Behavior Genetics, Vol. 35, No. 5, September 2005 (© 2005)
DOI: 10.1007/s10519-005-3556-x

Large, Consistent Estimates of the Heritability of Cognitive Ability in Two Entire Populations of 11-Year-Old Twins from Scottish Mental Surveys of 1932 and 1947

Beben Benyamin,[1] Valerie Wilson,[2] Lawrence J. Whalley,[3] Peter M. Visscher,[1] and Ian J. Deary[4,5]

Received 1 Oct. 2004—Final 16 Feb. 2005

Twin studies provide estimates of genetic and environmental contributions to cognitive ability differences, but could be based on biased samples. Here we report whole-population estimates using twins from unique mental surveys in Scotland. The Scottish Mental Surveys of 1st June 1932 (SMS1932) and 4th June 1947 (SMS1947), respectively, administered the same validated verbal reasoning test to almost everyone born in 1921 or 1936 and attending school in Scotland. There were 572 twin pairs from the SMS1932, and 517 pairs from the SMS1947. Information on zygosity was unavailable. A novel application of a mixture distribution was used to estimate genetic and environmental components of verbal reasoning variation by maximum likelihood. We found consistent heritability (~0.70) and shared environment (~0.21) estimates. The estimates did not change substantially when additional quantitative traits (height and weight) were added in a multivariate analysis. More generally for studies in genetics, the methodological innovation developed here implies that large (national) data collections can provide sufficient information on twin pairs to estimate genetic parameters, even without zygosity.

KEY WORDS: Heritability; intelligence; IQ; mixture distribution; twins.

INTRODUCTION

Intelligence differences in humans have a well-understood phenotypic structure (Carroll, 1993), strong predictive validity for health, education and occupational outcomes (Gottfredson and Deary, 2004; Neisser *et al.*, 1996), and correlates in brain structure and function (Gray and Thompson, 2004). The genetic and environmental contributions to variation in intelligence at different ages are of considerable interest, but are not fully understood (Plomin and Spinath, 2004). To date, twin, family and adoption studies suggest that, including studies at all ages, about 50% of the variation in human intelligence, as measured by psychometric tests, is attributable to additive genetic effects (Bouchard, 2004; Bouchard *et al.*, 1990; Plomin and Spinath, 2004; Plomin *et al.*, 2001). These sources also suggest that: the bulk of the substantial environmental effect is from sources not shared by siblings in the same rearing family; the genetic influence is especially strong on the general intelligence factor; the genetic contribution is stronger in adulthood than in childhood; and the effect of the shared rearing environment decreases almost to zero in early adulthood.

[1] School of Biological Sciences, University of Edinburgh, King's Buildings, West Mains Road, Edinburgh, EH9 3JT. UK.
[2] Scottish Council for Research in Education, University of Glasgow, 61 Dublin Street, Edinburgh, EH3 6NL. UK.
[3] Department of Mental Health, University of Aberdeen, Cornhill Royal Hospital, Cornhill Road, AB25 2ZH. UK.
[4] Department of Psychology, University of Edinburgh, 7 George Square, Edinburgh, EH8 9JZ. UK.
[5] To whom correspondence should be addressed at Department of Psychology, University of Edinburgh, 7 George Square, Edinburgh, EH8 9JZ. Tel: +44 131 650 3452; Fax: +44 131 651 1771; e-mail: I.Deary@ed.ac.uk.

A potentially serious and unanswered problem is the representativeness of the samples on which genetic and environmental contributions to variation in cognitive ability scores are based (Bouchard and McGue, 2003; Joseph, 2003). Volunteer samples could be especially prone to such effects, and attrition from population-referenced samples could have biasing effects. Therefore, it is novel and highly desirable to provide estimates of environmental and genetic contributions to intelligence variation based upon complete populations.

For this report we analysed data from unique, whole populations of 11-year-olds tested in the Scottish Mental Surveys of 1932 (SMS1932) (Scottish Council for Research in Education, 1933) and 1947 (SMS1947) (Deary *et al.*, 2004; Scottish Council for Research in Education, 1949). On June 1st 1932 and June 4th 1947, respectively, the Scottish Council for Research in Education organised the mental testing of all children attending Scottish schools and born in 1921 or 1936, respectively. In the Scottish Mental Survey of 1932 (SMS1932) 87,498 children were tested (Scottish Council for Research in Education, 1933), and 70,805 were tested in the Scottish Mental Survey of 1947 (SMS1947) (Scottish Council for Research in Education, 1949). The mental test used was a version of the Moray House Test No. 12 (Scottish Council for Research in Education, 1933).

In addition to the novelty of using data from an entire population we apply a finite mixture distribution model (Neale, 2003) that does not require zygosity of twin pairs to be known. This method provides a unique opportunity to perform genetic analysis on data collected (for non-biological purposes) from large population cohorts in, for example, the social sciences, as long as twin pairs can be identified from local identifiers such as family, school and date of birth.

METHODS

Study populations

Twin pairs were explicitly ascertained by the original researchers in the SMS1947, and some extra demographic and social information was collected from them (Mehrotra and Maxwell, 1949; Scottish Council for Research in Education, 1949, 1953). No attempt was made to establish zygosity. A total of 517 pairs were identified, 320 same-sex (SS) pairs and 197 opposite-sex (OS) pairs. In the SMS1932, twin pairs were not explicitly ascertained (Deary *et al.*, 2004; Scottish Council for Research in Education, 1933). They were identified for the present study by matching pairs of subjects for: surname, date of birth and school identifier. A total of 572 pairs were ascertained, 382 SS and 190 OS pairs. The zygosity status of these twin pairs is not known. The number of twin pairs as a proportion of the entire population was 0.64% and 0.70% for the 1932 and 1947 populations, respectively, slightly lower than the current rate of twinning in Caucasian populations (Imaizumi, 2003). In total, the two populations contained 1089 twin pairs, 702 SS and 387 OS twin pairs.

