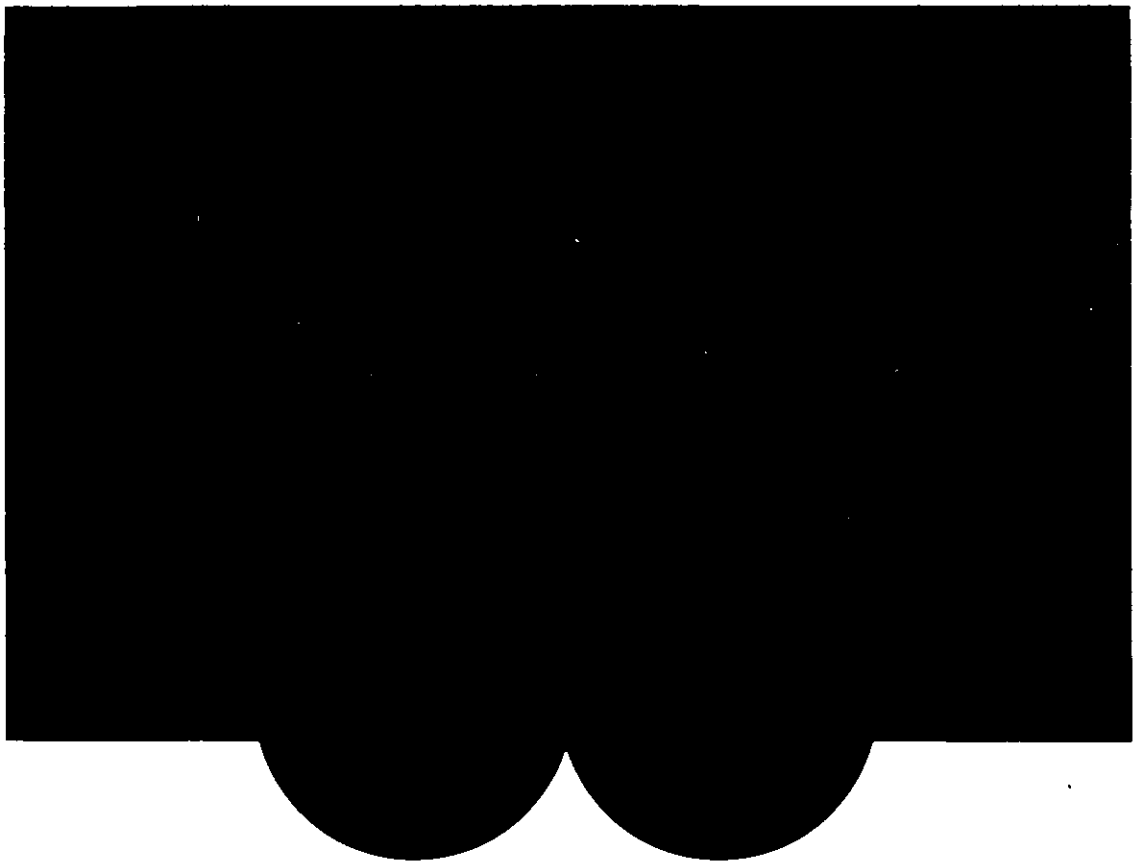


REVISITING DATA ON TWINS TO SEPARATE NATURE FROM NURTURE

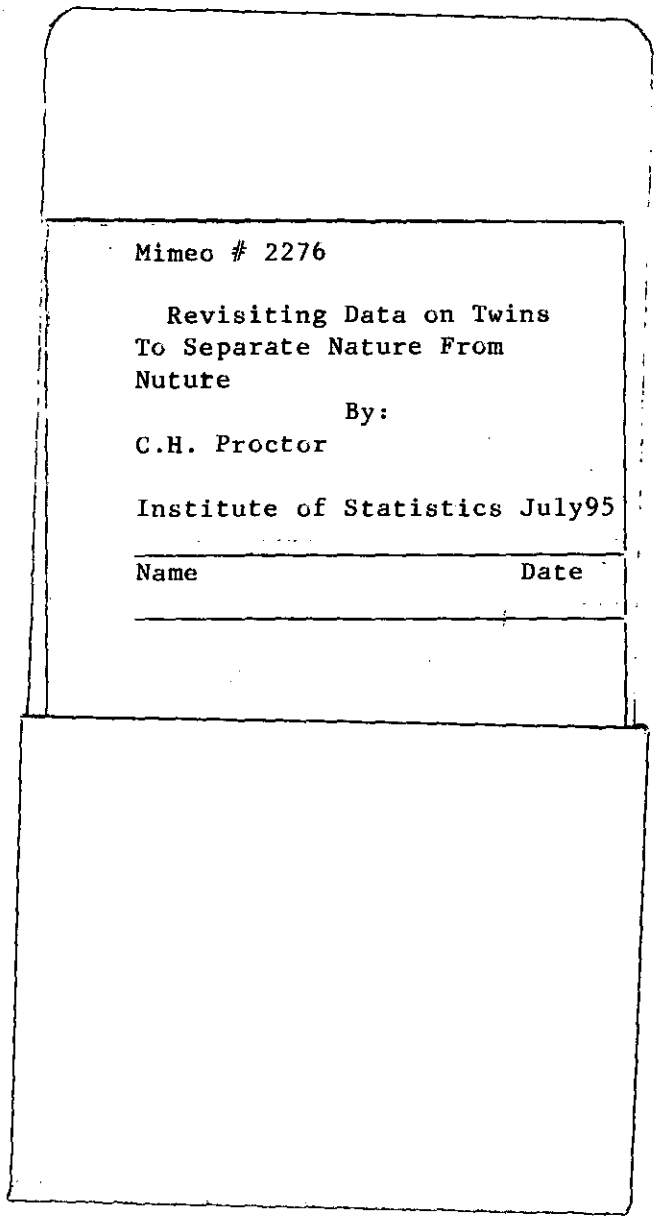
by

C. H. Proctor

Institute of Statistics Mimeograph Series No. 2276

July 1995

NORTH CAROLINA STATE UNIVERSITY
Raleigh, North Carolina



Mimeo # 2276

Revisiting Data on Twins To Separate Nature From Nuture

By:

C.H. Proctor

Institute of Statistics July95

Name Date

Revisiting Data on Twins to Separate Nature from Nurture

C. H. Proctor, NCSU, July 1995

Part I. Model Formulation

Fashions in statistical analysis methods change from time to time but carefully collected data endure. When in 1961 we examined data from mental and physical measurements on 80 like-sexed twins (Clark, *et al*, 1961), the method of choice was nonparametric. However, if one checks for normality and linearity of relationships with today's diagnostic methods (Proctor, 1987) there seems to be a hint of possible outliers but no skewness, no variance heterogeneity, no nonlinearity. Thus the methods of choice will, for today, be parametric.

The objectives of the present analysis will include those of the earlier study -- relating mental scores to physical scores and looking for birth order effects. A major additional objective of the present anlaysis is estimating and comparing the heritabilities of mental abilities tests and physical scores. The birth order question concerns means while the other two issues involve covariances and variances. We will deal first with the birth order differences and then we'll tackle the relationships.

Thirteen variables were listed for every subject. They were: M, R, V, T, L, B, Z, C, S, W, I, E, and P in the article's notation. T is the sum of M, R and V the mental test scores. Z, C, I, E, and P are also functions of the more basic L = head length, B = head breadth, S = stature and W = weight. While the derived variables have their uses, it was decided to confine this preliminary anlaysis to the three mental scores (M = Raven's Progressive Matrices, R =

Chicago Mental Abilities for Reasoning and V = Chicago Mental Abilities for Verbal) and the head sizes (L and B) and body sizes (S and W). Factor analysis results showed that two factors seem to underly these seven. The rotated factor loadings are given in Table 1.

Table 1. Rotated Factor Loadings of Seven Twin Variables (n = 153 full-data cases)

Mean[a]	St. Dev.		Factor I	Factor II	Name of Variable
38.0	9.7	M	.17	.75	Raven's Matrices
53.8	17.5	R	.03	.89	Chicago Primary Reasoning
75.6	25.5	V	.12	.89	Chicago Primary Verbal
186.4	7.8	L	.72	.18	Head Length
146.8	6.7	B	.66	.02	Head Breadth
1652.0	85.1	S	.84	.11	Stature
122.2	22.7	W	.85	.09	Weight

[a] Weight is in pounds and the others are in millimeters.

