Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Personality and Individual Differences 39 (2005) 497–509
PERSONALITY AND INDIVIDUAL DIFFERENCES
www.elsevier.com/locate/paid

Sex-related differences in general intelligence *g*, brain size, and social status

Helmuth Nyborg *

Department of Psychology, Research Unit for Differential Psychology, University of Aarhus, 8000 Aarhus C, Denmark

Received 24 November 2003; received in revised form 9 March 2004; accepted 7 December 2004
Available online 8 June 2005

Abstract

The question of a sex difference in intelligence has long divided the experts. IQ researchers sum standardized subtest scores to calculate *intelligence in general*, and find that males outscore females by about 3.8 points, whereas factor analysts derive the *g* factor scores from intertest-correlations and find no consistent sex differences in *general intelligence*. The latter finding is puzzling, as males have larger average brains than females, and brain size correlates .30–.45 with *g* (and IQ). Males thus "ought" to score a higher *g* than females.

The present study addressed this paradox by testing four hypotheses: (1) Inadequate analyses explain why researchers get inconsistent results, (2) The proper method will identify a male *g* lead, (3) The larger male brain "explains" the male *g* lead, (4) The higher male *g* average and wider distribution transform into an exponentially increased male–female ratio at the high end of the *g* distribution, and this largely explains male dominance in society.

All four hypotheses obtained support and explain in part why relatively few males dominate the upper strata in all known societies. The confirmation of hypothesis 3 suggests that the brain size—intelligence–dominance link may be partly biological.

Keywords: Sex-related differences; General intelligence; *g*; Brain size; Social status

* Tel.: +45 8768 0456/+45 8942 4900; fax: +45 8942 4900.
E-mail address: helmuth@psy.au.dk

0191-8869/$ - see front matter
doi:10.1016/j.paid.2004.12.011

1. Introduction

Experts have long disagreed about the existence of a sex difference in overall intelligence. Some (e.g. Lynn, 1994, 1997, 1999; Lynn, Irwing, & Cammock, 2002) find that males outscore females by about 3.8 IQ points, but most find no sex difference (e.g. Brody, 1992; Halpern & LaMay, 2000; Jensen, 1998).

This disagreement is confusing for theoretical reasons. First, males dominate all higher ranks of education, research, occupation, and political power structures that call for capacity to deal with complexity, which is just another way of defining *general intelligence g*. Second, males have, on average, larger brains than females, and brain size correlates positively with intelligence. Yet, the empirical evidence for a male *g* advantage is equivocal.

The present study addresses these paradoxes by testing four hypotheses: (1) Ambiguous definitions and methods explain the current disagreement among the experts, (2) The proper analytic approach will identify a male lead in *general intelligence g*, (3) The larger male brain partly explains the average male *g* lead, and, (4) Classical Distribution Theory illustrates how a small male mean SD *g* score advantage and a wider male SD dispersion score translate into an exponentially increased male–female ratio at the high end of the *g* distribution. This unequal ratio of high *g* males to females explains in part why males (always, according to Goldberg, 1977) dominate the intellectually most demanding top occupational and political strata.

2. Hypothesis 1: Ambiguous definitions of intelligence and inadequate use of analytic methods explain the empirical inconsistency

2.1. The IQ position

Lynn (1994, 1997, 1999) claims that males ought to outscore females in IQ in terms of the following two syllogisms (Nyborg, 2002). First, brain volume correlates with IQ. Males have on average larger brains than females. Ergo: Males have a higher mean IQ than females. Second, job status and income correlate with IQ. Males have on average higher job status and income than females. Ergo: Males have a higher mean IQ than females. Lynn averaged several empirical studies, and found a male lead of 3.8 IQ points.

The problem with this is, however, that the total summed IQ score is sensitive to test item bias. Females will be favoured by an overweight of subtests tapping verbal abilities, and males by a spatial ability subtest bias. This means that a sex difference in *intelligence in general*, that is IQ, may reflect a test bias or a real sex difference, but by just summing up standardized subtest scores we would never know the difference (Jensen, 1998).

