# New questions about old heritability estimates

PETER H. SCHÖNEMANN
*Purdue University, West Lafayette, Indiana*

This note summarizes the main results of three recent studies on the heritability of mental traits: (1) The inferences Jinks and Fulker (1970) derived from Shields's (1962) twin data are invalid since the assumptions of the genetic model are consistently violated by these data. A purely environmental model fits them better by a factor of 2. (2) Holzinger's heritability coefficient ($h^2$) is invalid since his derivations imply that dizygotic twins share no genes. (3) In contrast, Nichols's (1965) heritability coefficient (HR) follows from a strictly additive genetic model. (4) However, the needed assumptions are consistently violated by Osborne's (1980) personality data, which produce an excessive number of inadmissible HRs. A purely environmental model fits these data better by a factor of 14. Jointly, these results suggest that heritability estimates of mental traits in the literature should be viewed with caution.

The purpose of this advance note is to put some recent results about the heritability of mental traits on record until more complete accounts of this work can be published in the appropriate journals. In the meantime, the reader is referred to Schönemann (1987, 1988) and Schönemann and Schönemann (1988) for more detailed empirical evidence and the proofs of formal claims.

## IDENTICAL TWINS RAISED TOGETHER AND APART (SHIELDS'S DATA)

In 1962, Shields published a set of twin data, which has been reanalyzed by numerous other investigators, including Jinks and Fulker (1970) and, more recently, Farber (1981). Shields recorded the values of various physical and four psychological variables: the Dominoes Intelligence Test, a synonyms section of the Mill Hill Vocabulary Scale, and two self-rating scores for extraversion and neuroticism. The sample comprised 44 pairs of monozygotic twins raised apart (MZAs), 44 pairs of monozygotic twins raised together (MZTs), and 32 pairs of dizygotic twins raised together (DZTs). On reanalyzing these data, Jinks and Fulker (1970) concluded, among other things; (1) that "it is reassuring to find that [correlated environments and genotype-environment interactions] are by no means universal phenomena" (p. 347); (2) "high heritabilities recorded in Table 30" (p. 347) for the four measures employed by Shields; and (3) that "the high number of genes estimated to be controlling IQ ($>$ 20 and approximately 100) fully confirms that this trait is under polygenic control" (p. 348). According to Eysenck (1973, p. 262), "The Jinks and Fulker paper ... is the cornerstone on which any future argument about heritability must be based."

Address correspondence to Peter H. Schönemann, Department of Psychological Sciences, Peirce Hall 365, Purdue University, West Lafayette, IN 47907.

To account for Shields's (1962) data, Jinks and Fulker (1970) postulated a variance components model that provides for four mutually uncorrelated latent variables to explain the observed variable $y$:

$$y - \mu = g_b + g_w + e_b + e_w,$$

where $\mu$ is the population mean of $y$, $g_b$ a between-family genetic variable, $g_w$ a within-family genetic variable, $e_b$ a between-family environmental variable, and $e_w$ a within-family environmental variable. Thus, the model postulates for MZTs

$$y_1 - \mu = g_b + g_w + e_b + e_{w_1} \quad \text{(1st twin)}$$

$$y_2 - \mu = g_b + g_w + e_b + e_{w_2} \quad \text{(2nd twin)},$$

and for MZAs

$$y_1 - \mu = g_b + g_w + e_{b_1} + e_{w_1},$$

$$y_2 - \mu = g_b + g_w + e_{b_2} + e_{w_2},$$

where the six variables on the right are again uncorrelated, and now, in addition, $\text{var}(e_{b_1}) = \text{var}(e_{b_2}) = \text{var}(e_b)$.

