Why are Male IQ Scores More Variable?

Bernard S. Strauss and Jacob A. Strauss

Department of Molecular Genetics and Cell Biology, The University of Chicago and MIT Computer Science and Artificial Intelligence Laboratory

---

Abstract: The male female difference in the variance of quantitative traits including IQ can be accounted for by the observation of the concentration of genes involved in brain development on the X chromosome and by the mechanism of sex determination in mammals. Females have two X chromosomes with one randomly inactivated. Males have a single X chromosome, which results in increased variance. It is shown that these factors are sufficient to account for the observed differences.

The average (mean) scores for males and females for a variety of quantitative traits, including IQ, are almost identical. What is different is that the variance for males is higher [1,2]. This means that there are more males at both ends of the distribution. There are about 40 percent more mentally defective males than females [3] and it has been argued that the excess of males in the very high IQ regions could account for the differential representation of males in tenured positions on the faculties of research universities particularly in the physical sciences [4]. Although there are sociological and economic reasons that might account for this differential, there is a simple biological reason, related to the mechanism of sex determination in mammals, for the greater trait variance of males.

Females have two X chromosomes and in humans the X includes about 1100 genes. The Y chromosome on the other hand includes about 50 genes, most related to sex determination and fertility. The X chromosome also seems to have an unusually high representation of genes involved in brain development [5]. Nature compensates for the excess representation of X-linked genes. In humans this occurs by a process of random inactivation of most of the genes on one or the other of the two X chromosomes. As far as gene function is concerned, females are mosaics [6,7]. The net result of these biological facts is that males must be more variable for any trait involving the X chromosome [6,8].

If IQ were simply related to one gene on the X chromosome with no representative on the Y, and this gene had two alleles (A) and (a) with (A) providing 100 IQ units and (a) providing (0) units, and if all genes acted equally (no dominance) then males would be either X(A)Y or X(a)Y and would have values of 100 or 0. Given equal gene frequencies the mean would be 50 and the variance 50^2. Females would be X(A)X(A) or X(A)X(a) or X(a)X(a) in a ratio 1:2:1. Assuming that X inactivation occurs randomly, the mean would still be 50 but the variance would be half that of males: $50^2/2$. Of course IQ cannot be related to only one gene but is more likely affected by many genes on the X chromosome and on the autosomes. Assuming that each one of these genes can be considered as contributing to IQ in an additive fashion, the contribution of genes on the X chromosome can be generalized as shown in Model 1 (Fig. 1). For any number of loci the male variance is greater than the female variance by the fraction of all loci that are on

the X chromosome . If IQ were determined solely by genetic factors the variance ratio between males and females could take any value between 1 and 2. Of course this does not describe the real world and a model taking into account additional sources of variance is given in the supporting online material.

How well do observed variance ratios fit these models? All are consistent with the predicted range between 1 and 2. The observed variance ratio falls between 1.07 to 1.17 [1] depending on the component tested. In a study on the number of words used by males and females, the mean was statistically indistinguishable but the male variance was 1.40 times greater [9]. A simple prediction based on Model 1 is that between 7% and 17% or 40% of all the genes involved in these traits are located on the X chromosome.

If a major reason for the difference in variance is as straightforward as this analysis indicates then the selection of women at the highest IQ values merely requires increasing the size of the cohort.

References

1. Strand, S., Deary, I.J. & Smith, P. Sex differences in cognitive abilities test scores: a UK national picture. *Br J Educ Psychol* **76**, 463-80 (2006).
2. Hedges, L.V. & Nowell, A. Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science* **269**, 41-5 (1995).
3. Leonard, H. & Wen, X. The epidemiology of mental retardation: challenges and opportunities in the new millennium. *Ment Retard Dev Disabil Res Rev* **8**, 117-34 (2002).
4. Summers, L.H. Remarks at NBER Conference on Diversifying the Science and Engineering Workforce by Lawrence H. Summers, Cambridge Mass. Jan. 14, 2005. in *http://www.president.harvard.edu/speeches/summers_2005/nber.php* (2005).