2.2. The *g* position

A methodologically better approach is to factor analyse the intertest-correlations among subtests and derive *g*, which reflects *general intelligence* that shows higher reliability and validity than IQ scores. However, factor analysis gives inconsistent results: Females outscore males in some studies, males do better in other studies, and the remaining studies show no sex difference in *g* (Aluja-Fabregat, Colom, Abad, & Juan-Espinosa, 2000; Colom, García, Juan-Espinosa, & Abad, 2002; Colom & García-López, 2002; Colom, Juan-Espinosa, Abad, & García, 2000; Jensen, 1998). The reason for this inconsistency is methodological: *g* factor scores can be contaminated to some extent by group factors and test specificity (Jensen, 1998; Nyborg, 2001, 2002, 2003). This is usually a problem only where the task is to single out a small group (e.g. sex) difference, and where even a slight contamination of *g* may either drown a real difference and lead to a Type II error, or may erroneously indicate a male or a female superiority and lead to a Type I error. Jensen (1998) thus factor analysed five separate studies with inclusive test batteries and representative samples. He found a female Principal Factor (PF1) *g* lead in one test, a male PF1 *g* lead in another test battery, and no sex differences in the remaining studies (Jensen, 1998, pp. 538–543). He concluded that sex differences in *g* are either totally nonexistent or of uncertain directions and of inconsequential magnitude. Observed differences must be attributable to factors other than *g*.

However, the female *g* lead disappeared after Jensen eliminated the unusually large number of test items favouring females in the General Aptitude Test Battery, and repeated the factor analysis. In other words, the female *g* superiority was an artefact due to test bias that favoured females. Obviously, male test bias would erroneously produce male *g* superiority, and the practice of averaging contaminated *g*s from various studies is a non sequitur.

This supports hypothesis 1 that the disagreement among experts about the existence of a sex difference emanates from conceptual and analytic difficulties.

3. Hypothesis 2: The proper analytic approach will identify a male lead in *general intelligence g*

3.1. Introduction

Jensen's (1998) review of studies on sex differences further indicated problems with sampling, fewer than the minimum of 9 tests needed for proper factor analysis, and too little internal diversity to tap into a wide range of abilities.

This lead Nyborg (2003, pp. 199–201) to develop a rough and ready grading scale for ranking current studies. Briefly, a study earns one point for unbiased sampling, one for incorporating nine or more tests in the battery, one for assuring sufficient test diversity, one for applying a hierarchical factor analysis (HFA) with an orthogonal Schmid–Leiman transformation (S–L; Schmid & Leiman, 1957), one for inserting the point–biserial correlations (see later) into the intertest-correlation matrix for co-factoring and, finally, one point for testing whether sex loads statistically significantly on *g*. No points are given for the inclusion of correlated vector analysis in sex difference studies, for reasons given later. Studies of sex differences in *g* earning <5 points are deemed untrustworthy, as the risk of committing Type I or Type II errors is too large.

There are only two current studies that earn the maximum 6 points on the quality scale: The present study, and one by Colom et al. (2002). These two studies are analysed here to critically test hypothesis 2, which says, that the proper analytic approach will identify a male lead in *general intelligence g*.

3.2. The present study

3.2.1. Method

An exceptionally careful sample selection began in the late 1970s with a computer search in the Danish Folkeregister for every twentieth child that was either 8, 10, 12, 14 or 16 years old (±0.5 year), was either a boy or a girl, and was attending an elementary school in the Skanderborg communal district, situated either at the countryside, in a suburb, or in a larger city. Information about the socio-economic status of the parents, defined by father's occupational status, was also collected and categorized at five levels. If the twentieth child, or its parents, refused to participate in the 20+ years cohort-sequential study, the twenty-first (or in two cases the twenty-third) child on the computer list was invited. No particular pattern of reasons for refusing to participate could be spotted in retrospect. Five age group categories, 8, 10, 12, 14, or 16+ year, were established on the basis of this preliminary search protocol. The age group distributions of sex, geographical characteristics, and socio-economic background of the children were inspected when about 50% of the children were tested, and the age categories were then supplemented to full capacity, so that each age category finally included a total of 15 boys and 15 girls. Great care was taken to ensure that all categories ended up being representative with respect to the general Danish socio-economic population distribution while also conforming to the nationwide proportional representation of rural, suburban or city residency and school attendance. The total cohort-sequential study included two cross-sectional phases—1976 (including only some cognitive testing) and 1988–2000+ (full study), and a longitudinal study spanning 1976–2000+, which included two repeatedly tested groups of 30 boys and girls, each, and a control group of 30 boys and girls that were tested at age 10 and 16+ years, only. The first part of the present analysis is based on the sub-sample of 31 males (mean age 17.4, SD = 1.8) and 31 females (mean age 17.3, SD = 1.9) for whom WAIS and all the other test data were available. Later analyses include children as well (Section 5).