The correlations among the seven variables must, of course, be kept in view but scanning for birth order effects can most conveniently be done separately for each variable. There are differences by sex to be expected and the possibility of zygosity differences exists. Thus we coded the factors: Zygosity (monozygous versus dizygous), Sex (male versus female), Order (first versus second), and Pair (numbered from 1 to the number of pairs in each zygosity-sex group).

For each variable a linear model with the 15 lines of a full factorial on four factors was run. The mean squares were scanned by the LSVC program, which drops higher order interactions and provides the correct error term in this mixed model situation (Proctor, 1985). That is, Pairs are random and the other 3 factors are fixed.

The P-values gave evidence of differences by sex at the .0051 level for Raven's Matrices with females averaging 36 and males 40. The reasoning scores were a bit higher for first borns at 55 than second borns at 53 with a P-value of .050. The verbal scores for first borns were higher at 77.9 than the second borns 73.9, with a main effect P-value of .01, but there was a zygosity-by-sex interaction in that the order difference was most pronounced among dizygous twins.

The four physical measures all differed by sex - B and S were most pronounced and they also show birth order effects - males larger and first borns also a bit larger. For both L and W there was a zygosity-by-sex by order interaction and this is best viewed in Table 2.

Table 2. Means for Head Length and for Weight to Illustrate Pattern of Three-way Interaction

		Head Length (mm)		Weight (lbs.)		
		First	Second	First	Second	(n)
Dizygous	Females	183	182	118	114	(23)
	Males	190	192	128	136	(14)
Monozygous	Females	183	185	112	113	(20)
	Males	190	188	132	131	(23)
Two-tail, 5% LSD		= 2.0 mm		= 5.2 lbs.		

Whatever differences in variance between the dizygous and the monozygous twins there may be were glossed over in the scan for effects of birth order. They would not be large enough to disturb such findings. Now our attention

turns to the relative variance of the inherited component versus the other components and it will be these very differences in variance which become crucial.

The following model was formulated to match the nature of the present data and the structure of the seven variables. Among the variables are three potentially parallel pairs (R-and-V, L-and-B and S-and-W) in which the notion of "parallel" is of indicators for an underlying trait. R and V indicate "general symbolic intelligence," L and B indicate "general head size" and S and W indicate "general body size." When R and V are averaged to define a new variable, TT say, then M and TT become potentially parallel indicators of "general intelligence."

The two parallel variables will be denoted generally as y_i and $y_{i'}$. Zygosity will be denoted by the subscript j = D or j = M, family is denoted $\ell = 1, 2, \ldots, n(j)$, while birth order is indexed by $\mathcal{O} = 1$ or 2. The model must have fixed effects for sex and birth order in accord with our just concluded examination of means, but there is no need to carry along the terms. The model equations for the random portion are written:

$$y_{iD\ell\mathcal{O}} = \pi_{i\ell} + \Delta_{i\ell\mathcal{O}} + e_{i\ell} + \epsilon_{i\ell\mathcal{O}},$$
$$y_{i'D\ell\mathcal{O}} = \pi_{i'\ell} + \Delta_{i'\ell\mathcal{O}} + e_{i'\ell} + \epsilon_{i'\ell\mathcal{O}},$$
$$y_{iM\ell\mathcal{O}} = \lambda_{i\ell} + e_{i\ell} + \epsilon_{i\ell\mathcal{O}}, \text{ and}$$
$$y_{i'M\ell\mathcal{O}} = \lambda_{i'\ell} + e_{i'\ell} + \epsilon_{i'\ell\mathcal{O}}.$$

The e's are "family" environmental influences. The ϵ's are a combination of measurement error and individual effects, and since our data do not allow these to be separated we will use just the one term for both. For monozygous

twins there is just the one genetic effect, λ, but for the dizygous twins there is a common genetic term π and a specific genetic term Δ.

All terms have zero means since we are supposing the fixed effects have been subtracted. The ϵ variances are σ^2_{MIi} and $\sigma^2_{MIi'}$. The family, or twin-pair, environment variances are σ^2_{Fi} and $\sigma^2_{Fi'}$. The genetic variances are: $V(\lambda_i) = \sigma^2_{Gi}$, $V(\pi_{i\ell}) = \sigma^2_{Gi}/2$ and $V(\Delta_{i\ell}) = \sigma^2_{Gi}/2$ with the same situation holding for i'. That is, the identical genetic part is just one-half in the case of dizygous twins.

An analysis of variance would separate the between family from the within family sums of squares and, taking two variables at a time, yields six "mean squares" and "mean cross products." We can derive the six expressions for expectations by finding variances and covariances for the sums and differences of the twins in each family in terms of the theoretical variances and of the theoretical correlations.

A simple case is the within pair mean square for monozygous twins. The difference is $y_{iM\ell 1} - y_{iM\ell 2} = \epsilon_{i\ell 1} - \epsilon_{i\ell 2}$. Its variance is $2\sigma^2_{MIi}$. The within family covariance is $\mathrm{Cov}(\epsilon_{i\ell 1} - \epsilon_{i\ell 2}, \epsilon_{i\ell 1} - \epsilon_{i\ell 2}) \equiv 2\rho_{MI_{ii'}}\sigma_{MIi}\sigma_{MIi'}$ wherein the measurement-individual correlation is defined. Since ANOVA computations will assure that the coefficient of σ^2_{MIi} will be unity we divide by 2 to get the expectations of the mean squares and mean cross products as σ^2_{MIi}, $\rho_{MI_{ii'}}$, $\sigma_{MIi}\,\sigma_{MIi'}$ and $\sigma^2_{MIi'}$.

For dizygous twins the difference becomes $y_{iD\ell 1} - y_{iD\ell 2} = \Delta_{i\ell 1} - \Delta_{i\ell 2} + \epsilon_{i\ell 1} - \epsilon_{i\ell 2}$. Now the variance of this difference will be $\sigma^2_{Gi} + 2\sigma^2_{\epsilon i}$ and so the expectations of the mean squares are of the form $\sigma^2_{\epsilon i} + .5\sigma^2_{Gi}$. The expectation of the "mean cross product" is $\rho_{MIii'}\,\sigma_{MIi}\sigma_{MIi'} + .5\,\rho_{Gii'}\,\sigma_{Gi}\sigma_{Gi'}$.

The between family mean squares have expectations of $\sigma^2_{MIi} + 2\sigma^2_{Gi} + 2\sigma^2_{Fi}$ for monozygous twins and of $\sigma^2_{MIi} + 1.5\sigma^2_{Gi} + 2\sigma^2_{Fi}$ for the dizygous twins. The expectations of the mean cross products are correspondingly:

$$\rho_{MIii'}\sigma_{MIi}\sigma_{MIi'} + 2\rho_{Gii'}\sigma_{Gi}\sigma_{Gi'} + 2\rho_{Fii'}\sigma_{Fi}\sigma_{Fi'}, \text{ and}$$

$$\rho_{MIii'}\sigma_{MIi}\sigma_{MIi'} + 1.5\rho_{Gii'}\sigma_{Gi}\sigma_{Gi'} + 2\rho_{Fii'}\sigma_{Fi}\sigma_{Fi'} .$$

These symbols can be reduced somewhat by defining $\sigma^2_{Ti} = \sigma^2_{Gi} + \sigma^2_{Fi}$ with $\sigma^2_{Gi}/\sigma^2_{Ti} = H_i$ say and then $\sigma^2_{MIi} \equiv \alpha_i\sigma^2_{Ti}$ defines α_i. If the two variables y_i and $y_{i'}$ are parallel indicators of some underlying trait then we would expect $H_i = H_{i'} = H$ say, as well as $\alpha_i = \alpha_{i'} = \alpha$ say. If we define γ by $\sigma^2_{Ti} \equiv \gamma\,\sigma^2_{Ti}$ the cast of seven parameters is complete.

That is, the 12 mean squares and mean cross products can be modelled by the 7 parameters: σ^2_{Ti}, γ, H, α, ρ_G, ρ_F and ρ_{MI}. We used PROC NLIN in SAS, with weights inverse to the square of the mean squares (or cross products) divided by degrees of freedom to furnish a fit. The results appear in Table 3. The program code is in the Appendix.

When we computed the intraclass correlation coefficients (essentially the correlation between sibs) the average seemed to be above .95 and thus we were surprised, at first, by the seemingly high values for $\hat{\alpha}$ in Table 3. The excess would seem to be due to individual developmental effects and this component should also, it would seem, be included in the definition of heritability. Having defined $H = \sigma^2_G/(\sigma^2_G + \sigma^2_F)$ as a model parameter we now define $H' = \sigma^2_G/(\sigma^2_G + \sigma^2_F + \sigma^2_I)$ as a heritability parameter where σ^2_I is $(\alpha - .05)\ \sigma^2_T$ or that part of σ^2_{MI} due to the individual.