Twins who attended different schools, including brothers and sisters attending single sex schools (rarer at the primary level than at the secondary level) would not be identified in SMS1932. The computerised SMS1932 database is not complete. The ledgers containing the data from Fife, Angus and Wigtown have not been traced. There is no obvious bias entailed by these omissions as there is no reason to believe that genetic and environmental contributions to intelligence would be any different in those areas. In one area of Scotland, 10 twin pairs were from birth years 1922 and 1923. These were retained because all children were tested in the area from these birth years. Another two of the "twin" pairs are from triplets. Two pairs had no information on sex and were omitted.

Measures

The version of the Moray House Test No. 12 (MHT) was very like that used in the United Kingdom for selection from primary to secondary school education when children were about 11 years old (Deary *et al.*, 2004). Based on 71 questions, the maximum possible score on this test is 76. The MHT is a group-administered test with a time limit of 45 minutes, and contains a preponderance of verbal reasoning items, but also other material including numerical and spatial items. It was validated against the Stanford revision of the Binet test ($r \sim 0.8$) (Scottish Council for Research in Education, 1933). The MHT has high stability of individual differences over more than 60 years, with a correlation coefficient between MHT score between age 11 and almost 80 of 0.66 (Deary *et al.*, 2000, 2004).

Additional phenotypic data, including height, weight and hence body mass index (BMI, body mass in kg divided by the square of the height in m) were available for the SMS1947 twins. Height and weight were measured by a nurse either in the school or at

home (Scottish Council for Research in Education, 1949). Since these traits have well-known heritabilities, including these indices afforded a check on heritability estimates against other studies. Moreover, some studies have reported a moderate phenotypic correlation between IQ and height (Humphreys *et al.*, 1985; Johnson, 1991; Molinari, *et al.*, 2002). By performing multivariate genetic analyses, the present study allowed an examination of the environmental and genetic correlations between cognitive ability and height. Lastly, in the methods we have used to estimate genetic parameters, multiple traits provide more information on zygosity than data on a single trait.

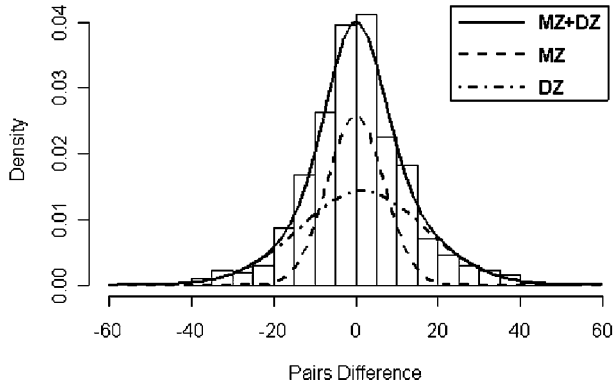

Fig. 1. Probability density function of same-sex (SS) pair difference for the combined Moray House Test score, after adjusting for sex and cohort effects, and the fitted curves for two underlying distributions, assumed to correspond to MZ and SSDZ pairs. The histogram represents the observed distribution of the SS pairs difference and the solid curve the sum of the fitted distributions from the mixture model. The estimated mean, variance, skewness and kurtosis from the single distribution of SS pair difference are 0.78 ± 0.49, 152.1 ± 8.6, 0.28 ± 0.10 and 1.40 ± 0.20, respectively. The estimated means (variances) of the inferred MZ and SSDZ distribution of pair difference are – 0.12 ± 0.42 (48.4 ± 4.1), 1.52 ± 0.83 (236.1 ± 18.1), respectively. The estimated proportion of MZ among SS pairs of 0.45 was used to weight the analysis.

Analysis

Zygosity status, i.e., whether a twin pair is either monozygotic (MZ) or dizygotic (DZ), was not available for the same-sex twin pairs. Opposite-sex pairs are DZ. The number of MZ twins as a fraction of all twin pairs (or as a fraction of SS pairs) can be estimated assuming that the probability that a DZ pair is same-sex is 0.5. From each of the populations, the proportion of MZ twin pairs was estimated as (Scarr-Salapatek, 1971):

1–2(proportion of OS twin pairs).

Intraclass correlations of SS and OS were obtained by partitioning the total variance into a between (σ^2b) and within (σ^2w) pair variance using ANOVA, and fitting sex and cohort (for the combined MHT score of the SMS 1932 and 1947) as covariates.

To partition the observed intraclass correlation of OS and SS twin pair phenotypic similarity into possible underlying causes, a model was fitted that partitions the covariance between twin pairs into an additive genetic (*A*), a common environmental (*C*), and a residual environmental (*E*) component of variance (Neale and Cardon, 1992). The variance components of MHT score and other traits were estimated by fitting a finite mixture distribution method (Neale, 2003) using the statistical package Mx (Neale *et al.*, 2002).