3.2.2. Test battery

The 62 subjects went through a very large battery of 20 ability tests, differing widely in content area. Negative scores were inverted. Briefly, the tests are (1) The Rod-and-Frame test (RFT) Frame dependence (signed errors, inverted: Nyborg & Isaksen, 1974), (2) RFT Response Variability (inverted: Nyborg & Isaksen, 1974), (3) RFT Field Dependence (unsigned errors, inverted: Asch & Witkin, 1948), (4) The Embedded-Figures Test (seconds/figure, inverted: Witkin, 1950), (5) The Money Left–Right Discrimination Test (errors, inverted: Money, 1965), (6) Mental Rotation Test (nos. of figures found, corrected for guessing: Vandenberg & Kuse, 1978), (7) Tapping Test, Left Hand (max. nos. of taps during 2×30 s), (8) Tapping Test, Right hand (max. nos. of taps during 2×30 s), (9) Oral Fluency (name as many animal names as possible, beginning with F within a minute), (10–20) All 11 Wechsler Adult Intelligence Scale raw scores (WAIS: Wechsler, 1958), that is, (10) WAIS Information, (11) WAIS Comprehension, (12) WAIS Arithmetic, (13) WAIS Similarities, (14) WAIS Digit Span, (15) WAIS Vocabulary, (17) WAIS Digit Symbol, (18) WAIS Picture Completion, (19) WAIS Block Design, (20) WAIS Picture Arrangement, and (21) WAIS Object Assembly.

3.2.3. Analysis

A preliminary PC analysis of the separate male and female data was first performed to check for identical factor structure in male and female data. The congruence coefficient was fairly close to unity (.92), which indicates virtual identity in the *g* factor structure for males and females.

According to Jensen (1998, p. 538), "The best method for determining the sex difference in psychometric *g*..." is to first fit the *d* effects (the sex difference on each subtest divided by the averaged male–female standard deviations into the formula (Jensen, 1998, p. 542, note 9), to calculate the point–biserial correlations (r_{pbs}), indicating the extent to which sex, as a dichotomous variable, loads on the metric sex differences), and then to insert the (twenty) r_{pbs} in the subtest inter-correlation matrix, and factor analyse them together with a very large number of highly varied tests. It is empirically established that the inclusion of the sex r_{pbs} in the correlation matrix has no effect on the factor structure and only negligible effects on the subtests' *g* loadings (Jensen, 1998, p. 542, note 9). The congruence coefficient for with–without inserted sex r_{pbs} was .9999.

The present study used an HFA analysis with the Schmid–Leiman (SL) transformation. This HFA/SL analysis provides estimates of test specificity at the lowest level, of group factors at the next level, and of general intelligence *g* at the highest second or third order level. This higher-order *g* factor shows high reliability and heritability, little dimensional contamination, has many close biological and brain correlates, and shows better predictive validity than other estimates of overall intelligence (Jensen, 1998). The predicted loading of sex on *g* was tested for one-sided significance at the 2.5% level.

A correlated vector analysis (Jensen, 1998, Appendix B) was also done, even if this type of analysis is a priori deemed unsuitable for the study of sex differences (Nyborg, 2003).

3.2.4. Results

The HFA/SL analysis permitted extraction of one second-order *g* factor and six first-order group factors.

Table 1 first outlines the type of tests used. Column two gives the observed sex difference on each test (*d* effects, where a minus sign indicates female superiority) and the average effect size and its average IQ equivalent. Third column gives the r_{pbs} (adjusted for unequal SDs; see Jensen, 1998, p. 542, note 9). The fourth column shows the tests' *g* loadings, and the last row gives the average *g* load.

As predicted, co-factored sex loads positively ($r_{\text{pbs}} = .274$) on the *g* factor dimension ($p = .03$, one-sided).

The purpose of the correlated vector calculation is to see whether a (sex) difference in *g* is a "Jensen Effect", that is, if the sizes of *d* differences in the vector of the 20 tests correlate with the sizes in the vector of the respective tests' *g* loadings. Jensen (1998) notes that this method works against finding statistically significant relationships by reducing the degrees of freedom to the number of tests. The *g* factor loadings of the 20 tests in Table 2 were corrected for attenuation by dividing each subtest's *g* loading by the square root of that subtest's reliability coefficient, as were the *d* values. This again works against the hypothesis of a statistically significant sex load on *g*, because it removes some part of the *g* vector correlation with the *d* vector, but it is preferable to no control at all (Jensen, 1998). The reliability coefficients for the ten WAIS tests were taken from Jensen (1980).

