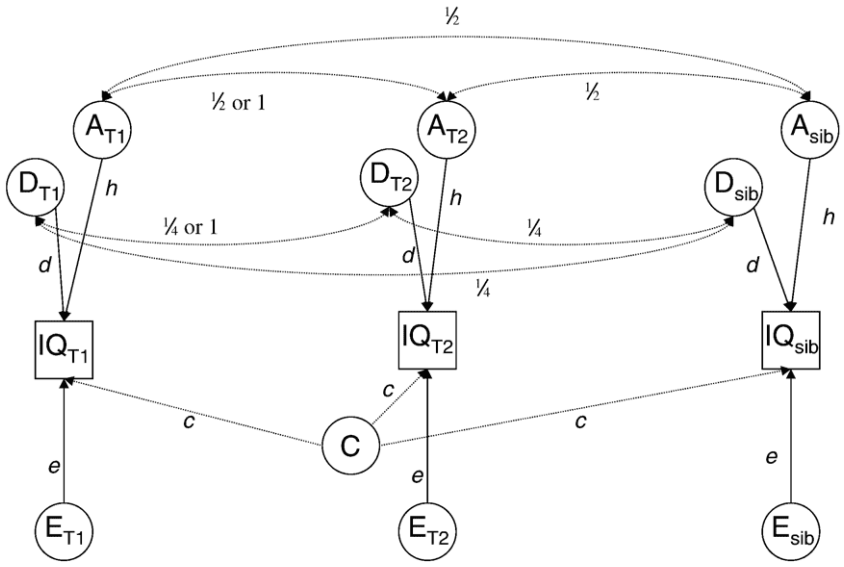

Fig. 4. Path model for the twin and sibling correlations under the assumption of social homogamy with T1 representing the oldest twin, T2 the youngest twin, sib the sibling, *A* additive genetic value (*h* its factor loading), *E* unique environmental value (*e* its factor loading), *D* dominance variation (*d* its factor loading), *C* common environmental value (*c* its factor loading).

their siblings. In this model it is possible to estimate both Var(C) and Var(D). Table 1 gives the derived expected correlations between family members under the assumption of social homogamy.

2.5.3. Model fitting

The Rasch-based IQ measures were first analyzed by fitting a general covariance matrix to the data from DZ and MZ twin families. In these general models, several assumptions were tested, such as equality of means and variances between MZ and DZ twins and between twins and siblings. The model was also used to test for sex and age effects on the means. Within the models best fitting model was chosen based on likelihood-ratio tests. Next, the Rasch IQ measures were fit to both the phenotypic assortment model and the social homogamy model using the statistical modeling package Mx (Neale, Boker, Xie, & Maes, 2003). Between models the best fitting model was chosen by minimizing Akaike's Information Criteria (AIC; $X^2 - 2 * df$). The scripts can be downloaded from www.psy.vu.nl/mxbib.

Because the general model indicated that the MZ twin correlation was about twice the DZ twin correlation, we chose to include dominance genetic variance, and therefore not to estimate β in the model which assumed phenotypic assortment; β was fixed to zero. Heritabilities in parents and their offspring were assumed to be equal (see Rijsdijk, Vernon, & Boomsma, 2002; Reynolds et al., 1996).

The best fitting genetic model resulting from the Mx analyses on the Rasch IQ measures (which are estimates) was also estimated using the raw item data directly in BUGS (Van den Berg et al., 2007; Van den Berg, Beem, & Boomsma, 2006). For clarity we report about the credibility regions as confidence intervals, although in Bayesian statistics one generally speaks of credibility regions instead of confidence intervals. The results on the estimated Rasch IQ measures are somewhat biased for two reasons: the precision of the estimates is not equal across generations, since the estimates in the parents were based on 36 items from the Advanced version and the estimates in the offspring were based on 60 items from the Standard version. The reliabilities for the scales might be different and by modeling the observed item data directly, one adjusts the model parameters for attenuation effects due to scale unreliability that might be different across test version. Secondly, scale reliability does not only differ across test versions, but is also dependent on the location on the scale: estimation precision is usually better for people with average scores than people at the extremes of the scales. By modeling the item data directly one gets parameter estimates that take all these scale effects into account, yielding results that are corrected for attenuation effects (see also Van den Berg et al., 2007).