As shown in more detail in Schönemann (1987), this model implies

(a) $E(\text{MSBA}+\text{MSWA}) = E(\text{MSBT}+\text{MSWT})$,

where $E(\text{MSBT})$ and $E(\text{MSWT})$ denote the expected values of the between and within sums of squares for the MZTs, and $E(\text{MSBA})$ and $E(\text{MSWA})$ denote those for the MZAs, and

(b) $E(\text{MSBA}-\text{MSWA}) \leq E(\text{MSBT}-\text{MSWT})$,

(c) $r_A - r_T \leq 0$,

where $r_A$ denotes the intraclass correlation (based on mean squares) for MZAs, and $r_T$ denotes the intraclass correlation for MZTs. These predictions are violated by Shields's (1962) data (a) in 7 out of 8 cases (88%), (b) in 8 out of 8 cases (100%), and (c) in 7 out of 8 cases (88%).

To take account of these systematic violations, an alternative *correlated environments model* was constructed on the following premises: (1) Jinks and Fulker's (1970) assumption "that separated twins are randomly assigned across environments" (p. 325) is absurd so that an additional covariance parameter, $\varrho_{e_1,e_2}$, is needed to account for correlated environments between MZAs. (2) An additional variable $s$ (for *separation variable*) is needed for the MZAs to predict a larger $E(MSB + MSW)$ for the MZAs and is assumed to vary only between families. A possible interpretation of $s$ might be length of separation (see, e.g., Farber, 1981). (3) The observed variables contain no genetic component at all. On fitting to the Shields data the resulting four-parameter environmental model

$$\text{MZTs:} \quad y_1 - \mu = e + z_1$$
$$y_2 - \mu = e + z_2,$$

$$\text{MZAs:} \quad y_1 - \mu = e_1 + z_1 + s$$
$$y_2 - \mu = e_2 + z_2 + s,$$

with $\text{cov}(e_1, e_2) = \varrho_{e_1 e_2} \text{var}(e) \neq 0$,

where $z_1$ and $z_2$ denote measurement errors, one obtains positive least-squares estimates for all variances and high positive environmental correlations for the MZAs:

Neuroticism, $\varrho_{e_1 e_2} = .91$,  Extraversion, $\varrho_{e_1 e_2} = .94$,

Mill Hill, $\varrho_{e_1 e_2} = .72$,  Dominoes, $\varrho_{e_1 e_2} = 1.04$.

Hence, the fit will deteriorate only slightly if one fixes $\varrho_{e_1 e_2} = 1$. Thus one arrives at a three-parameter correlated environments model (CE3) that contains the same number of parameters as the genetic model proposed by Jinks and Fulker (1970), so that the fit of both models can be compared. The results are as follows: (1) Qualitatively: Under the genetic model, the least-squares estimates of the between-family variance component, $\text{var}(e_b)$, are negative in 12 out of 12 cases. Under the purely environmental model, all variance estimates are positive. (2) Quantitatively: On comparing the fit of both models in terms of the unweighted residuals ($e'e/y'y = 1 - \eta^2$, where now $e$ is the least-squares error and $y$ the criterion) and also in terms of the predicted and observed intraclass correlations for the MZTs and MZAs, the correlated environments model fits roughly twice as well as the genetic model proposed by Jinks and Fulker (1970). In terms of the predicted intraclass correlations, the fit ratio favoring the environmental model is 2.17, and, in terms of the error measures $1 - \eta^2$ left by both models after a least-squares fit, it is 2.10 for the four larger, pooled groups. For the eight smaller within-sex analyses, it is 4.54.

Since a purely environmental model fits Shields's (1962) data better than the genetic model of Jinks and Fulker (1970), their claim that performance on the four tests is largely inherited is unsupported.

## HERITABILITY ESTIMATES

This section summarizes some formal arguments developed more fully by Schönemann (1988).

## Holzinger's Derivation of $h^2$

To derive his heritability ratio $h^2$, Holzinger (Newman, Freeman, & Holzinger, 1937, pp. 94–116) started out with a variance component model for the MZs that postulates four latent variables, $g$ (genetic component), $e_1$ and $e_2$ (environmental components), and $z_1$ and $z_2$ (measurement errors):