Table 1
Type of test, d effects, point–biserial correlations, and g loadings for 20 variables with eigenvalues >1

Tests	Effect d	Point–biserial correlation r_{pbs}	g Loading secondary factor
RFT Frame Dependence (signed errors, inverted)	0.39	0.20	0.37
RFT Response Variability (errors, inverted)	0.40	0.20	0.47
RFT Field Dependence (unsigned errors inverted)	0.36	0.18	0.41
Embedded-Figures Test (seconds/figure inverted)	0.21	0.10	0.53
Money left–right Discrimination Test (errors inverted)	0.52	0.27	0.61
Mental Rotation (Figures found, corrected for guessing)	0.41	0.21	0.46
Tapping (Left hand)	0.58	0.30	0.31
Tapping (Right hand)	0.30	0.15	0.35
Oral fluency	−0.08	−0.04	0.23
WAIS Information	0.42	0.21	0.55
WAIS Comprehension	−0.22	−0.11	0.39
WAIS Arithmetic	0.13	0.06	0.47
WAIS Similarities	0.34	0.17	0.46
WAIS Digit Span	0.16	0.08	0.23
WAIS Vocabulary	0.35	0.18	0.47
WAIS Digit Symbol	−0.54	−0.28	0.01
WAIS Picture Completion	0.42	0.22	0.40
WAIS Block Design	0.08	0.04	0.60
WAIS Picture Arrangement	0.04	0.02	0.35
WAIS Object Assembly	−0.06	−0.03	0.46
Point–biserial factor loading of sex			0.274*
Average effect size	0.21		
Average IQ equivalent	3.15		
Average factor loading			0.41

Point–biserial correlations (adjusted for unequal SDs) were factored in to reflect the loading of sex on the g-dimension. $N = 31$ males (mean age 17.4, SD = 1.8) and 31 females (mean age 17.3, SD = 1.9).
* Significant at p (one-sided) = .016.

The predicted Pearson uncorrected and corrected correlations of g with d vectors are .570 ($p = .005$) and .589 ($p = .003$) (both one-sided), respectively. However, outliers or other peculiarities in the g factor loading or d scales could produce a bias that would remain unrecognised unless also a Spearman's rank-order correlation (r_S) is calculated, and its size compared to that of the Pearson r. The uncorrected and corrected r_S are .319 ($p = .085$) and .412 ($p = .035$; one-sided), respectively. The discrepancy between r and r_S points to a bias, and the just about significant Spearman r_S suggests that the observed sex difference is a weak Jensen effect. However, as mentioned previously, the degrees of freedom are restricted to the number of tests, not to the number of subjects, and this makes the correlated vector analysis vulnerable to a Type II error, that is, to rejecting a real sex load on g.

3.3. The Colom et al. (2002) study

There is only one other study in the entire literature on general intelligence, taking a similar high quality methodological approach (Colom et al., 2002). The study included a large sample

Table 2
Correlated vector analysis based on hierarchical g loadings and average sex differences (d) for 31 male and 31 female adult subjects on 20 tests

Column	A	B	C	D	E	F	G	H	I	J
	Factor loadings					Mean sex differences				
	Uncorrected			Corrected		Uncorrected			Corrected	
Test	g	Rank	Reliability	g	Rank	d	Rank	Reliability	d	Rank
RFT Frame Dependence (signed errors inverted)	0.373	7	0.867	0.401	5	0.391	14	0.867	0.424	14
RFT Response Variability (errors inverted)	0.465	14	0.658	0.573	14	0.398	15	0.658	0.497	15
RFT Field Dependence (unsigned errors inverted)	0.413	10	0.892	0.437	7	0.361	13	0.892	0.386	11
Embedded-Figures Test (seconds/figure inverted)	0.530	17	0.475	0.769	20	0.208	9	0.475	0.301	9
Money Left–Right Discrimination Test (inverted)	0.611	20	0.785	0.690	18	0.518	19	0.785	0.601	18
Mental Rotation (Figures found, corrected f. guess.)	0.460	12	0.586	0.600	15	0.408	16	0.586	0.540	17
Tapping (Left hand)	0.312	4	0.732	0.365	4	0.578	20	0.732	0.700	19
Tapping (Right hand)	0.347	5	0.634	0.435	6	0.300	10	0.634	0.378	10
Oral fluency	0.228	3	0.675	0.277	3	−0.080	3	0.675	−0.097	3
WAIS Information	0.546	18	0.718	0.644	16	0.421	17	0.718	0.504	16
WAIS Comprehension	0.387	8	0.596	0.501	9	−0.221	2	0.596	−0.285	2
WAIS Arithmetic	0.466	15	0.738	0.542	11	0.127	7	0.738	0.147	7
WAIS Similarities	0.458	11	0.795	0.514	10	0.343	11	0.795	0.387	12
WAIS Digit Span	0.226	2	0.664	0.277	2	0.160	8	0.664	0.196	8
WAIS Vocabulary	0.474	16	0.730	0.555	12	0.349	12	0.730	0.412	13
WAIS Digit Symbol	0.007	1	0.518	0.009	1	−0.545	1	0.518	−0.780	1
WAIS Picture Completion	0.404	9	0.318	0.717	19	0.422	18	0.318	0.758	20
WAIS Block Design	0.596	19	0.747	0.690	17	0.085	6	0.747	0.098	6
WAIS Picture Arrangement	0.351	6	0.590	0.457	8	0.040	5	0.590	0.051	5
WAIS Object Assembly	0.461	13	0.660	0.567	13	−0.062	4	0.660	−0.076	4