To test for GE interaction, the average scores of MZ twin pairs were correlated with the absolute differences within a pair. Differences within MZ twin pairs can be attributed to the environment, and differences between MZ twin pairs can be attributed to genotype and environmental effects shared in twins. Thus, if the averages and differences are correlated — and there are no shared environmental effects — this suggests that people with a certain genotype are more sensitive to environmental influences than people with another genotype. Since the scores are estimates, and the precision of a Rasch IQ estimate is dependent of the location on the scale (see above), the correlation estimate and its p-value might be incorrect. Therefore the correlation between the latent trait average and latent absolute difference was estimated by modeling the observed item data in BUGS and applying the Rasch model.

Table 2
Descriptives for all subjects of the sum IQ score on the Raven Progressive Matrices

	N	Minimum	Maximum	Mean	SD
Fathers	94	4	36	27.0	6.5
Mothers	95	9	36	25.9	6.0
Male siblings	44	24	56	43.8	7.8
Female siblings	57	30	59	46.4	6.5
Male twins	114	13	50	36.7	8.6
Female twins	110	19	50	36.6	7.1

Parents received the Advanced Progressive Matrices Test (maximum achievable score=36) and offspring received Standard Progressive Matrices (maximum achievable score=60).

Table 3
Expected phenotypic correlations, variances and covariances for the general reference model and for the genetic models with phenotypic assortment and social homogamy

	Reference model			Phenotypic Assortment			Social Homogamy		
	Var	Cov	r	Var	Cov	r	Var	Cov	r
Twin	1.18			1.18			1.20		
Sibling	1.82			1.79			1.81		
Spouse	2.60	.85	.33	2.64	.84	.32	2.56	.70	.27
Twin MZ		.74	.63		.69	.58		.74	.61
Twin DZ		.30	.25		.41	.34		.48	.39
Twin-sibling		.55	.37		.50	.34		.58	.39
Parent-twin		.62	.35		.68	.38		.61	.35
Parent-sibling		.83	.38		.83	.38		.75	.35

3. Results

Descriptive statistics of the Raven IQ sum scores are in Table 2. For the estimated IQ measures based on the Rasch scaling, no significant sex differences were observed: neither in the total group, nor within groups (parents, siblings, twins). There was no age by sex interaction in the offspring. The variance in the siblings was significantly larger than in the twins, which could partly be explained by age differences. Therefore the age effect was retained in all models. Phenotypic correlations, variances and covariances estimated in the general reference (non-genetic) model are given in Table 3 and model fit is in Table 4. The distribution of Rasch IQ scores looked more or less normal (see Fig. 5). The distribution of the estimated measures in twins showed a slight negative skew and in the sibling and parental data a slight positive skew.

In the genetic analyses, the larger variance in the siblings was modeled using a scalar effect in addition to the age effect to account for their variance, assuming that the components of

Table 4
Fit indices for the general (non-genetic) reference model, best fitting phenotypic assortment genetic model and best fitting social homogamy genetic model

Model	-2LL	# free parameters	# *df*	AIC
Reference	1633.95	13	500	633.95
Phenotypic Assortment	1635.71	9	504	627.71
Social Homogamy	1636.08	10	503	630.08

genetic and environmental variance were proportional to those observed in twins. Fitting of the model assuming phenotypic assortment (see Figs. 1 and 2) showed that including dominance variation in this model does not lead to a significantly better fit (*d*; −Δ2LL=.88; 95% confidence interval 0, .75). There was also no significant contribution of cultural transmission (*f*; −Δ2LL=.86; 95% confidence interval −.30, .44) and therefore no GE correlation (*s*). A simple model with only additive genetic effects and non-shared environmental effects explained the data best. The expected phenotypic correlations, variances and covariances are given in Table 3; the model fit is in Table 4. In this model genetic variation contributes 58% to the variation in intelligence in children as well as adults. The remaining 42% is explained by unique environmental variation.

In the model assuming social homogamy (see Figs. 3 and 4) there is a significant contribution of dominance variance (fixing *d* to 0 leads to a significantly worse fit Δ-2LL=2.93), and social environment (fixing *c* to 0 leads to a significantly worse fit Δ-2LL=8.16). Additive genetic variance could however be dropped (Δ-2LL=0.73). As a model with only dominance genetic variance is a priori not sensible, this additive genetic component was retained. Tables 3 and 4 present the expected variances, covariances, correlations and model fit indices. In this model additive genetic variation contributes 15% to the variation in intelligence in children as well as adults, dominance deviation explains 19% in variation in IQ, and shared environment explains 27%. The remaining 39% is explained by non-shared environmental variation. Comparing the phenotypic assortment model and the social homogamy model, the model assuming phenotypic assortment appears superior as it showed a higher likelihood while having fewer parameters.