$$\text{MZs:} \quad y_1 - \mu = g + e_1 + z_1, \quad y_2 - \mu = g + e_2 + z_2$$

where $\mu = E(y_k)$, $E(g) = E(e_k) = E(e_k) = E(z_k) = 0$, and $\text{var}(g, e_1, e_2, z_1, z_2) = \text{diagonal}$ (i.e., all latent variables are uncorrelated) and $\text{var}(z_k) = \text{var}(z)$, $\text{var}(e_k) = \text{var}(e)$. However, halfway through his derivations, Holzinger discarded the measurement error variables by setting var $(z) = 0$ (p. 113). Therefore, the model Holzinger actually worked with predicts for the MZs:

$$E(MSBM) = 2\text{var}(g) + \text{var}(e),$$

$$E(MSWM) = \text{var}(e),$$

$$\varrho_M = \text{var}(g)/[\text{var}(g) + \text{var}(e)],$$

so that the genetic variance ratio is simply given by the intraclass correlation of the MZs.

$$\text{DZs:} \quad y_1 - \mu = g_1 + e_1, \quad y_2 - \mu = g_2 + e_2.$$

Holzinger did not spell out the critical correlation between the two genetic variables $g_1$ and $g_2$. Suppose it is given by $\varrho_{12}$. Then,

$$E(MSBD) = (1 + \varrho_{12})\text{var}(g) + \text{var}(e),$$

$$E(MSWD) = (1 - \varrho_{12})\text{var}(g) + \text{var}(e)$$

$$\varrho_D = \varrho_{12}\text{var}(g)/[\text{var}(g) + \text{var}(e)].$$

This implies that his heritability coefficient can be written

$$h^2 = (\varrho_M - \varrho_D/(1 - \varrho_D)$$

$$= (1 - \varrho_{12})\text{var}(g)/[(1 - \varrho_{12})\text{var}(g) + \text{var}(e)].$$

On setting $h^2 = \text{var}(g)/[\text{var}(g) + \text{var}(e)]$, one obtains $\varrho_{12}\text{var}(g)\text{var}(e) = 0$, so that $\varrho_{12} = 0$ if both variances are nonzero. As a result, $\varrho_D = 0$. The necessary condition $\varrho_{12} = 0$ violates the basic assumption of genetic theory that dizygotic twins, on average, share half of the genes.

## Nichols's (1965) Heritability Index

Nichols's (1965) heritability index

$$[HR := 2(r_M - r_d)/r_M],$$

in contrast, can be derived as a valid deduction from a strictly additive genetic variance component model AG3, which is a special case of Fisher's (1918) model when gene/environment correlation, interactions, and dominance are all zero. This model postulates three classes of latent variables strictly additive genetic variables $a$, environmental variables $e$, and measurement error $z$. For the

$$\text{MZs:} \quad y_1 - \mu = a + e + z_1, \quad y_2 - \mu = a + e + z_2,$$

with $E(a) = E(e) = E(z_k) = 0$ and $\text{Var}(a, e, z_1, z_2) =$ diagonal.

In contrast to Holzinger's model, the environmental variable $e$ is now constant within families (since both twins are raised together). This model implies the expected mean squares:

MZs:

$E(\text{MSBM}) = 2\text{var}(a) + 2\text{var}(e) + \text{var}(z)$,

$E(\text{MSWM}) = \text{var}(z)$,

$$\varrho_M = [\text{var}(a) + \text{var}(e)]/[\text{var}(a) + \text{var}(e) + \text{var}(z)],$$

DZs:

$$y_1 - \mu = a_1 + e + z_1, \quad y_2 - \mu = a_2 + e + z_2,$$

where the latent variables again have zero means, and $\text{Var}(a_1, a_2, e_1, e_2, z_1, z_2) = $ diagonal *except* for $\text{cov}(a_1, a_2)$. Since dizygotic twins, on average, share half the genes, the correlation $\varrho_{12}$ between $a_1$ and $a_2$ is 1/2 by the correlation formula for common elements (Hogben, 1950, pp. 360f.). Therefore,