The basic principle of the mixture model is illustrated in Figure 1, where two normal distributions are fitted to the distribution of same-sex pair difference. The combined MHT score, after adjustment for a sex and a cohort effect, was used as an example. There is strong statistical evidence for the hypothesis that there are two distributions against the hypothesis that there is a single normal distribution (likelihood-ratio-test, 2 degrees of freedom, $p = 6.1 \times 10^{-10}$). From the data, the estimated proportion of MZ pairs among SS pairs was 0.45 and was used to weight the analysis. The estimates of the variances for the two underlying distributions, assumed to correspond to MZ and DZ pairs, are 48.4 and 236.1, respectively. We assume that the smaller within-pair variance corresponds to MZ twins because, for any trait with genetic variance, the variance within MZ pairs is less than the variance within DZ pairs. These estimates of the within-pair variance correspond to $2(1 - r)$ times the phenotypic variance, where r is the MZ or DZ correlation. Given the estimate of the total phenotypic variance of 243.4 (Table III), the inferred estimates of the MZ and DZ intra-class correlations are 0.90 and 0.51, respectively, which correspond well to the results from the full mixture model, presented later. In the full model, both the within and between same-sex variances are partitioned into two groups simultaneously, and appropriate weights are given to all sources of information by maximum likelihood.

Sex and cohort effects (for combined MHT score) were fitted as covariates for all traits. The mixture proportion, in this case the proportion of

Table I. Descriptive statistics of Scottish Mental Surveys' twin data

Traits	Sex	*N*	Mean	SD	CV (%)
Height (m) (SMS1947)	M	488	1.37	0.07	4.98
	F	521	1.37	0.07	5.12
	Total	1009	1.37	0.07	5.11
Weight (kg) (SMS1947)	M	480	31.14**	4.33	13.89
	F	519	30.13**	4.41	14.65
	Total	999	30.61	4.40	14.37
BMI (kg/m^2) (SMS1947)	M	480	16.68**	1.68	10.09
	F	519	16.11**	1.59	9.88
	Total	999	16.38	1.66	10.13
MHT Score (SMS1932)	M	508	28.45	15.31	53.81
	F	572	28.46	14.89	52.31
	Total	1080	28.46	15.08	52.99
MHT Score (SMS1947)	M	451	30.88*	16.77	54.31
	F	498	33.41*	15.81	47.32
	Total	949	32.21	16.31	50.64
Combined MHT Score	M	959	29.59	16.05	54.24
	F	1070	30.76	15.51	50.42
	Total	2029	30.21	15.78	52.23

Note: Column *N* is the number of individuals, and column SD and CV are the standard deviations and coefficient of variation, respectively. * and ** denote significant differences between male and female at 5% and 0.1%, respectively.

MZ pairs among SS pairs, is assumed to be known, and the variance components are estimated using maximum likelihood.

Male and female specific variance components were estimated and a likelihood-ratio-test was used to test the hypothesis that these components were the same. The covariance between same-sex DZ pairs and opposite-sex DZ pairs is not necessarily equal; for example, if there are sex-specific genetic effects or if the shared environmental covariance differs between the two types of DZ pairs. This difference was estimated by allowing the genetic correlation between males and females to differ from unity, and tested using a likelihood-ratio-test.

Since additional phenotypic data were available from the SMS 1947 population, multivariate genetic analyses were performed on the 1947 data, with height, weight, BMI and MHT score as phenotypes. Given that additional phenotypes provide additional zygosity information, the mixture model is likely to give more precise estimates of parameters with multiple traits (Neale, 2003).

RESULTS

Descriptive statistics of the MHT score and additional traits are presented in Table I. As is known, girls scored higher in the MHT in the SMS1947, and there was an increase in the mean MHT score from the 1932 population to the 1947 population (Scottish Council for Research in Education, 1949). In SMS1932, the proportion of monozygotic (MZ) among same-sex (SS) twins was 0.50 ± 0.04. The estimate was smaller in SMS1947 (0.38 ± 0.06), but the two are not significantly different ($p > 0.05$). When the twin data were combined, the proportion of MZ in SS was 0.45 ± 0.03, close to the estimated proportion of MZ in Caucasian populations (Imaizumi, 2003). A summary of the estimated proportion of MZ twins in the sample of SS

Table II. Intraclass correlations for same-sex (SS) and opposite-sex (OS) twins from the Scottish Mental Surveys of 1932 and 1947 (SMS1932, SMS1947)

Traits	Twin	σ^2b	σ^2w	*t*	SE(*t*)
Height (SMS1947)	SS	0.0033	0.0015	0.69	0.03
	OS	0.0024	0.0021	0.54	0.05
Weight (SMS1947)	SS	14.9	5.64	0.73	0.03
	OS	8.16	8.43	0.49	0.05
BMI (SMS1947)	SS	2.16	0.85	0.72	0.03
	OS	1.0	1.14	0.47	0.06
MHT Score (SMS1932)	SS	153.0	74.2	0.67	0.03
	OS	122.0	89.8	0.58	0.05
MHT Score (SMS1947)	SS	198.2	78.8	0.72	0.03
	OS	153.0	93.1	0.62	0.04
Combined MHT Score	SS	169.9	76.2	0.69	0.02
	OS	135.4	91.0	0.60	0.03

Note: σ^2b, σ^2w are between and within twin pair variances obtained from ANOVA after adjusting for sex and cohort effects (for combined MHT Score); *t* and SE(*t*) are the estimated intraclass correlation and its corresponding approximate standard error.