The phenotypic assortment model was also estimated in BUGS, this time on the raw item data. The estimate for *h* was a bit higher, leading to a heritability estimate of 67% (95% confidence interval: 52%, 79%). Similar to the analyses on the Rasch estimates, the parameter for the effect of a sib's age (in years) was not significantly different from zero (.18, 95% confidence interval: −0.02, 0.37). Estimated variance of the unobserved intelligence scores was 1.99 in the parents, 1.08 in the twins and 1.46 in the siblings (after age correction). The 67% point estimate can be regarded as the estimate for the heritability that we would get with an infinite number of similar test items, that is, corrected for attenuation effects (cf. Van den Berg et al., 2007).

The estimate for the correlation between average intelligence and difference between MZ twins is −.30, which is significantly different from 0, $p<.05$ (95% confidence interval: −.08, −.52). This suggests that the environment is relatively more important in explaining individual differences for low IQ groups than for high IQ groups. This GE interaction effect explains 9% of the variance in the scores (Jinks & Fulker, 1970). In the models fitted above, the G*E variance is attributed to environmental effects not shared by family members.

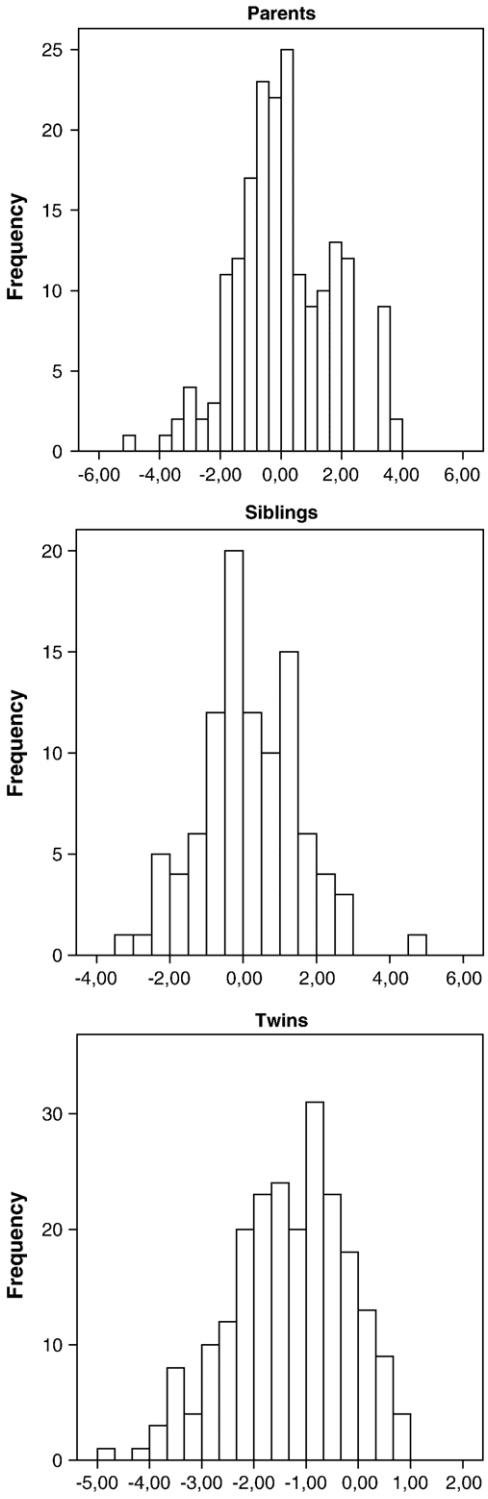

Fig. 5. The distribution of estimated, Rasch-based, IQ scores for parents, siblings and twins.

4. Discussion

In this study several quantitative genetic models to study the heritability of intelligence were evaluated using data from twins, one of their siblings and both parents. With a Rasch measurement model, a measure of IQ based on the Raven Progressive Matrices test was estimated in all participants.