Table 3. Parameter Estimates in Fitting to Mean Squares for Four Pairs of Variables. Heritability is estimated as $H' = H/(1+\alpha-.05)$.

	Variables Playing Roles of *ith* and *i'th*			
Parameter	Reasoning & Verbal	Head Length & Breadth	Stature & Weight	Raven's Matrices & Ave. of R and V
α	.15	.28	.08	.22
H	.27	.60	.56	.38
σ^2_{Ti}	270	38.1	4007	71.3
γ	1.83	.57	.10	3.06
$\rho_{MI,ii'}$	.38	-.07	.30	.29
$\rho_{G,ii'}$	.51	-.09	.51	.77
$\rho_{F,ii'}$	.90	.14	.84	.90
H'	.25	.49	.54	.32
$\%CV(\hat{\gamma})$	15	59	19	45
$R^2\%$	98	74	97	91

Notice that the two physical size pairs gave higher heritabilities than the mental tests showed. Model fit was fairly good for the R-V pair and also for the S-W pair of variables, but the sampling standard error of $\hat{\gamma}$ begins to rise when fitting to the L-B pair and the M-TT pair.

The correlations (ρ_{MI}, ρ_G and ρ_F) were quite similar except for the head sizes where $\rho_{MI} = -.07$, $\rho_G = -.09$ and $\rho_F = .14$. Surprisingly (or perhaps not so surprisingly), there is a very high correlation between the family effects of all the co-indicators except for the head sizes. For the body sizes the reason likely lies in nutrition differences among the families. For the mental tests it arises from the family culture. It is a pity we don't have some

duplicate (test halves, say) measurements that would allow separating the measurement errors from the individual effects.

We come now to our third objective of discovering the possible relationships between body size and mental measurements. With the four random components of genetic, family, individual and measurement effects, there are four levels of relationships to examine. Although we cannot separate the individual from the measurement variances, one can speculate that whatever correlation appears in the estimates of ρ_{MI} is entirely due to measurement and thus could become the basis of an estimate of correlated measurement error. The raw materials for this speculation, which is left to the reader, appear in Table 4 where estimates for α and ρ_{MI} are given for every pair of the seven variables. The other correlations appear in Table 5.

The estimates arise from successive applications of the nonlinear fitting procedure as developed for parallel variables to all pairs of variables. Thus one might expect to see fitting problems and indeed we do, as indicated by estimates of a few correlations on or beyond the boundaries of their ranges. Although the estimates for heritabilities become averages and thus lose their utilities, the estimates of correlations should nonetheless retain their reality.

In Table 5 we notice that fairly often (5 of 21 cases) the twin-pair correlations were estimated on their boundaries. That is, the data were such as to push them into unrealistic levels. In 2 of the 21 cases the genetic correlations become extreme. Roughly speaking, the variety of family situations seems even more diverse than the genetic ones among these 80 twin-pairs.

The cross correlations among the physical measurements seem quite logical. Genetic correlations between head sizes and body sizes are moderate, consistent, and positive (.22, .26, .32 and .33) while the family correlations with weight (.48 to head length and .38 to head breadth) could easily be driven by fatty tissue - presence or absence.

Table 4. Estimates of Proportion of Individual-measurement Variance, α (below diagonal) and of Individual-measurement Correlations, ρ_{MI} (above diagonal).

	M	R	V	L	B	S	W
M	1.00	.29	.08	-.25	.11	-.01	.02
R	.43	1.00	.38	-.30	-.13	.00	-.35
V	.21	.15	1.00	-.24	- 1	.19	.73
L	.42	.28	.18	1.00	-.07	.15	.47
B	.34	.28	.25	.28	1.00	-.40	.14
S	.09	.08	.08	.10	.09	1.00	.30
W	.22	.17	.12	.18	.16	.08	1.00

Table 5. Estimates of Genetic, ρ_G, and of Family, ρ_F, Correlations Between all Pairs of Variables (genetic correlations above, and family correlations below diagonal).

$\hat{\rho}_G \rightarrow$

$\hat{\rho}_F \downarrow$	M	R	V	L	B	S	W
M	1.00	1.1	.64	1.1	-.12	-.34	-.26
R	.72	1.00	.51	.43	.83	-.07	.97
V	.94	.90	1.00	.14	.77	-.51	.04
L	-1	.42	.53	1.00	-.09	.22	.26
B	.15	-1	-1	.14	1.00	.32	.33
S	1.1	.59	.66	1.2	.02	1.00	.51
W	.42	.03	.28	.48	.38	.84	1.00

The size and consistency of genetic correlations between the mental abilities and head sizes are suggestive (.43, .83, .14 and .77). Among this group of high school age twins the size of head seems tied to their mental

abilities scores. Perhaps the most interesting sector of relationships, the environmentally induced (family) correlations between mental and physical traits, appears muddied by extreme value estimates. Tallness seems to go with higher mental test scores and weight also to a slight extent, while head length and breadth are not consistently related.

Estimates of standard errors are provided by the nonlinear fitting routine (PROC NLIN) based on the number of data points (n = 12 in our examples) and seem fairly reasonable when the model fits. For example, the estimate $\hat{H} = .27$ for the R and V pair is given the standard error of .10. When we broke the 80 twins into 20 groups of 4 (basically one each of the zygosity by sex types) and applied a Tukey Jackknife, the less biased estimate became .28 and the standard error was .17. In our experience the Jackknife method overestimates the standard error and this is due to an occasional case (among the 20) of poor fit.

For the estimate of heritability, H', we have only the Tukey Jackknife. There, for the R and V pair still, the less biased estimate was .25 and the standard error was .15. A 60% sampling coefficient of variation is a bit disconcerting but one needs to recognize the rather small sample sizes of 40 monozygous and 40 dizygous twins-pairs and the wide variety of family situations - even in this southeastern Michigan subculture.

References

Clark, P. J., Vandenberg, S. G. and Proctor, C. H. (1961). "On the relationship of scores on certain psychological tests with a number of anthropometric characters and birth order in twins," *Human Biology*, vol. 335:163-180.

Proctor, Charles (1987). "Calculating significance levels for checking regression assumptions," *American Statistical Association Proceedings of the Social Statistics Section*, Alexandria, VA.

Proctor, C. H. (1985). "Simple inferences for infinite and finite population variance components and for fixed effects variation components," Institute of Statistics Mimeograph Series No. 1661, North Carolina State University, Raleigh, NC.

Appendix - Print of Data (4 pages) and Program Code (5 pages) with Output (10 pages)

SEVEN VARIABLES

OBS	FAM	O	ZYG	SEX	M	R	V	L	B	S	W
1	1	1	M	F	48	53	66	183	140	1682	103
2	1	2	M	F	35	42	61	188	138	1699	108
3	2	1	M	F	47	69	88	189	136	1620	118
4	2	2	M	F	53	74	84	186	140	1611	118
5	3	1	M	F	35	68	92	185	145	1503	128
6	3	2	M	F	42	61	86	186	140	1551	148
7	4	1	M	F	34	42	73	183	151	1562	110
8	4	2	M	F	26	38	68	185	147	1550	92
9	5	1	M	F	49	71	95	174	145	1614	124
10	5	2	M	F	38	72	97	186	143	1630	120
11	6	1	M	F	50	90	122	191	143	1736	116
12	6	2	M	F	46	82	101	191	141	1701	114
13	7	1	M	F	25	30	42	184	143	1551	112
14	7	2	M	F	28	37	43	186	143	1556	103
15	8	1	M	F	25	74	64	180	146	1603	122
16	8	2	M	F	41	78	65	179	144	1579	120
17	9	1	M	F	23	19	52	193	146	1680	119
18	9	2	M	F	23	36	59	191	145	1696	113
19	10	1	M	F	26	40	61	174	139	1552	86
20	10	2	M	F	16	37	50	174	138	1548	87
21	11	1	M	F	40	63	88	183	137	1695	106
22	11	2	M	F	40	54	90	192	147	1624	122
23	12	1	M	F	43	53	58	178	147	1650	114
24	12	2	M	F	48	60	85	177	150	1659	122
25	13	1	M	F	44	69	95	186	137	1695	106
26	13	2	M	F	34	86	106	185	140	1679	110
27	14	1	M	F	29	50	84	183	148	1587	100
28	14	2	M	F	43	69	96	186	146	1627	104
29	15	1	M	F	21	64	72	191	155	1632	143
30	15	2	M	F	33	59	63	183	158	1606	122
31	16	1	M	F	30	56	65	178	151	1610	97
32	16	2	M	F	26	51	75	176	154	1598	104
33	17	1	M	F	38	64	102	169	142	1509	88
34	17	2	M	F	31	48	95	174	145	1530	99
35	18	1	M	F	31	44	65	176	140	1652	102
36	18	2	M	F	33	31	60	177	138	1676	110
37	19	1	M	F	41	88	129	198	147	1736	138
38	19	2	M	F	52	85	130	194	147	1743	138
39	20	1	M	F	27	34	48	188	141	1555	111
40	20	2	M	F	38	55	46	189	138	1563	110
41	21	1	M	M	45	47	56	175	159	1486	68
42	21	2	M	M	37	40	55	177	150	1468	66
43	22	1	M	M	40	30	48	177	145	1450	96
44	22	2	M	M	41	25	59	181	144	1481	112
45	23	1	M	M	48	55	80	195	148	1735	120
46	23	2	M	M	50	58	60	196	147	1701	118
47	24	1	M	M	42	51	72	190	150	1792	142
48	24	2	M	M	46	55	57	191	148	1790	132
49	25	1	M	M	.	43	64	183	151	1661	117
50	25	2	M	M	.	49	54	171	150	1671	116
51	26	1	M	M	45	51	83	187	135	1624	101
52	26	2	M	M	31	57	101	181	137	1620	99
53	27	1	M	M	44	60	83	195	152	1725	130