Table III. Variance components of MHT score and additional traits from univariate analyses using mixture distribution model

Traits	Sex	σ_a^2 (95% CI)	σ_c^2 (95% CI)	σ_e^2 (95% CI)	σ_T^2 (95% CI)	a^2 (95% CI)	c^2 (95% CI)	e^2 (95% CI)
Height (SMS1947)	Male	0.0040 (0.003–0.005)	0.0005 (0–0.001)	0.0002 (0.0001–0.0006)	0.0047 (0.004–0.005)	0.85 (0.70–0.95)	0.10 (0.01–0.23)	0.05 (0.03–0.13)
	Female	0.0029 (0.002–0.004)	0.0017 (0.0004–0.003)	0.0003 (0.0002–0.0005)	0.0049 (0.004–0.006)	0.59 (0.38–0.86)	0.35 (0.08–0.54)	0.06 (0.03–0.11)
	All	0.0038 (0.003–0.005)	0.0007 (0–0.001)	0.0003 (0.0002–0.0004)	0.0048 (0.004–0.005)	0.80 (0.65–0.95)	0.14 (0–0.28)	0.06 (0.04–0.09)
Weight (SMS1947)	Male	11.04 (6.05–16.04)	5.22 (0.55–10.00)	1.54 (0.51–2.94)	17.80 (15.49–20.65)	0.62 (0.34–0.89)	0.29 (0.03–0.53)	0.09 (0.03–0.17)
	Female	15.52 (10.40–19.34)	3.23 (0.27–8.31)	1.29 (0.65–2.57)	20.05 (17.52–23.10)	0.77 (0.52–0.93)	0.16 (0.01–0.39)	0.06 (0.03–0.13)
	All	13.90 (11.06–16.89)	3.62 (0.64–6.58)	1.42 (0.85–2.24)	18.94 (17.08–21.12)	0.73 (0.58–0.90)	0.19 (0.04–0.33)	0.08 (0.04–0.12)
BMI (SMS1947)	Male	2.27 (1.67–2.79)	0.29 (0–0.94)	0.11 (0.06–0.24)	2.67 (2.33–3.08)	0.85 (0.62–0.97)	0.11 (0–0.33)	0.04 (0.02–0.09)
	Female	2.17 (1.58–2.66)	0.28 (0–0.91)	0.14 (0.08–0.26)	2.59 (2.26–2.98)	0.84 (0.61–0.97)	0.11 (0–0.33)	0.05 (0.03–0.10)
	All	2.22 (1.86–2.63)	0.29 (0–0.69)	0.13 (0.08–0.20)	2.63 (2.38–2.93)	0.84 (0.70–0.97)	0.11 (0–0.25)	0.05 (0.03–0.08)
MHT Score (SMS1947)	Male	199.59 (135.85–250.76)	60.53 (19.17–116.28)	19.76 (8.20–45.10)	279.87 (243.77–323.74)	0.71 (0.49–0.87)	0.22 (0.07–0.39)	0.07 (0.03–0.16)
	Female	144.53 (88.31–199.14)	85.70 (31.98–143.50)	16.27 (6.49–32.42)	246.50 (215.26–284.37)	0.59 (0.35–0.81)	0.35 (0.14–0.54)	0.07 (0.03–0.13)
	All	176.56 (136.76–217.29)	68.05 (26.67–110.12)	17.34 (9.50–29.48)	261.95 (235.70–292.80)	0.67 (0.52–0.84)	0.26 (0.11–0.40)	0.07 (0.04–0.11)
MHT Score (SMS1932)	Male	184.46 (115.62–234.45)	32.98 (0–88.25)	20.82 (8.07–45.67)	238.25 (208.33–274.48)	0.77 (0.49–0.95)	0.14 (0–0.36)	0.09 (0.03–0.19)
	Female	151.22 (95.29–199.93)	38.61 (0–90.96)	26.37 (16.82–41.72)	216.20 (190.59–246.88)	0.70 (0.44–0.91)	0.18 (0–0.40)	0.12 (0.08–0.20)
	All	164.82 (121.37–207.99)	36.32 (0–76.36)	25.14 (16.76–37.83)	226.28 (205.15–250.79)	0.73 (0.53–0.91)	0.16 (0–0.32)	0.11 (0.07–0.17)
Combined SMS1932 &	Male	192.10 (150.09–227.88)	46.46 (17.27–83.27)	19.99 (10.89–35.17)	258.55 (234.64–285.94)	0.74 (0.58–0.87)	0.18 (0.07–0.31)	0.08 (0.04–0.14)
SMS1947 MHT Score	Female	148.07 (109.52–185.41)	60.10 (24.75–98.58)	22.34 (15.20–32.20)	230.52 (210.04–253.87)	0.64 (0.47–0.80)	0.26 (0.11–0.41)	0.10 (0.07–0.14)
	All	170.87 (142.20–200.08)	51.00 (22.40–79.55)	21.55 (15.56–29.54)	243.42 (226.43–262.37)	0.70 (0.58–0.83)	0.21 (0.10–0.32)	0.09 (0.06–0.12)

Note: a, c, e, and T are the additive genetic, common and shared environmental and total variance for each trait, respectively; a^2, c^2 and e^2 are the standardized variance for a, c, and e, respectively.

twins is given in Appendix Table I. The estimated proportions of MZ among SS twins were used in the mixture analysis.

Intraclass correlations for SS and opposite-sex (OS) twins for each trait are presented in Table II. For all traits, the intraclass correlations of the SS twins were consistently higher than that of OS twins. This suggests that genetic variation contributed to the individual phenotypic differences.

Univariate analyses

From univariate analyses, the additive genetic, shared environmental and specific environmental variance components and the corresponding 95% confidence interval (CI) of all traits are presented in Table III. There were no significant differences between variance components for MHT scores estimated from the 1932 and 1947 populations ($p > 0.05$). Therefore the combined MHT score data after adjusting for the cohort effect provides more precise estimates of variance components for MHT score. A large proportion of phenotypic variance in MHT score, approximately 70% (95% CI 58%–83%), was attributed to additive genetic effects. In addition, an estimated proportion of 21% (95% CI 10%–32%) of the phenotypic variance was due to common environmental effects shared by twin pairs. These estimates imply a large repeatability ($\sim$0.90) of MHT score in the Scottish population at age 11 in the 1930s and 1940s. There was no significant difference between variance components in males and females for all traits ($p > 0.05$, likelihood-ratio-test). The genetic correlation between males and females was not significantly different from unity for all traits ($p > 0.05$). This suggests that the same set of genes influenced the phenotypes in males and females.