Correlations were higher in MZ twins than in first-degree relatives (siblings, DZ twins and parent–offspring pairs). The spousal correlation for the Rasch IQ estimates was significant and moderately high (0.33). A model assuming that this correlation is due to phenotypic assortment proved superior to a model assuming that the correlation was due to purely environmental factors that are transmitted from generation to generation. Corrected for scale unreliability effects, additive genetic effects account for 67% of the variation in intelligence and the remainder is explained by random environmental factors, including measurement error. Non-additive genetic effects (*D*) and cultural transmission effects (*f*) were not significant. Some other studies have suggested that non-additive genetic effects plays a role in the heredity of intelligence (Chipuer, Rovine, & Plomin, 1990; Fulker 1979; Jinks & Fulker, 1970). We did not find evidence for genetic non-additivity, which seems consistent with most of the behavior genetics literature on IQ (though we recognize that studies that explicitly addressed this issue are scarce).

The absence of common environmental effects shared by family members is in line with the findings from for instance the CAP study where an adoptive parent's IQ does not predict the IQ of the adopted child (Phillips & Fulker, 1989). Prior studies on intelligence in children have reported environmental influences that are shared by siblings, a finding that we did not replicate. Usually familial environmental effects are only seen in children and tend to disappear in adolescence (e.g., Posthuma et al., 2001; Scarr & Weinberg, 1983). There may be several reasons why other studies found such effects and we did not. First of all, our final model assumed phenotypic assortment. When phenotypic assortment is not controlled for, an analysis based on only MZ and DZ twin correlations overestimates shared environmental influences and underestimates additive genetic variance. Therefore, in studies where only twins are used, part of the variance that is labeled shared environmental influences may actually include genetic variance due to assortative mating. Secondly, the absence of shared environmental influences may be related to the IQ measure that was used. IQ was assessed with the Raven Progressive Matrices, a test conceptually more related to performal IQ than verbal IQ. Thus, the findings of our relatively high heritability estimate relative to other studies in children and the absence of shared environmental influences may be due to the measure that was used in addition to modeling the effects of assortative mating.

One important assumption in the modeling was that heritability was equal across generations, and the same genes are expressed. Regarding the first assumption, the heritability estimate based on the estimated scores, uncorrected for reliability, (58%), is comparable to the 64% reported by (Rijsdijk et al., 2002), who collected Raven data in Dutch 16-year-old twin pairs, and also comparable to the heritability observed in adults by Reynolds et al. (1996). Regarding the second assumption, it is known that intelligence scores are highly stable and that in children, this stability is partly due to a common genetic factor explaining IQ at different ages: the genes that influence IQ in early childhood are largely the same genes that influence IQ at later ages (e.g., Bartels et al., 2002). Also in adulthood, stability in intelligence is largely due to the same set of genetic factors (e.g. Plomin et al., 1994; Van den Berg, Posthuma & Boomsma, 2004; see also DeFries, Plomin, & LaBuda, 1987; Plomin, Fulker, Corley, & DeFries, 1997). Thus, although there is some evidence that the genetic correlation across age is not perfect, the conclusions from our models are not likely to be severely biased.

Our conclusions are based on the assumption of phenotypic assortment. A model with phenotypic assortment provided a more parsimonious explanation of the present data than a model with social homogamy. This finding is in contrast to that obtained by Reynolds et al. (1996) who studied twins born between 1911 and 1935. They reported that social homogamy could explain spousal similarity and that phenotypic assortment was not significant. However, their analysis was based on the (unlikely) assumption that there is no correlation between genotypes and the environment in which prospective partners meet. Alternatively, it is possible that nowadays, social homogamy plays a less important role than in the early 20th century.

There was evidence for GE interaction, suggesting that the environment is relatively more important in explaining individual differences for low IQ groups than for high IQ groups. Similar findings were reported by Jinks & Fulker (1970), Jensen (1970) and Finkel and Pedersen (2001), although their effect sizes were smaller. The GE interaction effect is in agreement with findings from Turkheimer et al. (2003) who showed that the relative influence of genotype is larger for children from parents of high social-economic status (SES) than for children from low SES parents.

We found that the mean IQ score in the older siblings was higher and also that there was more variance in siblings than in twins, even though the same test was used. This could not be fully explained by age differences among the siblings. The finding is, however, consistent with results obtained by Thurstone (1928) who showed a positive relationship between group mean and group variance with scaled intelligence scores. Such a phenomenon cannot be observed in normed IQ scores by definition. Future research should determine whether it is merely a scaling effect or whether it perhaps reflects increased variability due to individual differences in the timing of puberty.