SEVEN VARIABLES

OBS	FAM	O	ZYG	SEX	M	R	V	L	B	S	W
54	27	2	M	M	39	40	60	197	148	1709	126
55	28	1	M	M	40	43	46	186	156	1699	152
56	28	2	M	M	37	51	56	174	159	1710	133
57	29	1	M	M	44	66	81	203	154	1693	134
58	29	2	M	M	.	72	90	202	154	1702	141
59	30	1	M	M	37	37	46	180	148	1795	140
60	30	2	M	M	42	20	40	182	145	1807	141
61	31	1	M	M	27	36	37	185	148	1723	134
62	31	2	M	M	27	29	22	184	148	1705	120
63	32	1	M	M	45	72	98	186	160	1767	140
64	32	2	M	M	48	82	114	181	161	1768	143
65	33	1	M	M	51	46	128	193	157	1715	150
66	33	2	M	M	49	62	124	192	155	1674	138
67	34	1	M	M	44	47	92	192	140	1663	114
68	34	2	M	M	52	45	94	195	141	1632	122
69	35	1	M	M	44	64	64	186	149	1600	114
70	35	2	M	M	27	46	60	187	146	1589	115
71	36	1	M	M	29	49	38	193	150	1709	122
72	36	2	M	M	36	33	37	190	154	1697	140
73	37	1	M	M	27	20	39	192	147	1748	126
74	37	2	M	M	29	18	25	195	144	1755	134
75	38	1	M	M	48	62	71	199	153	1704	219
76	38	2	M	M	44	53	88	198	149	1685	220
77	39	1	M	M	33	30	86	199	145	1714	148
78	39	2	M	M	44	46	93	191	147	1665	134
79	40	1	M	M	39	54	51	190	156	1735	136
80	40	2	M	M	43	41	59	193	153	1734	136
81	41	1	M	M	47	73	86	191	151	1752	156
82	41	2	M	M	34	70	78	194	149	1760	162
83	42	1	M	M	54	63	92	208	154	1681	154
84	42	2	M	M	48	62	101	205	153	1689	145
85	43	1	M	M	32	39	93	191	148	1752	126
86	43	2	M	M	33	35	87	185	148	1760	124
87	44	1	D	F	48	85	73	185	146	1645	117
88	44	2	D	F	40	82	65	188	146	1649	130
89	45	1	D	F	39	69	119	197	147	1656	95
90	45	2	D	F	48	76	126	190	142	1613	90
91	46	1	D	F	50	75	115	179	151	1587	146
92	46	2	D	F	53	57	99	185	155	1644	138
93	47	1	D	F	44	62	59	183	141	1632	104
94	47	2	D	F	44	46	43	176	136	1569	96
95	48	1	D	F	36	31	65	169	135	1391	62
96	48	2	D	F	35	36	45	180	140	1440	86
97	49	1	D	F	35	68	87	190	145	1656	124
98	49	2	D	F	20	67	82	182	142	1654	116
99	50	1	D	F	42	75	116	191	138	1690	156
100	50	2	D	F	.	65	105	188	149	1590	138
101	51	1	D	F	50	88	110	185	144	1671	134
102	51	2	D	F	52	82	100	178	141	1674	121
103	52	1	D	F	23	72	85	182	135	1700	126
104	52	2	D	F	.	48	80	182	133	1630	115
105	53	1	D	F	41	65	119	174	146	1557	122
106	53	2	D	F	.	54	97	180	140	1614	147

SEVEN VARIABLES

OBS	FAM	O	ZYG	SEX	M	R	V	L	B	S	W
107	54	1	D	F	37	47	69	172	145	1594	96
108	54	2	D	F	33	37	46	175	146	1580	106
109	55	1	D	F	32	65	93	180	145	1642	110
110	55	2	D	F	42	67	71	173	145	1575	108
111	56	1	D	F	29	61	85	179	142	1589	116
112	56	2	D	F	37	61	85	181	143	1526	103
113	57	1	D	F	21	32	60	179	140	1605	124
114	57	2	D	F	17	30	32	181	134	1526	98
115	58	1	D	F	17	72	90	183	141	1650	132
116	58	2	D	F	32	60	80	191	146	1631	114
117	59	1	D	F	26	25	55	192	144	1587	113
118	59	2	D	F	22	41	41	174	141	1614	121
119	60	1	D	F	28	31	26	193	145	1635	138
120	60	2	D	F	11	34	24	191	140	1596	136
121	61	1	D	F	47	47	67	183	151	1672	120
122	61	2	D	F	40	53	63	184	145	1699	148
123	62	1	D	F	46	68	96	187	142	1675	132
124	62	2	D	F	43	58	82	177	137	1645	131
125	63	1	D	F	30	51	49	183	143	1601	114
126	63	2	D	F	27	52	49	179	140	1585	110
127	64	1	D	F	34	58	65	177	146	1591	112
128	64	2	D	F	47	52	78	182	140	1513	94
129	65	1	D	F	11	41	61	190	153	1640	124
130	65	2	D	F	36	41	63	182	139	1527	92
131	66	1	D	F	42	77	68	186	147	1541	103
132	66	2	D	F	44	56	67	187	147	1523	94
133	67	1	D	M	58	91	128	192	146	1808	136
134	67	2	D	M	54	73	129	195	144	1683	124
135	68	1	D	M	44	46	79	184	136	1545	101
136	68	2	D	M	42	34	42	189	136	1626	102
137	69	1	D	M	44	43	70	187	150	1638	110
138	69	2	D	M	43	36	58	187	151	1588	103
139	70	1	D	M	36	40	63	196	153	1857	144
140	70	2	D	M	42	39	63	186	159	1691	130
141	71	1	D	M	40	75	118	189	166	1697	155
142	71	2	D	M	.	57	78	190	155	1684	158
143	72	1	D	M	34	21	53	186	150	1745	122
144	72	2	D	M	45	31	70	194	147	1650	125
145	73	1	D	M	50	70	93	200	158	1749	132
146	73	2	D	M	45	67	109	198	156	1760	148
147	74	1	D	M	50	81	101	198	150	1679	107
148	74	2	D	M	41	47	96	195	144	1695	108
149	75	1	D	M	31	76	122	195	158	1706	134
150	75	2	D	M	43	70	75	199	148	1706	131
151	76	1	D	M	23	29	62	181	159	1745	127
152	76	2	D	M	26	29	42	190	152	1730	158
153	77	1	D	M	52	66	114	190	155	1758	134
154	77	2	D	M	42	69	120	182	159	1745	127
155	78	1	D	M	48	51	62	187	147	1733	126
156	78	2	D	M	30	35	49	186	154	1705	160
157	79	1	D	M	23	48	78	179	162	1678	123
158	79	2	D	M	38	62	87	187	165	1680	157
159	80	1	D	M	28	38	62	199	155	1844	153