$E(\text{MSBD}) = 1.5\text{var}(a) + 2\text{var}(e) + \text{var}(z)$,

$E(\text{MSWD}) = .5\text{var}(a) + \text{var}(z)$,

$$\varrho_D = [.5\text{var}(a) + \text{var}(e)]/[\text{var}(a) + \text{var}(e) + \text{var}(z)],$$

so that $\varrho_M - \varrho_D = .5\text{var}(a)/[\text{var}(a) + \text{var}(e) + \text{var}(z)]$. Hence,

$$\text{HR} := 2(\varrho_M - \varrho_D)/\varrho_M = \text{var}(a)/[\text{var}(a) + \text{var}(e)]$$

is indeed the narrow heritability ratio if AG3 fits. In contrast,

$$h^2 := (\varrho_M - \varrho_D)/(1 - \varrho_D)$$
$$= .5\text{var}(a)/[.5\text{var}(a) + \text{var}(z)]$$

measures nothing of interest, because it does not contain $\text{var}(e)$. For what follows, it should be noted that AG3 implies a number of conditions that should be checked before relying on HR:

(a) $E(\text{MSWD} - \text{MSWM}) = \text{var}(a)/2 \geq 0$,

(b) $E(\text{MSBM} + \text{MSWM}) = E(\text{MSBD} + \text{MSWD})$

(c) $E(\text{MSBD} - \text{MSWD}) - E(\text{MSBM} - \text{MSWM})/2$

$= \text{var}(e) \geq 0$,

(d) $\varrho_D \geq \varrho_M/2$.

### Jensen's (1967) Index

Jensen's (1967) index (JHR) is as follows:

$$\text{JHR} := 2(r_M - r_D) = r_M \, \text{HR}$$
$$= \text{var}(a)/[\text{var}(a) + \text{var}(e) + \text{var}(z)].$$

JHR is not an improvement over Nichols's HR, as some believe, because the denominator is inflated by the unknown measurement error variance, $\text{var}(z)$. If JHR is divided by the reliability ($= r_M$ under the strictly additive model AG3), then one arrives again at HR.

### IDENTICAL AND FRATERNAL TWINS (OSBORNE PERSONALITY DATA)

Osborne (1980) reported the $h^2$ and HR for two personality test batteries, Jenkins's (1959) "How well do you know yourself" questionnaire (19 variables, 2 subsamples: male, female), and Cattell's (1958) "High school personality questionnaire" (13 variables, 3 subsamples: white, black, total). One of Cattell's variables ($B$) is supposed to be a power test of intelligence.

Across all studies and variables, 59% of the HRs exceeded 1 and 14% were negative. Thus, only 28% of the HRs fell into the admissible range [0,1]. The main qualitative violation of the genetic model underlying HR was $r_D < r_M/2$; that is, monozygotic twins are more similar than the genetic model predicts. As for the Shields data, a least-squares fit of the genetic model to these data produced many (59%) negative variance estimates.

As an alternative, Schönemann and Schönemann (1988) fitted a purely environmental model E3 for the MZ/DZ data, which has the same number of free parameters as the genetic model from which HR was derived: To obtain a model that predicts larger intraclass correlations for the MZs than for the DZs, we postulated two environmental variables, a baseline variable $e_D$, which enters the scores of both MZs and DZs, and a second component $e_M$, which contributes only to the scores of the MZs, thus raising their environmental variance. In addition, the model E3 postulates a measurement error variable $z$. Thus, we postulate:

MZs:

$$y_1 - \mu = e_M + e_D + z_1, \quad y_2 - \mu = e_M + e_D + z_2$$
$$E(e_M, e_D, z_1, z_2) = \emptyset',$$
$$\text{Var}(e_M, e_D, z_1, z_2) = \text{diag}, \quad \text{var}(z_1) = \text{var}(z_2).$$

DZs:

$$y_1 - \mu = e_D + z_1, \quad y_2 - \mu = e_D + z_2,$$

with the same stochastic restrictions on the remaining variables as for the MZs. This model implies

$E(\text{MSBM} + \text{MSWM}) \geq E(\text{MSBD} + \text{MSWD})$,

$E(\text{MSBM} - \text{MSBD}) \geq 0$, $E(\text{MSWM}) = E(\text{MSWD})$,

$$1 \geq \varrho_M = [\text{var}(e_M) + \text{var}(e_D)]/$$
$$[\text{var}(e_M) + \text{var}(e_D) + \text{var}(z)]$$
$$\geq \text{var}(e_D)/[\text{var}(e_D) + \text{var}(z)]$$
$$= \varrho_D \geq 0.$$