Variability in fluid intelligence as measured by the Raven is largely explained by additive genetic effects that are transmitted from parents to offspring. In accordance with adoption studies (Alarcón et al., 2003; Fulker & DeFries, 1983; Humphreys & Davey, 1988; Rice et al., 1989; cf. Scarr & Weinberg, 1978, 1983), we found no evidence for cultural transmission: all influence from parents on their children's IQ was explained by the transmission of genes. However, in the approach that was taken, cultural transmission was modeled as a direct effect of parental IQ on offspring environment. Although this model does not seem unreasonable for IQ, it might be that Raven IQ

does not capture those aspects of the parental phenotype that are most salient in determining the child's Raven IQ.

The present study design is not suited to uncover GE correlations other than one resulting from simultaneous genetic and cultural transmission. But what we can conclude is that if there is GE correlation, the role of parents seems limited to responding to the needs and interests as indicated by the child. We found no indication that intelligent parents provide their offspring with intelligence promoting circumstances. More likely, children with a genetic predisposition for either a low or a high IQ ask for a specific type of stimulation. In other words, an evocative gene–environment correlation (where individuals are reacted to on the basis of their genetically influenced phenotype) or an active GE correlation (where individuals seek or create environments correlated with their genetic inclinations) seems a more probable mechanism than a passive GE correlation (Scarr & McCartney, 1983). Only the last type of correlation could in principle have been detected by our extended twin family design.

In conclusion, individual differences in intelligence are largely accounted for by genetic differences. Environmental factors are significantly more important in children with a genetic predisposition for low IQ than in children with a genetic predisposition for high IQ. Environmental factors influencing IQ are generally not shared among siblings.

For future research we recommend to implement extended twin designs similar to the one used in this study. Although our study consisted of a high number of participants (516), we only included a limited number of families. In our sample we had only limited power to detect effects of genetic dominance and perhaps they will reach significance in a larger sample. Measures of cognition that include aspects of e.g. verbal cognition, correlates of IQ such as brain volume and function, and inclusion of twins and sibs of different ages should shed more light on, for instance, the presence of cultural transmission for verbal IQ, how genetic effects on IQ are mediated, and the extent to which results generalize to younger and older children. Moreover, we recommend that future genetic research on intelligence focuses on the exact nature of the GE interaction and the possible existence of evocative and active GE correlation.

Acknowledgements

We would like to acknowledge the following grants: Quantitative genetic modelling of variation in cognitive brain maturation (*NWO 051.02.060*, *668.772*); Spinozapremie (*NWO/SPI 56-464-14192*); CNCR (Center for Neurogenomics and Cognition Research); Database Twin Register (*NWO 575-25-006*); Twin-family database for behavior genetics and genomics studies (*NWO 480-04-004*) and we would like to thank the participating twin families.

References

Abad, F., Colom, R., Rebollo, I., & Escorial, S. (2004). Sex differential item functioning in the Raven's Advanced Progressive Matrices: Evidence for bias. *Personality and Individual Differences*, *36*, 1459–1470.

Agrawal, N., Sinha, S. N., & Jensen, A. R. (1984). Effects of inbreeding on Raven matrices. *Behavior Genetics*, *14*, 579–585.

Alarcón, M., Plomin, R., Corley, R. P., & DeFries, J. C. (2003). Multivariate parent–offspring analyses of specific cognitive abilities. In S. A. Petrill, R. Plomin, J. C. DeFries, & J. K. Hewitt (Eds.), *Nature, nurture, and the transition to early adolescence.* Oxford: Oxford University Press.

Bartels, M., Rietveld, M. J. H., Van Baal, G. C. M., & Boomsma, D. I. (2002). Genetic and environmental influences on the development of intelligence. *Behavior Genetics*, *32*, 237–249.

Bashi, J. (1977). Effects of inbreeding on cognitive performance. *Nature*, *266*, 440–442.

Bishop, E. G., Cherney, S. S., & Hewitt, J. K. (2003). Developmental analysis of IQ. In S. A. Petrill, R. Plomin, J. C. DeFries, & J. K. Hewitt (Eds.), *Nature, nurture, and the transition to early adolescence.* Oxford: Oxford University Press.