Correlations	A with F	D with I
Pearson r (one-tailed)	.570; p = .005	.589; p = .003
Spearman rank-order r_S (one-tailed)	.319; p = .085	.412; p = .035

Illustrating the extent to which the sex differences load on the g dimension. Reliability coefficients for the first ten tests are the lower-bound estimate of each test's communality. Reliability coefficients for the WAIS sub-tests were taken from Jensen (1980).

of 703 females and 666 males, participating in the Spanish standardization of the WAIS-III test, and found a male IQ advantage of 3.6 points in *intelligence in general*—close to the average 3.8 IQ points found by Lynn (1994, 1997) and to the 3.15d equivalent IQ points in the present study (Table 1). More importantly, Colom and colleagues fitted the r_{pbs} among each subtests' score and the dichotomous sex variable into the matrix of sub-test inter-correlations, performed a HFA/SL type factor analysis of the full matrix, and found that sex loaded .159 on g. Unfortunately, this value was combined with other sex-loaded g values that cancelled out each other. The conclusion: "Null sex difference in general intelligence".

However, a statistical test of the observed sex load on g gives a highly significant male lead in g (N = 1,369; $p < .0001$, Fisher z = .16036; Nyborg, 2003). This confirms the existence of a significant sex difference in *general intelligence* g, even in the WAIS-III test battery that was originally derived of test items with a large sex bias.

3.4. Conclusion

The only two existing high quality studies verify independently, that there is a significant average g difference in male favour, thus confirming hypothesis 2.

4. Hypothesis 3: The larger male brain partly explain the male lead in *g*

4.1. Introduction

Many studies (Anderson, 2003; Ankney, 1992, 1995; Gignac, Vernon, & Wickett, 2003; Haier, 2003; Lynn, 1994, 1999; Rushton, 1992; Rushton & Ankney, 1996) demonstrate that males have, on average, a larger brain (d = .30–.35) that contains about 15% more neurons than the female brain (Packenberg & Gundersen, 1997). Moreover, overall brain size correlates .10–.45 with IQ. The marked differences in correlations reflect measurement error plus the type of measure used—from rough brain size by taping head circumference to precisely scanned brain volume. The rule seems to be: The more exact the measure, the higher the correlation.

Lynn used the sex difference in brain size to predict the male IQ lead, and failed only by a mere −.2 IQ point (see Table 3).

Given a male lead in brain size of d = .78, and a mean correlation of .35 between brain image size and IQ (Rushton & Ankney, 1996, though Gignac et al., 2003, suggest a mean size–IQ correlation of .40, and which Schoenemann, Budinger, Sarich, & Wang, 2000, raises further to .45 for the brain volume–g relationship), the multiplication of the male lead in brain size with the brain size–IQ correlation gives an SD of .27, which, when multiplied by 15 translates into a male lead of 4.05 IQ points. In other words, Lynn's theoretical prediction of intelligence from brain size matches the empirically observed male average IQ lead of 3.85 quite well. The major problem with this calculation is, that it is based on IQ scores or *intelligence in general*, rather than on the more precise and less contaminated HFA/SL *general intelligence* g measure.

Table 3
Prediction of sex differences in *intelligence in general* IQ or *general intelligence g* from observed sex differences in scanned brain volume, or from head circumference, which is a rough proxy for brain size (re-calculated from Nyborg, 2001, 2002, 2003)

Study	A: Observed sex difference in brain volume[a] or circumference[b]	B: Correlation between volume[a] or circumference[b] and IQ[1] or *g* factor score[2]	C: A × B	D: Predicted male lead in IQ[3] or *g* converted to IQ[4]	E: Observed male lead in IQ[5], *d* equivalent IQ[6] (Table 1) or *d*[7] (based on sex loaded r_{pbs})
	d	*r*	SD units	C × 15	
Intelligence in general IQ data (Lynn, 1999, average over several studies)	.78[a]	.35[a1]	.27	4.05[3]	3.85[5]
General intelligence g data (Nyborg, 2001, 2002, 2003, Fig. 10.2)	.90[b]	.21[b2]	.19	2.84[4]	.315[6], or .57[7] (IQ 8.55)

Note: [7]After the r_{pbs} was factored in on the *g* dimension to indicate how much sex loads on the *g* factor (.274; see Table 1), the formula given by Jensen (1998, p. 543, note 12) was used to derive its *d* value, which, when multiplied by 15 provides its IQ equivalent.