SEVEN VARIABLES

OBS	FAM	O	ZYG	SEX	M	R	V	L	B	S	W
160	80	2	D	M	55	70	105	204	154	1818	174

```
NOTE: Copyright(c) 1989 by SAS Institute Inc., Cary, NC USA.
NOTE: SAS (r) Proprietary Software Release 6.08  TS404
      Licensed to NORTH CAROLINA STATE UNIVERSITY, Site 0027585013.

NOTE: Running on IBM Model 3090 Serial Number 114928.


 Welcome to the SAS Information Delivery System.

 This message is displayed for users in the SAS log when the NEWS
 option is specified.  You can replace this message with your own by
 editing the NEWS file.

NOTE: The SASUSER library was not specified. SASUSER library will now be the same as the WORK library.
NOTE: All data sets and catalogs in the SASUSER library will be deleted at the end of the session.  Use the NOWORKTERM option to
      prevent their deletion.

NOTE: SAS system options specified are:
      SORT=6

NOTE: The initialization phase used 0.15 CPU seconds and 1935K.
1          OPTIONS NONOTES;                                                                                    00000062
2          DATA D;INPUT M R V T L B C I E Z W S P;                                                              00000070
3            /***                                                                                               00000080
4          IF _N_=141 THEN B=.;                                                                                00000090
5          IF _N_=65 THEN R=.;                                                                                 00000100
6          IF _N_=109 THEN M=.;                                                                                00000110
7          IF _N_=105 THEN V=.;                                                                                00000120
8          IF ((_N_=75)+(_N_=76)) THEN DO;W=.;S=.;END;;                                                        00000130
9            ***/                                                                                              00000140
10         TITLE 'CLARK TWINS';                                                                                00000150
11         SEX='M';                                                                                            00000160
12         ZYG='D';                                                                                            00000170
13         TITLE2 'FULL DATA';                                                                                 00000180
14         PR=CEIL(_N_/2);                                                                                     00000190
15         GM=1;                                                                                               00000200
16         IF PR<21 THEN SEX='F';IF PR<44 THEN ZYG='M';                                                        00000210
17         IF ((PR>43)*(PR<67)) THEN SEX='F';                                                                  00000220
18         LABEL M='PROGRESSIVE MATRICES TEST' R='PRIMARY MENTAL A. REASONING'                                  00000230
19         V='PRIMARY MENTAL A. VERBAL' T='TOTAL MENTAL SCORES' L='HEAD LENGTH'                                 00000240
20         B='HEAD BREADTH' Z='BIZYGOMATIC BREADTH' C='HEAD CIRCUMFERENCE'                                      00000250
21         S='STATURE' W='WEIGHT' I='CEPHALIC INDEX' E='CROSS-SECTIONAL AREA'                                   00000260
22         P='RECIPROCAL PONDERAL INDEX';                                                                      00000270
23         CARDS;                                                                                              00000280
23                                                                                                             00000280
184         ;;                                                                                                 00001890
185        DATA D;SET D;                                                                                       00001900
186        O=2-MOD(_N_,2);                                                                                     00001910
187          FAM=CEIL(_N_/2);                                                                                  00001920
188          *CHANGING DEFINITION OF T;                                                                        00001930
189        T=(R+V)/2;                                                                                          00001940
190         *  USE SBS FOR TUKEY JACKKNIFE SUBGROUPS;                                                         00001950
191        SBS=MOD(FAM,20);IF SBS=0 THEN SBS=20;                                                               00001960
192        ARRAY VVVV (J) M R V T L B S W C I;ARRAY VVV (J) MM RR VV TT LL BB                                  00001970
193         SS WW CC II;                                                                                       00001980
194        DO OVER VVVV;VVV=VVVV;END;                                                                          00001990
195        PROC SORT;BY ZYG;                                                                                   00002000
196        %MACRO _BD;                                                                                         00002010
```

```
197        PROC GLM OUTSTAT=OT NOPRINT;                                                                    00002020
198        CLASSES SEX O  FAM;                                                                             00002030
199        MODEL MM RR VV TT CC II SS WW=SEX O FAM/SS1;                                                     00002040
200        MANOVA H=FAM;                                                                                   00002050
201        BY ZYG;                                                                                         00002060
202        DATA DOT;SET OT;KEEP _NAME_ RR VV _SOURCE_ DF ;                                                 00002070
203         /***                                                                                           00002080
204        IF _SOURCE_='SEX*O' THEN DELETE;                                                                00002090
205          ***/                                                                                          00002100
206        IF ((_SOURCE_='SEX')+(_SOURCE_='O')) THEN DELETE;                                               00002110
207        IF (_NAME_='RR')+(_NAME_='VV') THEN  OUTPUT;                                                    00002120
208        DATA DOT;SET DOT;                                                                               00002130
209        RETAIN OMS OOMS;                                                                                00002140
210        MS=RR/DF;OUTPUT;                                                                                00002150
211        IF MOD(_N_,2)=1 THEN DO;OMS=RR/DF;OOMS=VV/DF;END;                                               00002160
212        IF MOD(_N_,2)=0 THEN DO;MS=VV/DF;R=OOMS/SQRT(OMS*MS);                                           00002170
213        SG=MS/OMS;  OUTPUT;END;                                                                         00002180
214         TITLE3 'RR VV';                                                                                00002190
215        DATA DOT;SET DOT;N=_N_;                                                                         00002200
216        SUBS=MOD(N,3);IF SUBS=0 THEN SUBS=3;PROC SORT;BY SUBS;PROC MEANS NOPRINT                        00002210
217        ;;                                                                                              00002220
218         VAR MS;OUTPUT OUT=MNS MEAN=MMS ;BY SUBS;DATA MNS;SET MNS;                                      00002230
219         RETAIN OMMS;IF _N_=1 THEN OMMS=MMS;IF _N_=3 THEN G=OMMS/MMS;                                   00002240
220        DATA DOT;MERGE DOT MNS;BY SUBS;                                                                 00002250
221         NSUBS=4-SUBS;PROC SORT;BY NSUBS;DATA DOT;SET DOT;RETAIN STARTG STARTMS;                        00002260
222        IF _N_=1 THEN STARTG=1/G;                                                                       00002270
223        IF _N_=1 THEN STARTMS=MS;                                                                       00002280
224        DATA DRT;SET DOT;KEEP R;IF R=. THEN DELETE;                                                     00002290
225        PROC TRANSPOSE PREFIX=VR  OUT=DTT;                                                              00002300
226        DATA DTT;SET DTT;                                                                               00002310
227        H=4*(VR2-VR4)+2*(1/19)*(VR2+VR1-VR3-VR4);                                                       00002320
228        RF=.5*((1/19)+2)*VR4 +1.5*(1/19)*VR3-(2*(1/19)+H)*VR1;                                          00002330
229         RF=RF/(1-H);                                                                                   00002340
230        RG=2*(1/19)*(VR1-VR3)+H*VR1;RG=RG/H;                                                            00002350
231        RM=VR3;                                                                                         00002360
232        DATA ODTT;SET ODTT DTT;                                                                         00002370
233        DATA DOT;SET DOT;                                                                               00002380
234        KEEP N MS STARTG STARTMS DF;                                                                    00002390
235        PROC NLIN  DATA=DOT OUTEST=OOT NOPRINT;                                                         00002400
236        PARMS A=.3 H=.3 T=65 G=3 RM=.5 RG=.5 RF=.5;                                                     00002410
237        IF _ITER_=0 THEN IF _N_=1 THEN DO;                                                              00002420
238          T=STARTMS;G=STARTG;                                                                           00002430
239         END;                                                                                           00002440
240         BOUNDS -1<=RF<=1.2 ,-1<=RG<=1.2 ,-1<=RM<=1.2, 0<=A<=.8,0<=H<=1.2,T>0 ,                         00002450
241           G>0;                                                                                        00002460
242        _WEIGHT_=DF/MS**2;                                                                              00002470
243        IF N=7 THEN MODEL.MS=A*T ;                                                                      00002480
244        IF N=1 THEN MODEL.MS=A*T+.5*H*T;                                                                00002490
245        IF N=10 THEN MODEL.MS=A*T+2*T;                                                                  00002500
246        IF N=4 THEN MODEL.MS=A*T+(2-.5*H)*T;                                                            00002510
247        IF N=9 THEN MODEL.MS=G*(A*T);                                                                   00002520
248        IF N=3 THEN MODEL.MS=G*(A*T+.5*H*T);                                                            00002530
249        IF N=12 THEN MODEL.MS=G*(A*T+2*T);                                                              00002540
250        IF N=6 THEN MODEL.MS=G*(A*T+(2-.5*H)*T);                                                        00002550
251        IF N=8 THEN MODEL.MS=SQRT(G)*(RM*A*T);                                                          00002560
252        IF N=2 THEN MODEL.MS=SQRT(G)*(RM*A+.5*RG*H)*T;                                                  00002570
253        IF N=11 THEN MODEL.MS=SQRT(G)*(RM*A*T +2*(RG*H+RF*(1-H))*T);                                    00002580
254        IF N=5 THEN MODEL.MS=SQRT(G)*(RM*A*T+(1.5*RG*H+2*RF*(1-H))*T);                                  00002590
```

```
255          *TAKE DERIVATIVES;                                                    00002600
256         IF N=7 THEN DO;DER.A=T;DER.T=A;DER.G=0;DER.H=0;DER.RM=0;DER.RG=0;      00002610
257           DER.RF=0;END;                                                       00002620
258         IF N=1 THEN DO;DER.A=T;DER.T=A+.5*H;DER.G=0;DER.H=.5*T;DER.RM=0;      00002630
259         DER.RG=0;DER.RF=0;END;                                                00002640
260         IF N=10 THEN DO;DER.A=T;DER.T=A+2;DER.G=0;DER.H=0;DER.RM=0;           00002650
261         DER.RG=0;DER.RM=0;END;                                                00002660
262         IF N=4 THEN DO;DER.A=T;DER.T=A +2-.5*H;DER.G=0;DER.H=-.5*T;           00002670
263          DER.RM=0;DER.RG=0;DER.RF=0;END;                                      00002680
264         IF N=9 THEN DO;DER.A=G*T;DER.T=G*A;DER.G=A*T;DER.H=0;DER.RM=0;        00002690
265         DER.RG=0;DER.RF=0;END;                                                00002700
266         IF N=3 THEN DO;DER.A=G*T;DER.T=G*(A+.5*H);;DER.G=A*T+.5*H*T;          00002710
267         DER.H=G*.5*T;                                                         00002720
268         DER.RM=0;DER.RG=0;DER.RF=0;END;                                       00002730