Bond, T. G., & Fox, C. M. (2001). Applying the Rasch model: fundamental measurement in the human sciences. Mahwah, N.J.: Erlbaum.

Boomsma, D. I. (1993). Current status and future prospects in twin studies of the development of cognitive abilities, infancy to old age. In T. J. Bouchard, & P. Propping (Eds.), *Twins as a Tool of Behavioral Genetics* (pp. 67–82). Chichester: John Wiley & Sons.

Boomsma, D., Busjahn, A., & Peltonen, L. (2002). Classical twin studies and beyond. *Nature Reviews Genetics*, *3*, 872–882.

Boomsma, D. I., De Geus, E. J. C., Vink, J. M., Stubbe, J. H., Distel, M. A., Hottenga, J. J., et al. (2006). Netherlands Twin Register: From twins to twin families. *Twin Research and Human Genetics*, *9*, 849–857.

Boomsma, D. I., & Molenaar, P. C. (1987). Constrained maximum likelihood analysis of familial resemblance of twins and their parents. *Acta Geneticae Medicae Gemellogiae*, *36*, 29–39.

Boomsma, D. I., Orlebeke, J. F., & Van Baal, G. C. (1992). The Dutch Twin Register: Growth data on weight and height. *Behavior Genetics*, *22*, 247–251.

Boomsma, D. I., Vink, J. M., Van Beijsterveldt, T. C., De Geus, E. J., Beem, A. L., Mulder, E. J., et al. (2002). Netherlands Twin Register: A focus on longitudinal research. *Twin Research*, *5*, 401–406.

Bouchard, T. J., Jr., & McGue, M. (1981). Familial studies of intelligence: A review. *Science*, *212*, 1055–1059.

Bouchard, T. J., Jr., & McGue, M. (2003). Genetic and environmental influences on human psychological differences. *Journal of Neurobiology*, *54*, 4–45.

Cavalli-Sforza, L. L., & Bodmer, W. F. (1971). Quantitative characters, polygenic inheritance, and environmental interactions. *The Genetics of Human Populations* (pp. 508–633). San Francisco: W.H. Freeman and Company.

Cherny, S. S., & Cardon, L. R. (1994). General cognitive ability. In J. C. DeFries, R. Plomin, & D. W. Fulker (Eds.), *Nature and nurture during middle childhood* (pp. 46–56). Oxford: Blackwell.

Chipuer, H. M., Rovine, M. J., & Plomin, R. (1990). LISREL modeling: Genetic and environmental influences on behavior revisited. *Intelligence*, *14*, 11–29.

Cloninger, C. R. (1980). Interpretation of intrinsic and extrinsic structural relations by path analysis: Theory and applications to assortative mating. *Genetical Research*, *36*, 135–145.

Deary, I. J., Spinath, F. M., & Bates, T. C. (2006). Genetics of intelligence. *European Journal of Human Genetics.*, *14*, 690–700.

DeFries, J. C., Plomin, R., & LaBuda, M. C. (1987). Genetic stability of cognitive development from childhood to adulthood. *Developmental Psychology*, *23*, 4–12.

Eaves, L. J., Eysenck, H. J., & Martin, N. G. (1989). Genes, culture and personality: an empirical approach. London: Academic.

Finkel, D., & Pedersen, N. L. (2001). Sources of environmental influence on cognitive abilities in adulthood. In E. L. Grigorenko, & R. J. Sternberg (Eds.), *Family environment and intellectual functioning* (pp. 173–194). Mahwah: Lawrence Erlbaum Associates.

Fulker, D. W. (1979). Nature and nurture: Heredity. In H. J. Eysenck (Ed.), *The Structure and Measurement of Intelligence.* Berlin: Springer.

Fulker, D. W. (1982). Extensions of the classical twin method. In B. C. Weir, E. J. Eisen, M. M. Goodman, & G. Namkoong (Eds.), *Proceedings of the second international conference on quantitative genetics* (pp. 395–406). Sunderland, MA: Sinauer.

Fulker, D. W., & DeFries, J. C. (1983). Genetic and environmental transmission in the Colorado Adoption Project: Path analysis. *British Journal of Mathematical and Statistical Psychology*, *36*(Pt 2), 175−188.

Gibson, W. A. (1959). Three multivariate models: factor analysis, latent structure analysis, and latent profile analysis. *Psychometrika*, *24*, 229−252.