Eighty percent (95% CI 65%–95%) of the phenotypic variation in height was due to genetic effects, under the assumed ACE model. Only 14% (95% CI 0%–28%) of the height variation between individuals was due to common environmental effects shared by twin pairs. A large proportion of variance due to additive genetic effects was also observed for weight (73%, 95% CI 58%–90%) and BMI (84%, 95% CI 70%–97%).

Multivariate analyses

Multivariate analyses of MHT score with height, weight, and BMI were performed on data from the SMS1947 population. A full 4-trait analysis could not be carried out because of the dependence between BMI, height and weight. Therefore, 2 trivariate analyses were performed: height, BMI and MHT score; and weight, BMI and MHT score. From the multivariate analyses, 59% (95% CI 43%–78%) of the phenotypic variance of MHT score of SMS1947 was attributed to genetic variance compared to 67% (95% CI 52%–84%) from the univariate analysis. There was a slightly higher proportion of variance due to common environmental effects in the multivariate analyses (31% compared to 26%). Estimates of correlation coefficients (genetic [rG], common [rC] and specific [rE] environmental and phenotypic [rP]) were derived from the estimated (co)variance components. Estimated phenotypic correlations between the measured traits were consistent with estimates of Pearson's correlations that ignored the twin structure of the data. The estimated genetic correlation coefficients between MHT score and the additional traits (height, weight, BMI) were not significantly different from zero ($p > 0.05$). There was a significant phenotypic correlation coefficient between height and MHT score (0.28 [95% CI 0.21–0.35]; $p < 0.001$), and the common environmental correlation coefficient between height and MHT score was high (0.74 [95% CI 0.29–1.0]). Appendix Table II shows the estimates of all variance components from multivariate analyses, and the corresponding estimates of correlation coefficients are presented in Appendix Table III.

DISCUSSION

This study is the first to estimate genetic and environmental variance components for verbal reasoning ability (MHT score) in entire school-attending populations. Hence, we may assume that there was no bias in parameter estimates due to ascertainment. The estimates for the additive genetic contribution to differences in MHT scores were 70% in the combined univariate analysis, with a 95% CI of 58–83%, and highly consistent across the two whole populations. Estimates of heritability in later-born cohorts of children of similar ages from twin studies with known zygosity vary from 40% to 70% (Bartels *et al.*, 2002; Bishop *et al.*, 2003; Knopik and DeFries, 1998; Plomin *et al.*, 2001).

The heritability estimate of height was consistent with the published estimates using twin designs with

known zygosity (Schousboe *et al.*, 2004; Silventoinen, 2003; Silventoinen *et al.*, 2003). The heritability estimates of weight and BMI in SMS 1947 are slightly higher than estimates from a number of twin studies (Schousboe *et al.*, 2003), but similar to others (Pietilainen *et al.*, 1999). Analyses of these traits have been useful to provide a check on the reliability of the mixture distribution methods. In addition, the similarity of variance components estimated from multivariate analyses compared to univariate analyses indicated that the estimates are precise.

From multivariate analyses, although a significant phenotypic correlation between MHT score and height was estimated, there was no significant evidence for a non-zero genetic correlation. The significant phenotypic correlation between MHT score and height was attributed to common environmental effects, i.e., the same common environmental effects (e.g., SES and better nutrition) influence both height and intelligence. The mixture distribution maximum likelihood method that was used has not been used before to estimate variance components in twin studies without known zygosity. From simulation studies, the estimate of the heritability using a mixture distribution appeared slightly biased upwards (Neale, 2003). However, when new simulations were run with approximately the same population parameters as estimated in this study, the upward bias in the estimate of the heritability was only ~2% (M. Neale, personal communication). The information to separate the distribution of same-sex pair differences (for analysis within pairs) and pair sums (for analysis between pairs) into two underlying distributions comes from the contrast between the variance and kurtosis of the distribution (see Appendix A). Therefore, if the distribution of a trait is strongly kurtotic then the mixture distribution may result in biased results. For the MHT scores considered in this study, kurtosis was estimated as −0.71 (SE 0.15). Although this is significantly different from the expectation under normality, the estimate indicates that the distribution is platykurtic, which is the opposite of what is expected and observed when the differences between SS pair observations are analysed. Hence, the departure from normality is unlikely to have caused a major bias. Nevertheless, the impact of departures from normality on the behaviour of the mixture distribution approach requires further research.

Heath *et al.* (2003) showed that a latent class analysis can be used for zygosity diagnosis. The authors applied their analysis to discrete data from standard questions for zygosity diagnosis and fitted a 2-class latent class model, where the two classes are assumed to correspond to MZ and DZ groupings. In principle this method can be used for any discrete data on twins, and can be viewed as a discrete-trait version of the mixture model for quantitative traits that we have used.

An alternative approach would be to use separate ANOVA for OS and SS data to estimate intraclass correlations and to estimate heritability from these assuming that the mixture proportion in the SS pairs is known (see Appendix B for details). However, although these derivations help to understand and quantify the relationship between the population parameters and estimates in a least squares framework, this method does not use all information efficiently and could lead to severe bias if the male–female genetic correlation deviates from unity.