269         IF N=12 THEN DO;DER.A=G*T;DER.T=G*(A+2);DER.G=(A+2)*T;DER.H=0;        00002740
270         DER.RM=0;DER.RG=0;DER.RF=0;END;                                       00002750
271         IF N=6 THEN DO;DER.A=G*T;DER.T=G*(A+2-.5*H);DER.G=A*T+(2-.5*H)*T;     00002760
272          DER.H=-.5*G*T;DER.RM=0;DER.RG=0;DER.RF=0;END;                        00002770
273         IF N=8 THEN DO;DER.A=SQRT(G)*RM*T;DER.T=SQRT(G)*RM*A;                 00002780
274          DER.G=.5*RM*A*T/SQRT(G);DER.H=0;DER.RM=SQRT(G)*A*T;DER.RG=0;DER.RF=0; 00002790
275           END;                                                                00002800
276         IF N=2 THEN DO;DER.A=SQRT(G)*RM*T;DER.T=SQRT(G)*(RM*A+.5*RG*H);      00002810
277         DER.G=.5*(RM*A+.5*RG*H)*T/SQRT(G);DER.H=SQRT(G)*.5*RG*T;              00002820
278          DER.RM=SQRT(G)*A*T;DER.RG=SQRT(G)*.5*H*T;DER.RF=0;END;               00002830
279         IF N=11 THEN DO;DER.A=SQRT(G)*RM*T;DER.T=SQRT(G)*(RM*A+2*(RG*H       00002840
280          +RF*(1-H)));DER.G=.5*(RM*A+2*(RG*H+RF*(1-H)))*T/SQRT(G);             00002850
281         DER.H=SQRT(G)*2*(RG-RF)*T;                                            00002860
282          DER.RM=SQRT(G)*A*T;DER.RG=SQRT(G)*2*H*T;DER.RF=SQRT(G)*2*(1-H)       00002870
283          *T;END;                                                              00002880
284         IF N=5 THEN DO;DER.A=SQRT(G)*RM*T;DER.T=SQRT(G)*(RM*A+1.5*RG*H+2*RF*  00002890
285          (1-H));                                                             00002900
286         DER.H=SQRT(G)*(1.5*RG-2*RF)*T;                                        00002910
287         DER.G=.5*(RM*A+1.5*RG*H+2*RF*(1-H))*T/SQRT(G);DER.RM=SQRT(G)*A*T;     00002920
288          DER.RG=SQRT(G)*1.5*H*T;DER.RF=2*SQRT(G)*(1-H)*T;END;                 00002930
289         OUTPUT OUT=NOT P=PMS R=RMS;                                           00002940
290         DATA ONOT;SET ONOT NOT;                                               00002950
291           DATA FOT;SET FOT OOT;                                               00002960
292         %MEND _BD;                                                            00002970
293          ;;                                                                   00002980
294         DATA FOT;INPUT A H T G RM RG RF ;                                     00002990
295         CARDS;                                                                00003000
295                                                                               00003000
297         ;;                                                                    00003020
298         DATA ODTT;INPUT VR1-VR4 ;CARDS;                                       00003030
298                                                                               00003030
300         ;;                                                                    00003050
300           DATA ONOT; INPUT MS;CARDS;                                          00003050
300                                                                               00003050
302         ;;                                                                    00003070
303         PROC PRINTTO PRINT=FT12F001 NEW;                                      00003080
304         DATA D1;SET D;IF SBS^=1;%_BD                                          00003090
WARNING: Variable SS already exists on file WORK.OT.
305           DATA D1;SET D;IF SBS^=2;%_BD                                        00003100
WARNING: Variable SS already exists on file WORK.OT.
306           DATA D1;SET D;IF SBS^=3;%_BD                                        00003110
WARNING: Variable SS already exists on file WORK.OT.
307           DATA D1;SET D;IF SBS^=4;%_BD                                        00003120
WARNING: Variable SS already exists on file WORK.OT.
```

```
308         DATA D1;SET D;IF SBS^=5;%_BD                                                00003130
WARNING: Variable SS already exists on file WORK.OT.
309         DATA D1;SET D;IF SBS^=6;%_BD                                                00003140
WARNING: Variable SS already exists on file WORK.OT.
310         DATA D1;SET D;IF SBS^=7;%_BD                                                00003150
WARNING: Variable SS already exists on file WORK.OT.
311         DATA D1;SET D;IF SBS^=8;%_BD                                                00003160
WARNING: Variable SS already exists on file WORK.OT.
312         DATA D1;SET D;IF SBS^=9;%_BD                                                00003170
WARNING: Variable SS already exists on file WORK.OT.
313         DATA D1;SET D;IF SBS^=10;%_BD                                               00003180
WARNING: Variable SS already exists on file WORK.OT.
314         DATA D1;SET D;IF SBS^=11;%_BD                                               00003190
WARNING: Variable SS already exists on file WORK.OT.
315         DATA D1;SET D;IF SBS^=12;%_BD                                               00003200
WARNING: Variable SS already exists on file WORK.OT.
316         DATA D1;SET D;IF SBS^=13;%_BD                                               00003210
WARNING: Variable SS already exists on file WORK.OT.
317         DATA D1;SET D;IF SBS^=14;%_BD                                               00003220
WARNING: Variable SS already exists on file WORK.OT.
318         DATA D1;SET D;IF SBS^=15;%_BD                                               00003230
WARNING: Variable SS already exists on file WORK.OT.
319         DATA D1;SET D;IF SBS^=16;%_BD                                               00003240
WARNING: Variable SS already exists on file WORK.OT.
320         DATA D1;SET D;IF SBS^=17;%_BD                                               00003250
WARNING: Variable SS already exists on file WORK.OT.
321         DATA D1;SET D;IF SBS^=18;%_BD                                               00003260
WARNING: Variable SS already exists on file WORK.OT.
322         DATA D1;SET D;IF SBS^=19;%_BD                                               00003270
WARNING: Variable SS already exists on file WORK.OT.
323         DATA D1;SET D;IF SBS^=20;%_BD                                               00003280
WARNING: Variable SS already exists on file WORK.OT.
324         DATA D1;SET D;                                                              00003290
325          %_BD                                                                       00003300
WARNING: Variable SS already exists on file WORK.OT.
326                                                                                     00003310
326          PROC PRINTTO;                                                              00003310
327         PROC PRINT DATA=ONOT;TITLE3 'INPUT MEAN SQUARES & CROSS PRODUCTS';          00003320
328         PROC PRINT DATA=ODTT;TITLE3 'ESTIMATES FROM CORRELATIONS';                  00003330
329         DATA FOT;SET FOT;IF _TYPE_='FINAL';                                         00003340
330         HT=H/(1+A-.05);                                                             00003350
331         PROC PRINT DATA=FOT;                                                        00003360
332         TITLE3 'ESTIMATES AND TUKEY JACKKNIFE CALCULATIONS FOR EACH';               00003370
333         PROC IML;                                                                   00003380
IML Ready
334         START TJK;                                                                  00003390
334                                                                                     00003390
335                                                                                     00003400
335          PRINT 'THIS PROGRAM DOES THE TUKEY JACKKNIFE';                             00003400
335                                                                                     00003400
336         PRINT 'LAST VALUE IN TJKIN IS ESTIMATE BASED ON ALL DATA',                  00003410
337          'AND OTHER VALUES ARE BASED ON OMIT-ONE-SUBSAMPLE DATA', TJKIN;            00003420
337                                                                                     00003420
338         NRP=NROW(TJKIN)-1;                                                          00003430
338                          NRPL1=NRP-1;                                               00003430
338                                       NRPP1=NRP+1;                                  00003430
338                                                                                     00003430
339         PV=J(NRP,1);                                                                00003440
```

```
339                                                                                                  00003440
340         PV=TJKIN[NRPP1,]#J(NRP,1,NRP)-NRPL1#TJKIN[1:NRP,];                                       00003450
340                                                                                                  00003450
341         PRINT 'PSUEDO VALUES ARE IN PV' ,PV;                                                     00003460
341                                                                                                  00003460
342         BE=TJKIN[NRPP1,];                                                                        00003470
342                                                                                                  00003470
343         LBE=SUM(PV)/NRP;                                                                         00003480
343                                                                                                  00003480
344         SE=SSQ(PV-LBE#J(NRP,1))/(NRP*(NRP-1));                                                   00003490
344                                                SE=SQRT(SE);                                      00003490
344                                                                                                  00003490
345         PRINT 'BE IS BIASED EST., LBE IS LESS BIASED EST. AND SE IS STD. ERR.'                   00003500
346          ,BE LBE SE;                                                                             00003510
346                                                                                                  00003510
347         PB=(LBE-BE)/SE;                                                                          00003520
347                          PCTSCV=100*SE/LBE;                                                      00003520
347                                                                                                  00003520
348         PRINT 'PB IS PROPORTION BIAS AND PCTSCV IS PERCENT SAMPLING CV',                         00003530
349           PB PCTSCV;                                                                             00003540
349                                                                                                  00003540
350         FINISH;                                                                                  00003550
350                                                                                                  00003550
351         START CONTROL;                                                                           00003560
351                                                                                                  00003560
352         USE FOT;                                                                                 00003570
352                                                                                                  00003570
353         READ ALL VAR {A H T G RM RG RF HT} INTO FOT;                                             00003580
353                                                                                                  00003580
354                                                                                                  00003590
354          DO I1=1 TO 8;                                                                           00003590
354                                                                                                  00003590
355                                                                                                  00003600
355          TJKIN=FOT[,I1];                                                                         00003600
355                                                                                                  00003600
356         RUN TJK;                                                                                 00003610
356                                                                                                  00003610
357         END;                                                                                     00003620
357                                                                                                  00003620
358         FINISH;                                                                                  00003630
358                                                                                                  00003630
359         RUN CONTROL;                                                                             00003640
359                                                                                                  00003640
Exiting IML.
```