The present study allows us to test whether Lynn's prediction generalises to *general intelligence g*, using the different approaches exemplified in Table 3. The first is to multiply the observed sex difference in head circumference ($d = .90$ in this study) with the correlation coefficient between head circumference and the *g* factor scores (.21), which gives a predicted male lead in *g* of .19 (or 2.84 equivalent IQ points), corresponding quite well to the observed 3.15*d* equivalent IQ (Table 1). The second approach derives the *d* value from the sex loaded point–biserial correlation r_{pbs} (.274, Table 1) by Jensen's formula (1998, p. 543, note 12), and finds a sex difference in male favour of $d = .57$. Multiplying this *d* by 15 gives an IQ lead of 8.55 equivalent points.

It is better to use the transformed r_{pbs} than factor scores for two reasons. The *g* factor score is necessarily contaminated to some extent by group factors and test specificity. Second, the *g* factor score is based on a *g*-weighted mean of that individual's standardised scores on each of the subtests, which may increase or decrease the mean sex difference as a function of the type of subtests included.

The analysis thus produced three results: (1) Lynn (1999) over-predicted the male IQ lead from brain size by .2 points (4.05–3.85) based on possibly biased data; (2) The present study under-predicted the male *g* lead from head size by −.31 IQ points (2.84–3.15) based on *d* equivalent IQ points, and (3) under-predicted male *g* by −5.40 IQ points (8.55–3.15) based on transforming the sex load of $r_{pbs} = .274$ to $d = .57$ or 8.55 IQ equivalent points. In addition, Colom et al. (2002) observed a sex load on *g* of .159 corresponding to 4.83 equivalent IQ points.

These results provide concurrent support for the third hypothesis, namely, that the larger average male brain explains a significant part of the average male lead in *g*.

5. Hypothesis 4: The *g* lead and wider distribution transform into an exponentially increased male–female ratio at the very high end of the *g* distribution, which partly explains male dominance in society

5.1. Introduction

There is strong selection for *g* all the way up from elementary education to the highest educational, occupational, and political strata (Gottfredson, 2003). Given the male average *g* lead and a wider distribution we can accordingly expect these upper strata to be increasingly conquered by high *g* males. Classical Distribution Theory illustrates this as we plot the separate distributions of male and female *g* factor scores and map the effects of sex differences in mean and dispersion on the male–female ratio at various points on the distributions.

To increase statistical power, the adult 16+ years sub-sample of 62 males and females was combined with an 8–15 years child sub-sample consisting of 59 boys (mean age 11.1, SD = 2.2) and 60 girls (mean age 11.0, SD = 2.3). Sex also loads significantly on children's *g* ($r_{pbs} = .231$; p (one-sided) = .006). Moreover, the male–female *g* factor congruence coefficients amounted to practical identity (.90 and .92 in the young and adult samples, respectively), and the combined young–old-female and young–old-male samples showed a congruence coefficient of .96, also suggesting identical factor structures. The demonstrated similarity in *g* factor structures across sex and age permits pooling young and adult *g* factor scores, to give a total population of 90 males (mean age 13.0, SD = 3.54) and 91 females (mean age 12.8, SD = 3.6).

The male *g* mean was .23 (SD = 1.03) and the female *g* mean −.23 (SD = .93), or a sex difference in *g* factor score of .46 SD (equal to 6.90 IQ points). The combined effects of the mean and dispersion differences on the male–female ratio are illustrated graphically in Fig. 1.

The ratio curve suggests an exponentially increasing male over-representation from average *g* and up, amounting to more than 8 males for each female at *g* = 3 SD (IQ 145).

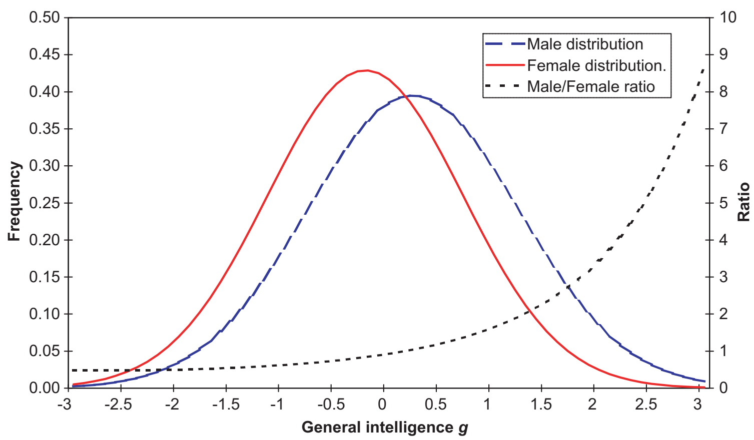

Fig. 1. Total sample male and female general intelligence *g* (HFA/S–L, see text) distributions and male–female ratio as a function of male $g = .23$ (SD = 1.03) and female $g = -.23$ (SD = .93). $N_{males} = 90$ (mean age 13.0, SD = 3.54) and $N_{females} = 91$ (mean age 12.8, SD = 3.6).