The proportion of MZ among same-sex twin pairs was assumed to be known in the analysis. A sensitivity analysis showed that varying this parameter from 0.35 to 0.65 has little effect on the estimate of heritability. For the combined analysis of MHT score, the estimate of the heritability for MZ proportions among same-sex twin pairs of 0.35 and 0.65 was 0.71 (95% CI 0.59–0.82) and 0.69 (95% CI 0.54–0.83), respectively.

In this study we fitted the commonly used ACE model. Other three-component parameterisation is also possible, for example a model with additive genetic (*A*), dominance (*D*), and residual environmental (*E*) effects. However, this (ADE) model predicts that MZ correlations are larger than twice the DZ correlations, which is not consistent with the reported OS and SS intraclass correlations in Table II. For all traits, the estimate of the C component was greater than zero, further suggesting that an ACE model is more appropriate than an ADE model.

Our study has wide implications for research into genetic variation of disease and non-disease related traits in human populations. Extremely large random samples, comprising 100,000s of individuals, have been collected or are being collected in a number of countries to answer fundamental questions in the fields of biomedical, educational and economics research. We have shown that, with a minimum of required information (the most important of which are date of birth and sex and a localized identifier such as school or household), a large number of twin pairs can be identified and that appropriate statistical methods are available to estimate genetic parameters without knowing zygosity. Hence, a genetic element

can be added to such studies, thereby greatly enhancing their value.

APPENDIX A. VARIANCE AND KURTOSIS FOR A MIXTURE OF TWO NORMAL DISTRIBUTIONS

For a trait x the i-th moment about the mean is defined as $m_i = E(x - E(x))^i$. The variance and kurtosis are m_2 and $K = \{m_4/(m_2)^2 - 3\}$, respectively. For a single distribution of a normally distributed trait, $m_2 = \sigma^2$, $m = 3\sigma^4$ and $K = 0$. For a mixture of two distributions with mixture proportion p,

$$m_2 = p\sigma_1^2 + (1-p)\sigma_2^2 \text{ and}$$

$$m_4 = 3p\sigma_1^4 + 3(1-p)\sigma_2^4$$

These moments also apply to within-pair variances in a twin design. For example, consider the difference in observations between pairs of twins in a mixture of DZ and MZ twin pairs with mixture proportion 0.5 and an ACE model with heritability of 0.6 and the proportion of variance due to common environmental effects of 0.2. The phenotypic variance is unity. Then, $m_2 = 0.350$, $m_4 = 0.435$ and $K = 0.55$. The principle of the mixture distribution approach is that the two unknown variances are estimated from the observed variance and kurtosis.

APPENDIX B. ESTIMATION OF PARAMETERS FROM ANOVA ON OS AND SS PAIRS

Consider the ACE model and scaled phenotypes so that the phenotypic variance is unity. Let p be the proportion of MZ twins among same-sex (SS) twin pairs. The between (B) and within (M) mean square component in an ANOVA will then be a mixture from two distributions (MZ and DZ). The expected values are,

$$E(B_{SS}) = e^2 + 2c^2 + \frac{1}{2h^2}(3+p) = 1 + c^2 + \frac{1}{2h^2}(1+p)$$

$$E(W_{SS}) = e^2 + \frac{1}{2}(1-p)h^2 = 1 - c^2 - \frac{1}{2}h^2(1+p),$$

with h^2, c^2 and e^2 the proportion of phenotypic variance due to additive genetic (A), common environmental (C) and residual environmental (E) effects. The intra-class correlation (t) from the between and within-pair analysis is,

$$\begin{aligned} t &= [(E(B) - E(W))/2]/[E((B) - E(W))/2 + E(W)] \\ &= [E(B) - E(W)]/[E(B) + E(W)]. \end{aligned}$$

For the SS pairs,

$$t_{SS} = c^2 + \frac{1}{2}(1+p)h^2.$$

For the opposite-sex (OS) pairs, assuming an ACE model but allowing for a genetic correlation of less than unity between the sexes and different heritabilities for males (m) and females (f),

$$t_{OS} = c^2 + \frac{1}{2}r_g h_m h_f.$$

If the heritability for males and females is the same then,

$$t_{OS} = c^2 + \frac{1}{2}r_g h^2.$$

Hence, under the assumption of equal heritabilities of males and females, there are two summary statistics (correlations), i.e. t_{SS} and t_{OS}, but three unknowns (c^2, h^2, r_g). It follows that,

$$2(t_{SS} - t_{OS}) = h^2[(1+p) - r_g], \quad \text{and}$$

$$[t_{OS}(1+p) - t_{SS}r_g]/[(1+p) - r_g] = c^2,$$

If we further assume that $r_g = 1$, then the estimates of h^2 and c^2 satisfy

$$h^2 = 2(t_{SS} - t_{OS})/p, \text{ and}$$

$$c^2 = [t_{OS}(1+p) - t_{SS}]/p.$$

Relative to the standard twin design with MZ and DZ pairs, the sampling variance of the estimate of the heritability from SS and OS pairs is increased by a factor of p^{-2}, for example by a factor of four if $p = 1/2$.

Appendix Table I. Estimated proportion of monozygotic (MZ) twin pairs

Cohort	SS pairs	OS pairs	Total	pMZ (SE)	pMZ|SS (SE)
SMS1932	382	190	572	0.34 (0.04)	0.50 (0.04)
SMS1947	320	197	517	0.24 (0.04)	0.38 (0.06)
Combined	702	387	1089	0.29 (0.02)	0.45 (0.03)

Note: pMZ and pMZ|SS are the estimated proportion of MZ twins in the population of twin pairs and the proportion of MZ twins in SS twin pairs, respectively. pMZ was estimated using the formula 1–2 (proportion of OS twin pairs). pMZ|SS was estimated as pMZ/(proportion of SS twin pairs).