ESTIMATES AND TUKEY JACKKNIFE CALCULATIONS FOR EACH

OBS	A	H	T	G	RM	RG	RF	_TYPE_	_NAME_	_ITER_	_SSE_	HT
1	0.15628	0.25236	271.795	1.85841	0.43176	0.45479	0.90459	FINAL		.	8.0357	0.22812
2	0.15498	0.26142	270.550	1.83100	0.47233	0.39533	0.91085	FINAL		.	9.2170	0.23658
3	0.14560	0.26879	281.953	1.78243	0.47949	0.39639	0.95666	FINAL		.	8.3330	0.24534
4	0.14958	0.27945	273.063	1.82858	0.47214	0.44491	0.94730	FINAL		.	9.6462	0.25415
5	0.15923	0.27763	267.885	1.77177	0.43154	0.42136	0.92770	FINAL		.	7.3858	0.25029
6	0.13820	0.28149	266.196	1.82931	0.43515	0.50229	0.86903	FINAL		.	12.5498	0.25868
7	0.14675	0.29164	261.804	1.82128	0.37349	0.55412	0.86675	FINAL		.	7.7687	0.26592
8	0.15768	0.24175	270.298	1.78075	0.41788	0.50174	0.93007	FINAL		.	7.6152	0.21825
9	0.15147	0.29644	262.411	1.88786	0.41753	0.46402	0.91472	FINAL		.	9.3372	0.26913
10	0.15623	0.27344	265.581	1.82070	0.42410	0.45957	0.88266	FINAL		.	8.5311	0.24718
11	0.15728	0.28094	265.953	1.78144	0.43523	0.44502	0.90905	FINAL		.	7.6184	0.25372
12	0.13012	0.30130	265.783	1.80319	0.42107	0.40791	0.92520	FINAL		.	13.0371	0.27895
13	0.15958	0.26361	266.427	1.72873	0.41487	0.50376	0.90644	FINAL		.	6.2258	0.23757
14	0.14457	0.20995	272.116	1.91278	0.40097	0.63138	0.83736	FINAL		.	7.3210	0.19181
15	0.15557	0.20902	279.662	1.73005	0.41786	0.61803	0.86582	FINAL		.	7.4288	0.18906
16	0.15482	0.27974	273.790	1.78361	0.45424	0.44732	0.91509	FINAL		.	6.6664	0.25320
17	0.16120	0.28585	258.342	1.78023	0.42657	0.47797	0.89535	FINAL		.	8.0897	0.25725
18	0.14567	0.27752	273.242	1.84017	0.48966	0.38242	0.92492	FINAL		.	8.9559	0.25329
19	0.16341	0.24181	252.586	1.86423	0.41806	0.53364	0.84847	FINAL		.	8.9191	0.21718
20	0.15741	0.13334	263.919	1.79819	0.48340	-0.37714	0.94800	FINAL		.	4.2310	0.12041
21	0.15236	0.26128	268.666	1.80979	0.43526	0.44851	0.90512	FINAL		.	8.4718	0.23702

FULL DATA

ESTIMATES AND TUKEY JACKKNIFE CALCULATIONS FOR EACH

THIS PROGRAM DOES THE TUKEY JACKKNIFE

LAST VALUE IN TJKIN IS ESTIMATE BASED ON ALL DATA
AND OTHER VALUES ARE BASED ON OMIT-ONE-SUBSAMPLE DATA

TJKIN
0.1562789
0.1549831
0.1455967
0.1495783
0.1592344
0.1381976
0.1467486
0.1576753
0.1514652
0.156232
0.1572798
0.1301245
0.1595846
0.1445669
0.1555664
0.1548225
0.1611961
0.1456689
0.1634099
0.1574113
0.1523594

PSUEDO VALUES ARE IN PV

PV
0.0778887
0.1025095
0.2808509
0.2052017
0.0217344
0.4214331
0.2589654
0.0513574
0.1693496
0.0787808
0.058872
0.5748227
0.0150801
0.3004169
0.0914262
0.1055612
-0.015538
0.2794789
-0.057599
0.0563733

BE IS BIASED EST., LBE IS LESS BIASED EST. AND SE IS STD. ERR.

BE	LBE	SE
0.1523594	0.1538483	0.0353758

CLARK TWINS
FULL DATA
ESTIMATES AND TUKEY JACKKNIFE CALCULATIONS FOR EACH

PB IS PROPORTION BIAS AND PCTSCV IS PERCENT SAMPLING CV

PB	PCTSCV
0.0420874	22.993964

THIS PROGRAM DOES THE TUKEY JACKKNIFE

LAST VALUE IN TJKIN IS ESTIMATE BASED ON ALL DATA
AND OTHER VALUES ARE BASED ON OMIT-ONE-SUBSAMPLE DATA

TJKIN
0.2523629
0.2614177
0.2687923
0.2794533
0.2776347
0.2814935
0.291643
0.2417544
0.2964413
0.2734384
0.2809415
0.3013017
0.2636055
0.2099512
0.2090167
0.279738
0.2858542
0.2775183
0.2418074
0.1333391
0.2612811

PSUEDO VALUES ARE IN PV

PV
0.430726
0.2586859
0.1185685
-0.08399
-0.049438
-0.122755
-0.315594
0.6322876
-0.406762
0.0302916
-0.112267
-0.49911
0.2171177
1.2365497
1.2543054
-0.0894
-0.205608
-0.047226
0.6312823
2.69218

FULL DATA

ESTIMATES AND TUKEY JACKKNIFE CALCULATIONS FOR EACH

BE IS BIASED EST., LBE IS LESS BIASED EST. AND SE IS STD. ERR.