5.2. The very high-end male g hypothesis

This hypothesis (Nyborg, 2003, p. 215) posits that the overrepresentation of high *g* males helps explain why males generally stand out in higher educational and socio-economic spheres, and why an official investigation of the power structure "should" find that 88% of the most influential people in Denmark are males (Christiansen, Møller, & Togeby, 2001). In general, *g* is the best single predictor of occupational status and income (e.g. Nyborg & Jensen, 2001) and of how people perform in life at large (e.g. Gottfredson, 2003). The combined evidence thus supports hypothesis 4, saying that the exponentially increasing overrepresentation of high *g* males may by and large explain why males typically out-compete females at the highest steps of the societal ladder.

6. Discussion

Jensen wrote in 1998 (p. 532) that the study of sex differences in general intelligence is "technically the most difficult to answer ... the least investigated, the least written about, and, indeed, even the least often asked".

The present study used a refined analytic approach in the attempt to cut through unclear definitions and unreliable measures, and confirmed hypothesis 1 that previous disagreement was attributable to conceptual and methodological problems.

With intelligence defined as *general intelligence g* (HFA/SL analysis), the present study confirmed hypothesis 2 that sex loads significantly on *g*. We now have three independent data sets showing the sex load on *g*: Children:$r_{pbs} = .231$, p (one-sided) = .006, Adults:$r_{pbs} = .274$, p (one-sided) = .016, and Adults in Colom et al., 2002:$r_{pbs} = .159$, $p = .0001$, respectively. In addition, the less accurate male and female factor scores indicated a sex difference in *g* in the combined Child–Adult sample of .46; $t(179) = 2.66$, p (one-sided) = .009. The difference in male and female dispersion SDs (1.03 and .93, respectively) was also significant (F-ratio variance = 1.50, p (one-sided) = .03). On the other hand, the correlated vector analysis only reached significance ($r_S = .41$, p (one-sided) = .035), suggesting bias or a weak Jensen effect.

The study provided support for hypothesis 3, which states that the sex difference in *g* can be explained, at least in part, in terms of an overall larger average male brain. However, a recent voxel-based morphometric MRI analysis shows that women have more white matter and fewer gray matter areas related to IQ, and the strongest IQ—gray matter correlations are in the female frontal and male frontal and parietal lobes (Haier, Jung, Yeo, Head, & Alkire, in press).

The study made it understandable how an exponentially growing male/female ratio at the high end of the *g* distribution—with the exact numerical ratio being a function of the size of the average sex difference in mean *g* and of the dispersion of *g*—provides part of an explanation of the male dominance in high society.

The present results derive from a more careful sampling of the two independent sub-samples than any before, but the small Ns call for caution in interpretation, even if it is harder to obtain a statistically significant difference in a small than in a large sample.

The general conclusion: Proper methodology identifies a male advantage in *g* that increases exponentially at higher levels, relates to brain size, and explains, at least in part, the universal male dominance in society.

References

Aluja-Fabregat, A., Colom, R., Abad, F., & Juan-Espinosa, M. (2000). Sex differences in general intelligence defined as *g* among young adolescents. *Personality and Individual differences, 28*(4), 813–820.

Anderson, B. (2003). Brain imaging and *g*. In H. Nyborg (Ed.), *The scientific study of general intelligence: Tribute to Arthur R. Jensen*. Oxford: Pergamon, Chapter 2.

Ankney, C. (1992). Sex difference in relative brain size: The mismeasure of woman, too? *Intelligence, 16*, 329–336.

Ankney, C. (1995). Sex difference in brain size and mental abilities: Comments on R. Lynn and Benbow, C.P. *Personality and Individual Differences, 18*, 423–424.

Asch, S., & Witkin, H. (1948). Studies in space orientation. I. Perception of the upright with displaced visual fields. *Journal of Experimental Psychology, 38*, 325–337.

Brody, N. (1992). *Intelligence* (2nd ed.). San Diego, CA: Academic Press.

Christiansen, P. M., Møller, B., & Togeby, L. (2001). *Den danske elite [The Danish elite]*. Copenhagen: Hans Reitzels Forlag.

Colom, R., García, L. F., Juan-Espinosa, M., & Abad, F. (2002). Null sex differences in general intelligence: Evidence from the WAIS-III. *Spanish Journal of Psychology, 5*(1), 29–35.