Appendix Table II. Variance components obtained from multivariate analysis of SMS 1947

Traits	σ_a^2 (95% CI)	σ_c^2 (95% CI)	σ_e^2 (95% CI)	σ_T^2 (95% CI)	a^2 (95% CI)	c^2 (95% CI)	e^2 (95% CI)
Height	0.0035 (0.0028–0.0042)	0.0009 (0.0002–0.0016)	0.0004 (0.0003–0.0006)	0.0048 (0.0043–0.0053)	0.74 (0.59–0.88)	0.18 (0.05–0.32)	0.08 (0.06–0.13)
Weight	13.30 (10.39–16.32)	3.96 (1.03–6.89)	1.72 (1.19–2.58)	18.97 (17.14–21.12)	0.70 (0.55–0.86)	0.21 (0.06–0.35)	0.09 (0.06–0.14)
BMI	2.21 (1.86–2.59)	0.29 (0–0.70)	0.14 (0.09–0.21)	2.64 (2.39–2.94)	0.84 (0.70–0.96)	0.11 (0–0.25)	0.05 (0.04–0.08)
MHT Score	157.05 (114.70–199.37)	81.60 (39.66–124.71)	26.04 (16.80–39.66)	264.69 (238.04–296.06)	0.59 (0.43–0.78)	0.31 (0.15–0.44)	0.10 (0.06–0.15)

Note: a, c, e, and T are the additive genetic, common and shared environmental and total variance for each trait, respectively; a^2, c^2 and e^2 are the standardized variance for a, c, and e, respectively. A full 4-trait analysis could not be carried out because of the dependence between BMI, Height and Weight. The results for Height, BMI and MHT Score are from a trivariate analysis of these traits. The results for Weight are from a trivariate analysis of Weight, BMI and MHT Score.

Appendix Table III. Phenotypic (rP), additive genetic (rG) common (rC) and specific environmental (rE) correlations estimated from multivariate analysis of SMS 1947, and their 95% confidence interval (CI)

Traits	rP (95% CI)	rA (95% CI)	rC (95% CI)	rE (95% CI)
Height and MHT Score	0.28 (0.21–0.35)	0.15 (−0.01–0.32)	0.74 (0.29–1.0)	0.06 (−0.21–0.31)
Weight and MHT Score	0.19 (0.11–0.26)	0.09 (−0.08–0.25)	0.47 (0.04–0.99)	0.13 (−0.13–0.37)
BMI and MHT Score	−0.01 (−0.09–0.06)	−0.003 (−0.15–0.14)	−0.13 (−1.0–1.0)	0.17 (−0.09–0.42)
Height and Weight	0.71 (0.67–0.74)	0.65 (0.56–0.72)	0.88 (0.45–1.0)	0.85 (0.75–0.91)
Height and BMI	−0.02 (−0.09–0.06)	−0.13 (−0.27–0.01)	0.43 (−1.0–1.0)	0.33 (0.05–0.56)
Weight and BMI	0.69 (0.65–0.73)	0.68 (0.61–0.74)	0.79 (−1.0–1.0)	0.73 (0.59–0.84)

Note: A full 4-trait analysis could not be carried out because of the dependence between BMI, Height and Weight. Rows 1, 3 and 5 are from a trivariate analysis of Height, BMI and MHT Score. Rows 2 and 6 are from a trivariate analysis of Weight, BMI and MHT Score, and row 4 is from a trivariate analysis of Weight, Height and MHT Score.

ACKNOWLEDGMENTS

We thank: Alison Pattie for collating and checking twin data from the SMS1947; Michael Neale for advice on the statistical analyses; Carole Hart with assistance in identifying twins from the SMS1932; Allan McRae for helping to create Figure 1; and Bill Hill and Ben Neale for helpful comments and discussions. Ian Deary is the recipient of a Royal Society-Wolfson Research Merit Award. Lawrence Whalley holds a Wellcome Trust Career Development Award. Peter Visscher holds a BBSRC Research Development Fellowship. We thank the referees for helpful comments.

REFERENCES

Bartels, M., Rietveld, M. J., Van Baal, J. C., and Boomsma, D. I. (2002). Heritability of educational achievement in 12-year-olds and the overlap with cognitive ability. *Twin Res.* **5**:544–553.

Bishop, E. G., Cherny, S. S., Corley, R., Plomin, R., DeFries, J. C., and Hewitt, J. K. (2003). Development genetic analysis of general cognitive ability from 1 to 12 years in a sample of adoptees, biological siblings, and twins. *Intelligence* **31**:31–49.

Bouchard, T. J. (2004). Genetic and environmental influences on human psychological differences. *Curr. Dir. Psychol. Sci.* **13**:140–144.

Bouchard, T. J., Lykken, D. T., McGue, M., Segal, N. L., and Tellegen, A. (1990). Sources of human psychological differences: the Minnesota Study of Twins Reared Apart. *Science* **250**:223–228.

Bouchard, T. J., and McGue, M. (2003). Genetic and environmental influences on human psychological differences. *J. Neurobiol.* **54**:4–45.

Carroll, J. B. (1993). *Human Cognitive Abilities: A Survey of Factor Analytic Studies*. Cambridge, UK: Cambridge University Press.

Deary, I. J., Whalley, L. J., Lemmon, H., Crawford, J. R., and Starr, J. M. (2000). The stability of individual differences in mental ability from childhood to old age: follow-up of the 1932 Scottish Mental Survey. *Intelligence* **28**:49–55.