Gilger, J. W. (1991). Differential assortative mating found for academic and demographic variables as a function of time of assessment. *Behavior Genetics*, *21*, 131−150.

Guttman, R. (1974). Genetic analysis of analytical spatial ability: Raven's progressive matrices. *Behavior Genetics*, *4*, 273−284.

Guttman, R., & Shoham, I. (1983). Parent–offspring resemblance and parental education. *Behavior Genetics*, *13*, 99−105.

Humphreys, L. G., & Davey, T. C. (1988). Continuity in intellectual growth from 12 months to 9 years. *Intelligence*, *12*, 183−197.

Jensen, A. R. (1970). IQ's of identical twins reared apart. *Behavior Genetics*, *1*, 133−148.

Jinks, J. L., & Fulker, D. W. (1970). Comparison of the biometrical genetical, MAVA, and classical approaches to the analysis of human behavior. *Psychological Bulletin*, *73*, 311−349.

Keller, M. C., Coventry, W. L., Heath, A. C., & Martin, N. G. (2005). Widespread evidence for non-additive genetic variation in Cloninger's and Eysenck's personality dimensions using a twin plus sibling design. *Behavior Genetics*, *35*, 707−721.

Lynch, M., & Walsh, B. (1998). Genetics and analysis of quantitative traits. Sunderland: Sinauer Associates.

Lynn, R., Allik, J., & Irwing, P. (2004). Sex differences on three factors identified in Raven's Standard Progressive Matrices. *Intelligence*, *32*, 411−424.

Mackintosh, N., & Bennett, T. (2005). What do Raven's matrices measure? An analysis in terms of sex differences. *Intelligence*, *33*, 663−674.

Mascie-Taylor, C. G. (1989). Spouse similarity for IQ and personality and convergence. *Behavior Genetics*, *19*, 223−227.

Mascie-Taylor, C. G., & Vandenberg, S. G. (1988). Assortative mating for IQ and personality due to propinquity and personal preference. *Behavior Genetics*, *18*, 339−345.

Neale, M., Boker, S. M., Xie, G., & Maes, H. H. (2003). *Mx: statistical modelling*, (6 ed.). Richmond, VA: Department of Psychiatry, Medical College of Virginia.

Phillips, K., & Fulker, D. W. (1989). Quantitative genetic analysis of longitudinal trends in adoption designs with application to IQ in the Colorado Adoption Project. *Behavior Genetics*, *19*, 621−658.

Phillips, K., Fulker, D. W., Carey, G., & Nagoshi, C. T. (1988). Direct marital assortment for cognitive and personality variables. *Behavior Genetics*, *18*, 347−356.

Plomin, R., DeFries, J. C., Craig, I. W., & McGuffin, P. (2002). Behavioral genetics. In R. Plomin, J. C. DeFries, I. W. Craig, & P. McGuffin (Eds.), *Behavioral Genetics in the Post Genomic Era* (pp. 3−15). Washington D. C.: American Psychological Association.

Plomin, R., Fulker, D. W., Corley, R., & DeFries, J. C. (1997). Nature, nurture and cognitive development from 1 to 16 years: A parent–offspring adoption study. *Psychological Science*, *8*, 442−447.

Plomin, R., Pedersen, N., Lichtenstein, P., & McClearn, G. E. (1994). Variability and stability in cognitive abilities are largely genetic later in life. *Behavior Genetics*, *24*, 207−215.

Plomin, R., & Spinath, F. (2004). Intelligence: genetics, genes, and genomics. *Journal of Personality and Social Psychology*, *86*, 112−129.

Posthuma, D., & Boomsma, D. I. (2000). A note on the statistical power in extended twin designs. *Behavior Genetics*, *30*, 147−158.

Posthuma, D., De Geus, E. J., & Boomsma, D. I. (2001). Perceptual speed and IQ are associated through common genetic factors. *Behavior Genetics*, *31*, 593−602.

Rasch, G. (1966). An item analysis which takes individual differences into account. *British Journal of Mathematical and Statistical Psychology*, *19*, 49−57.

Raven, J. C. (1960). Guide to the Standard Progressive Matrices. London: H.K. Lewis & Co. Ltd.

Raven, J., Raven, J. C., & Court, J. H. (1998). Raven manual: Section 4 advanced progressive matrices. Oxford: Oxford Psychologists Press.