Colom, R., & García-López, O. (2002). Sex difference in fluid intelligence among high school graduates. *Personality and Individual Differences, 32*, 445–451.

Colom, R., Juan-Espinosa, M., Abad, F., & García, L. (2000). Negligible sex differences in general intelligence. *Intelligence, 28*(1), 57–68.

Gignac, C., Vernon, P. A., & Wickett, J. C. (2003). Factors influencing the relationship between brain size and intelligence. In H. Nyborg (Ed.), *The scientific study of general intelligence: Tribute to Arthur R. Jensen* (pp. 93–106). Oxford: Pergamon.

Goldberg, S. (1977). *The inevitability of patriarchy*. London: Temple Smith.

Gottfredson, L. S. (2003). *g*, jobs and life. In H. Nyborg (Ed.), *The scientific study of general intelligence: Tribute to Arthur R. Jensen* (pp. 293–342). Oxford, UK: Pergamon.

Haier, R. J. (2003). Positron emission tomography studies of intelligence: From psychometrics to neurobiology. In H. Nyborg (Ed.), *The scientific study of general intelligence: Tribute to Arthur R. Jensen*. Oxford: Pergamon [chapter 3].

Haier, R. J., Jung, R. E., Yeo, R. A., Head, K., & Alkire, M. T. (in press). The neuroanatomy of general intelligence: Sex matters.

Halpern, D. F., & LaMay, M. L. (2000). The smarter sex: A critical review of sex differences in intelligence. *Educational Psychology Review, 12*(2), 229–246.

Jensen, A. R. (1980). *Bias in mental testing*. London: Methuen.

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

Lynn, R. (1994). Sex differences in intelligence and brain size: A paradox resolved. *Personality and Individual Differences, 17*, 257–271.

Lynn, R. (1997). Sex differences in intelligence: Data from a Scottish standardization of the WAIS-R. *Personality and Individual Differences, 24*(2), 289–290.

Lynn, R. (1999). Sex difference in intelligence and brain size: A developmental theory. *Intelligence, 27*(1), 1–12.

Lynn, R., Irwing, P., & Cammock, T. (2002). Sex differences in general knowledge. *Intelligence, 30*(1), 27–39.

Money, J. (1965). *A standardized road-map test of direction sense*. Baltimore: The Johns Hopkins Press.

Nyborg, H. (2001). *Early sex differences in general and specific intelligence: Pitting biological against chronological age (Addendum)*. Second annual conference of the international society for intelligence research, Cleveland, OH.

Nyborg, H., (2002). *IQ and g: The art of uncovering the sex difference in general intelligence*. Presented at the third annual conference of the international society for intelligence research, Vanderbilt University, Nashville, TN, 5–17th December. p. 37.

Nyborg, H. (2003). Sex differences in *g*. In H. Nyborg (Ed.), *The scientific study of general intelligence: Tribute to Arthur R. Jensen*. Oxford: Pergamon [chapter 10].

Nyborg, H., & Isaksen, B. (1974). A method for analysing performance in the rod-and-frame test. II. Test of the statistical model. *Scandinavian Journal of Psychology, 15*, 124–126.

Nyborg, H., & Jensen, A. R. (2001). Occupation and income related to psychometric *g*. *Intelligence, 29*, 45–55.

Packenberg, B., & Gundersen, H. J. G. (1997). Neocortical neuron number in humans: Effect of sex and age. *The Journal of comparative Neurobiology, 384*, 312–320.

Rushton, J. P. (1992). Cranial capacity related to sex, rank and race in a stratified sample of 6,325 military personnel. *Intelligence, 16*, 401–413.

Rushton, J. P., & Ankney, C. D. (1996). Brain size and cognitive ability: Correlations with age, sex, social class, and race. *Psychonomic Bulletin and Review, 3*, 21–36.

Schmid, J., & Leiman, J. M. (1957). The development of hierarchical factor solutions. *Psychometrika, 22*, 53–61.

Schoenemann, P. T., Budinger, T. F., Sarich, V. M., & Wang, W. (2000). Brain size does not predict general cognitive ability within families. *Proceedings of the National Academy of Sciences, 97*(9), 4932–4937.

Vandenberg, S., & Kuse, A. (1978). Mental rotations, a group test of three-dimensional visualization. *Perceptual and Motor Skills, 47*, 599–604.

Wechsler, D. (1958). *The measurement and appraisal of adult intelligence* (4th ed.). Baltimore: Williams & Wilkins.

Witkin, H. (1950). Individual differences in ease of perception of embedded figures. *Journal of Personality, 51*, 1–15.