Deary, I. J., Whiteman, M. C., Starr, J. M., Whalley, L. J., and Fox, H. C. J. (2004). The impact of childhood intelligence on later life: following up the Scottish Mental Surveys of 1932 and 1947. *J. Pers. Soc. Psychol.* **86**:130–147.

Gottfredson, L. S., and Deary, I. J. (2004). Intelligence predicts health and longevity: but why? *Curr. Dir. Psychol. Sci.* **13**:1–4.

Gray, J. R., and Thompson, P. M. (2004). Neurobiology of intelligence: science and ethics. *Nat. Rev. Neurosci.* **5**:471–482.

Health, A. C., Nyholt, D. R., Neuman, R., Madden, P. A. F., Bucholz, K. K., Todd, R. D., Nelson, E. C., Montgomery, G. W., and Martin, N. G. (2003). Zygosity diagnosis in the absence of genotypic data: an approach using latent class analysis. *Twin. Res.* **6**:22–26.

Humphreys, L. G., Davey, T. C., and Park, R. K. (1985). Longitudinal correlation analysis of standing height and intelligence. *Child. Dev.* **56**:1465–1478.

Imaizumi, Y. (2003). A comparative study of zygotic twinning and triplet rates in eight countries, 1972–1999. *J. Biosoc. Sci.* **35**:287–302.

Johnson, F. W. (1991). Biological factors and psychometric intelligence: a review. *Genet. Soc. Gen. Psychol. Monogr.* **117**:313–357.

Joseph, J. (2003). *The Gene Illusion: Genetic Research in Psychiatry and Psychology Under the Microscope*. Ross-on-Wye, UK: PCCS.

Knopik, V. S., and DeFries, J. C. (1998). A twin study of gender-influenced individual differences in general cognitive ability. *Intelligence* **26**:81–89.

Mehrotra, S. N., and Maxwell, J. (1949). The Intelligence of twins: a comparative study of eleven-year-old twins. *Popul. Stud.* **3**:295–302.

Molinari, E., Sartori, A., Ceccarelli, A., and Marchi, S. (2002). Psychological and emotional development, intellectual capabilities, and body image in short normal children. *J. Endocrinol. Invest.* **25**:321–328.

Neale, M. C. (2003). A finite mixture distribution model for data collected from twins. *Twin Res.* **6**:235–239.

Neale, M. C., Boker, S. M., Xie, G., and Maes, H. H. (2002). *Mx: Statistical Modelling* (6th ed.). Richmond, VA: Department of Psychiatry.

Neale, M. C., and Cardon, L. R. (1992). *Methodology for Genetic Studies of Twins and Families*. Dordrecht: Kluwer.

Neisser, U., Boodoo, G., Bouchard, T. J., Boykin, A. W., Brody, N., Ceci, S. J., Halpern, D. F., Loehlin, J. C., Perloff, R., Sternberg, R. J., and Urbina, S. (1996). Intelligence: knowns and unknowns. *Am. Psychol.* **51**:77–101.

Pietilainen, K. H., Kaprio, J., Rissanen, A., Winter, T., Rimpela, A., Viken, R. J., and Rose, R. J. (1999). Distribution and heritability of BMI in Finnish adolescents aged 16y and 17y: a study of 4884 twins and 2509 singletons. *Int. J. Obes. Relat. Metab. Disord.* **23**:107–115.

Plomin, R., and Spinath, F. (2004). Intelligence: genetics, genes, and genomics. *J. Pers. Soc. Psychol.* **86**:112–129.

Plomin, R., DeFries, J. C., McClearn, G. E., and McGuffin, P. (2001). *Behavioral Genetics* (4th ed.). New York: Worth.

Scarr-Salapatek, S. (1971). Race, social class, and IQ. *Science* **174**:1285–1295.

Schousboe, K., Visscher, P. M., Erbas, B., Kyvik, K. O., Hopper, J. L., Henriksen, J. E., Heitmann, B. L., and Sørensen, T. I. A. (2004). Twin study of genetic and environmental influences on adult body size, shape and composition. *Int. J. Obes.* **28**:39–48.

Schousboe, K., Willemsen, G., Kyvik, K. O., Mortensen, J., Boomsma, D. I., Cornes, B. K., Davis, C. J., Fagnani, C., Hjelmborg, J., Kaprio, J., de Lange, M., Luciano, M., Martin, N. G., Pedersen, N., Pietilainen, K. H., Rissanen, A., Saarni, S., Sørensen, T. I. A., van Baal, G. C. M., and Harris, J. K. (2003). Sex differences in heritability of BMI: a comparative study of results from twin studies in eight countries. *Twin Res.* **6**:409–421.

Scottish Council for Research in Education. (1933). *The Intelligence of Scottish Children*. London: University of London Press.

Scottish Council for Research in Education. (1949). *The Trend of Scottish Intelligence*. London: University of London Press.

Scottish Council for Research in Education. (1953). *Social Implications of the 1947 Scottish Mental Survey*. London: University of London Press.

Silventoinen, K. (2003). Determinants of variation in adult body height. *J. Biosoc. Sci.* **35**:263–285.

Silventoinen, K., Sammalisto, S., Perola, M., Boomsma, D. I., Cornes, B. K., Davis, C., Dunkel, L., Lange, M., Harris, J. R., Hjelmborg, J. V. B., Luciano, M., Martin, N. G., Mortensen, J., Nistico, L., Pedersen, N. L., Skytthe, A., Spector, T. D., Stazi, M. A., Willemsen, G., and Kaprio, J. (2003). Heritability of adult body height: A comparative study of twin cohorts in eight countries. *Twin Res.* **6**:399–408.

Edited by Stacey Cherny