5. Graves, J.A., Koina, E. & Sankovic, N. How the gene content of human sex chromosomes evolved. *Curr Opin Genet Dev* **16**, 219-24 (2006).
6. Lange, K. & Sobel, E. Variance component models for X-linked QTLs. *Genet Epidemiol* **30**, 380-3 (2006).
7. Payer, B. & Lee, J.T. X chromosome dosage compensation: how mammals keep the balance. *Annu Rev Genet* **42**, 733-72 (2008).
8. Pan, L., Ober, C. & Abney, M. Heritability estimation of sex-specific effects on human quantitative traits. *Genet Epidemiol* **31**, 338-47 (2007).
9. Mehl, M.R., Vazire, S., Ramirez-Esparza, N., Slatcher, R.B. & Pennebaker, J.W. Are women really more talkative than men? *Science* **317**, 82 (2007).

Fig.1. The predicted male/female variance ratio for a quantitative trait as a function of the fraction α of related genes on the X chromosome to total related genes. The model assumes that all genes contribute additively in the same proportion, and are uncorrelated. Since $0 \leq \alpha \leq 1$, the resulting variance ratio falls between 1 and 2, with increased autosomal contribution diluting the observed variance ratio toward 1.

Model 1 Let M and F be random variables describing a quantifiable trait in males and females. Assume that M and F each have mean value $\bar{X}$. Let M and F each be the sum of random variables A_i noting autosomal loci and M_i or F_i noting x-linked loci. Let α be the fraction of the n sites that are x-linked:

$$M = \sum_{i=1}^{(1-\alpha)n} A_i + \sum_{i=1}^{\alpha n} M_i \qquad F = \sum_{i=1}^{(1-\alpha)n} A_i + \sum_{i=1}^{\alpha n} F_i$$

Assuming that the mean contribution of each locus is the same: $\mu \equiv E[A_i] = E[M_i] = E[F_i] = \bar{X}/n$, the variance of each locus is:

$$\begin{aligned} Var(M_i) &= \frac{1}{2}(0-\mu)^2 + \frac{1}{2}(2\mu-\mu)^2 &&= \mu^2 \\ Var(F_i) &= \frac{1}{4}(0-\mu)^2 + \frac{1}{2}(\mu-\mu)^2 + \frac{1}{4}(2\mu-\mu)^2 &&= \frac{\mu^2}{2} \\ Var(A_i) &= \frac{1}{4}(0-\mu)^2 + \frac{1}{2}(\mu-\mu)^2 + \frac{1}{4}(2\mu-\mu)^2 &&= \frac{\mu^2}{2} \end{aligned}$$

The Variance Ratio, assuming that all loci are independent, is therefore:

$$\begin{aligned} \frac{\sigma_M^2}{\sigma_F^2} &= \frac{\sum_{i=1}^{(1-\alpha)n} Var(A_i) + \sum_{i=1}^{\alpha n} Var(M_i)}{\sum_{i=1}^{(1-\alpha)n} Var(A_i) + \sum_{i=1}^{\alpha n} Var(F_i)} \\ &= \frac{(1-\alpha)n\frac{\mu^2}{2} + \alpha n\mu^2}{(1-\alpha)n\frac{\mu^2}{2} + \alpha n\frac{\mu^2}{2}} \\ &= \frac{1+\alpha}{1} \end{aligned}$$

SUPPLEMENTAL MATERIAL

Fig S1. An extension to Model 1 that groups together additional sources of observed variance such as measurement errors or non-additive sources. The resulting model is no longer dependent only on the fraction of X-linked genes, but instead relates several parameters through a construct β. Since $\beta \geq 0$, the resulting range of variance ratios is again between 1 and 2.

Model 2 Define $\bar{X}$, M_i, and F_i as in model 1. Let n be the number of x-linked loci. Let A be a random variable noting all non-x-linked sources of trait variance such as autosomal and measurement contributions. Let $\bar{A}$ and σ_A^2 be A's mean and variance:

$$M = A + \sum_{i=1}^{n} M_i \qquad F = A + \sum_{i=1}^{n} F_i$$

In this case the mean values of M_i and F_i are:

$$\mu = \frac{\bar{X} - \bar{A}}{n}$$

The Variance Ratio, again assuming that all loci are independent, is:

$$\frac{\sigma_M^2}{\sigma_F^2} = \frac{\sigma_A^2 + n\mu^2}{\sigma_A^2 + n\mu^2/2}$$

Which, after introducing a change of variable $\beta \equiv \frac{\sigma_A^2}{2n\mu^2}$ reduces to:

$$\frac{\sigma_M^2}{\sigma_F^2} = \frac{\beta + 2}{\beta + 1}$$