On comparing the fit of the genetic model AG3 and the competing environmental model E3, the following results were obtained: (1) The overall percentage of inadmissible variance estimates dropped from 25% to 14% and their average magnitude from 3.68 for the genetic model to 1.17 for the environmental model. (2) Overall, in terms of sums of squared residuals, the environmental model fits the Osborne (1980) personality data better by a factor of 13.9. (See Schönemann & Schönemann, 1988, for more details.)

On the basis of these results we concluded: (1) Osborne's data contain no genetic component at all. (2) The strongest variance component of the personality variables is measurement error. (3) The greater similarity of MZs compared with DZs is probably caused by additional environmental variance induced by the greater physical resemblance of the MZs (e.g., Jones, 1949).

## DISCUSSION

Taken together, the results of the three interrelated heritability studies summarized here suggest that past pronouncements about the presumed heritability of mental traits, including intelligence (e.g., Jensen, 1970), should be viewed with extreme caution because (1) glaring statistical and conceptual errors that invalidate previous empirical inferences went undetected for decades, and (2) whenever a fair comparison was made between a genetic model and an environmental model with the same number of parameters, we found the environmental model produced fewer inadmissible parameter estimates and a better fit.

## REFERENCES

CATTELL, R. B. (1958). *Handbook for the IPAT High School Personality Questionnaire*. Champaign, IL: IPAT.

EYSENCK, H. J. (1973). *The measurement of intelligence*. Baltimore: Williams and Wilkins.

FARBER, S. L. (1981). *Identical twins reared apart: A reanalysis*. New York: Basic Books.

FISHER, R. A. (1918). The correlation between relatives on the supposition of Mendelian inheritance. *Transactions of the Royal Society of Edinburgh*, 52, 399-433.

HOGBEN, L. (1950). *Chance and choice by cardpack and chessboard* (Vol. 1). London: Parrish.

JENKINS, T. N. (1959). *How well do you know yourself?* New York: Executive Analysis Corportion.

JENSEN, A. R. (1967). Estimation of the limits of heritability of traits by comparison of monozygotic and dizygotic twins. *Proceedings of the National Academy of Sciences*, 58, 149-159.

JENSEN, A. R. (1970). Statement of Dr. Arthur R. Jensen before the General Subcommittee on Education, House Education and Labor Committee. *Congressional Record*, I-237.

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.

JONES, H. E. (1949). Environmental influences on mental development. In L. Charmichael (Ed.), *Manual of child psychology* (4th ed.). New York: Wiley.

NEWMAN, H., FREEMAN, F., & HOLZINGER, K. (1937). *Twins. A study of heredity and environment*. Chicago: University of Chicago Press.

NICHOLS, R. C. (1965). The National Merit twin study. In G. S. Vandenberg (Ed.). *Methods and goals in human behavior genetics*. New York: Academic Press.

OSBORNE, T. (1980). *Twins: Black and white*. Athens, GA: Foundation for human understanding.

SCHÖNEMANN, P. H. (1987). *Environmental versus genetic variance component models for the identical twin data by Shields* (Tech. Rep. No. 87-9). West Lafayette, IN: Purdue Mathematical Psychology Program.

SCHÖNEMANN, P. H. (1988). *A note on Holzinger's heritability estimate $h^2$* (Tech. Rep. No. 88-5). West Lafayette, IN: Purdue Mathematical Psychology Program.

SCHÖNEMANN, P. H., & SCHÖNEMANN, R. D. (1988). *Environmental versus genetic models for Osborne's personality data on identical and fraternal twins* (Tech. Rep. No. 88-6). West Lafayette, IN: Purdue Mathematical Psychology Program.

SHIELDS, J. (1962). *Monozygotic twins brought up apart and brought up together*. London: Oxford University Press.