Reynolds, C. A., Baker, L. A., & Pedersen, N. L. (1996). Models of spouse similarity: applications to fluid ability measured in twins and their spouses. *Behavior Genetics*, *26*, 73−88.

Rice, T., Carey, G., Fulker, D. W., & DeFries, J. C. (1989). Multivariate path analysis of specific cognitive abilities in the Colorado Adoption Project: Conditional path model of assortative mating. *Behavior Genetics*, *19*, 195−207.

Rietveld, M. J., van der Valk, J. C., Bongers, I. L., Stroet, T. M., Slagboom, P. E., & Boomsma, D. I. (2000). Zygosity diagnosis in young twins by parental report. *Twin Research*, *3*, 134−141.

Rijsdijk, F. V., Vernon, P. A., & Boomsma, D. I. (2002). Application of hierarchical genetic models to Raven and WAIS subtests: A Dutch twin study. *Behavior Genetics*, *32*, 199−210.

Rost, D. H., & Gebert, A. (1980). Zum Problem der Faktoreninterpretation bei Raven's Coloured Progressive Matrices [On the problem of factor interpretation in Raven's Coloured Progressive Matrices]. *Zeitschrift für Differentielle und Diagnostische Psychologie*, *1*, 255−273.

Scarr, S., & McCartney, K. (1983). How people make their own environments: A theory of genotype–environment effects. *Child Development*, *54*, 424−435.

Scarr, S., & Weinberg, R. A. (1978). The influence of "family background" on intellectual attainment. *American Sociological Review*, *43*, 674−692.

Scarr, S., & Weinberg, R. A. (1983). The Minnesota Adoption Studies — Genetic-differences and malleability. *Child Development*, *54*, 260−267.

Smith, E., & Smith, R. (Eds.). (2004). *Introduction to Rasch measurement: Theory, models and applications.* Maple Grove, MN: JAM Press.

Spearman, C. (1927). The abilities of man: Their nature and measurement. New York: MacMillan.

Thurstone, L. L. (1928). The absolute zero in intelligence measurement. *Psychological Review*, *35*, 175−197.

Turkheimer, E., Haley, A., Waldron, M., D'Onofrio, B., & Gottesman, I. I. (2003). Socioeconomic status modifies heritability of IQ in young children. *Psychological Science*, *14*, 623−628.

Van den Berg, S. M., Beem, L., & Boomsma, D. I. (2006). Fitting genetic models using Markov Chain Monte Carlo algorithms with BUGS. *Twin Research and Human Genetics*, *9*, 334−342.

Van den Berg, S. M., Glas, C. A., & Boomsma, D. I. (2007). Variance decomposition using an IRT measurement model. *Behavior Genetics*, *37*, 604−616.

Van den Berg, S. M., Posthuma, D., & Boomsma, D. I. (2004). A longitudinal genetic study of vocabulary knowledge in adults. *Twin Research*, *7*, 284−291.

Van der Ven, A. H. G. S., & Ellis, J. L. (2000). A Rasch analysis of Raven's standard progressive matrices. *Personality and Individual Differences*, *29*, 45−64.

Vigneau, F., & Bors, D. A. (2005). Items in context: assessing the dimensionality of Raven's Advanced Progressive Matrices. *Educational and Psychological Measurement*, *65*, 109−123.

Wadsworth, S. J., DeFries, J. C., Fulker, D. W., & Plomin, R. (1995). Cognitive ability and academic achievement in the Colorado Adoption Project: A multivariate genetic analysis of parent–offspring and sibling data. *Behavior Genetics*, *25*, 1−15.

Wainwright, M. A., Wright, M. J., Geffen, G. M., Luciano, M., & Martin, N. G. (2005). The genetic basis of academic achievement on the Queensland Core Skills Test and its shared genetic variance with IQ. *Behavior Genetics*, *35*, 133−145.

Watkins, M. P., & Meredith, W. (1981). Spouse similarity in newlyweds with respect to specific cognitive abilities, socioeconomic status, and education. *Behavior Genetics*, *11*, 1−21.

Watson, D., Klohnen, E. C., Casillas, A., Simms, E. N., Haig, J., & Berry, D. S. (2004). Match makers and deal breakers: Analyses of assortative mating in newlywed couples. *Journal of Personality*, *72*, 1029−1068.

Williams, T. (1975). Family resemblance in abilities: The Wechsler scales. *Behavior Genetics*, *5*, 405−409.