BE	LBE	SE
0.2612811	0.2784922	0.1663518

PB IS PROPORTION BIAS AND PCTSCV IS PERCENT SAMPLING CV

PB	PCTSCV
0.1034623	59.733014

THIS PROGRAM DOES THE TUKEY JACKKNIFE

LAST VALUE IN TJKIN IS ESTIMATE BASED ON ALL DATA
AND OTHER VALUES ARE BASED ON OMIT-ONE-SUBSAMPLE DATA

TJKIN
271.79487
270.55025
281.95266
273.06338
267.88536
266.19552
261.80421
270.29821
262.41056
265.58096
265.95274
265.78275
266.4271
272.11619
279.66201
273.78975
258.34158
273.24167
252.58555
263.91865
268.66611

PSUEDO VALUES ARE IN PV

PV
209.21965
232.86735
16.221659
185.11781
283.50023
315.60715
399.04207
237.65608
387.52143
327.28389
320.2201
323.44992
311.20727
203.11445
59.743899

FULL DATA
ESTIMATES AND TUKEY JACKKNIFE CALCULATIONS FOR EACH

171.31694
464.83205
181.73037
574.19669
358.86776

BE IS BIASED EST., LBE IS LESS BIASED EST. AND SE IS STD. ERR.

BE	LBE	SE
268.66611	278.13584	29.203977

PB IS PROPORTION BIAS AND PCTSCV IS PERCENT SAMPLING CV

PB	PCTSCV
0.3242618	10.499897

THIS PROGRAM DOES THE TUKEY JACKKNIFE

LAST VALUE IN TJKIN IS ESTIMATE BASED ON ALL DATA
AND OTHER VALUES ARE BASED ON OMIT-ONE-SUBSAMPLE DATA

TJKIN
1.8584131
1.8310034
1.7824267
1.8285845
1.7717655
1.829308
1.8212809
1.7807536
1.8878575
1.8206984
1.781442
1.803194
1.7287295
1.9127845
1.7300474
1.7836107
1.7802316
1.840173
1.8642304
1.7981908
1.8097884

PSUEDO VALUES ARE IN PV

PV
0.885918
1.4067026
2.3296608
1.4526618
2.5322221
1.4389149
1.5914296
2.3614498
0.3264749

FULL DATA

ESTIMATES AND TUKEY JACKKNIFE CALCULATIONS FOR EACH

1.6024979
2.3483701
1.9350819
3.349907
-0.147138
3.3248662
2.3071649
2.3713673
1.2324813
0.7753889
2.0301422

BE IS BIASED EST., LBE IS LESS BIASED EST. AND SE IS STD. ERR.

BE	LBE	SE
1.8097884	1.7727782	0.2016929

PB IS PROPORTION BIAS AND PCTSCV IS PERCENT SAMPLING CV

PB	PCTSCV
-0.183498	11.377224

THIS PROGRAM DOES THE TUKEY JACKKNIFE

LAST VALUE IN TJKIN IS ESTIMATE BASED ON ALL DATA
AND OTHER VALUES ARE BASED ON OMIT-ONE-SUBSAMPLE DATA

TJKIN
0.4317577
0.4723291
0.4794878
0.4721375
0.4315377
0.4351536
0.3734886
0.4178837
0.4175261
0.4241013
0.4352336
0.42107
0.4148675
0.4009715
0.417861
0.4542418
0.4265701
0.489658
0.4180613
0.4834025
0.4352602

PSUEDO VALUES ARE IN PV

PV
0.5018069
-0.269048
-0.405063

FULL DATA

ESTIMATES AND TUKEY JACKKNIFE CALCULATIONS FOR EACH

-0.265408
0.5059873
0.4372864
1.608921
0.7654134
0.7722088
0.6472794
0.4357654
0.7048735
0.8227208
1.0867446
0.7658453
0.0746093
0.6003728
-0.598298
0.7620386
-0.479444

BE IS BIASED EST., LBE IS LESS BIASED EST. AND SE IS STD. ERR.

BE	LBE	SE
0.4352602	0.4237306	0.1284569

PB IS PROPORTION BIAS AND PCTSCV IS PERCENT SAMPLING CV

PB	PCTSCV
-0.089754	30.315707

THIS PROGRAM DOES THE TUKEY JACKKNIFE

LAST VALUE IN TJKIN IS ESTIMATE BASED ON ALL DATA
AND OTHER VALUES ARE BASED ON OMIT-ONE-SUBSAMPLE DATA

TJKIN
0.4547895
0.3953255
0.3963917
0.4449134
0.4213569
0.5022919
0.5541219
0.5017368
0.4640191
0.4595691
0.445018
0.4079126
0.5037595
0.6313808
0.6180298
0.4473223
0.4779682
0.38242
0.533635
-0.377139
0.4485145

FULL DATA
ESTIMATES AND TUKEY JACKKNIFE CALCULATIONS FOR EACH

PSUEDO VALUES ARE IN PV

PV
0.3292899
1.4591056
1.4388491
0.5169369
0.9645099
-0.573255
-1.558025
-0.562707
0.1539282
0.2384784
0.5149483
1.219952
-0.60114
-3.025943
-2.772275
0.4711673
-0.111105
1.7043107
-1.168774
16.135924

BE IS BIASED EST., LBE IS LESS BIASED EST. AND SE IS STD. ERR.

BE	LBE	SE
0.4485145	0.7387087	0.860805

PB IS PROPORTION BIAS AND PCTSCV IS PERCENT SAMPLING CV

PB	PCTSCV
0.3371195	116.52834

THIS PROGRAM DOES THE TUKEY JACKKNIFE

LAST VALUE IN TJKIN IS ESTIMATE BASED ON ALL DATA
AND OTHER VALUES ARE BASED ON OMIT-ONE-SUBSAMPLE DATA

TJKIN
0.9045878
0.9108517
0.9566598
0.9473003
0.927699
0.8690321
0.8667473
0.9300746
0.9147177
0.8826648
0.9090452
0.9252011
0.9064409
0.8373565
0.8658174
0.9150929

ESTIMATES AND TUKEY JACKKNIFE CALCULATIONS FOR EACH

0.8953464
0.9249242
0.8484701
0.9480002
0.9051237

PSUEDO VALUES ARE IN PV

PV
0.9153066
0.7962907
-0.074063
0.1037677
0.476193
1.5908634
1.634276
0.431056
0.7228385
1.331842
0.8306159
0.5236539
0.8800961
2.1927013
1.6519434
0.7157084
1.090893
0.528914
1.9815425
0.0904702

BE IS BIASED EST., LBE IS LESS BIASED EST. AND SE IS STD. ERR.

BE	LBE	SE
0.9051237	0.9207455	0.1417516

PB IS PROPORTION BIAS AND PCTSCV IS PERCENT SAMPLING CV

PB	PCTSCV
0.1102053	15.39531

THIS PROGRAM DOES THE TUKEY JACKKNIFE

LAST VALUE IN TJKIN IS ESTIMATE BASED ON ALL DATA
AND OTHER VALUES ARE BASED ON OMIT-ONE-SUBSAMPLE DATA

TJKIN
0.2281187
0.2365807
0.2453387
0.2541459
0.250294
0.2586787
0.265916
0.2182539
0.2691336
0.24718

FULL DATA

ESTIMATES AND TUKEY JACKKNIFE CALCULATIONS FOR EACH

0.2537222
0.2789509
0.2375713
0.1918121
0.1890584
0.2531972
0.2572491
0.2532866
0.2171773
0.1204061
0.2370199

PSUEDO VALUES ARE IN PV

PV
0.4061417
0.2453641
0.078962
-0.088374
-0.015188
-0.174497
-0.312006
0.5935738
-0.373141
0.0439773
-0.080325
-0.559669
0.2265426
1.0959679
1.1482873
-0.07035
-0.147336
-0.072048
0.6140288
2.4526822

BE IS BIASED EST., LBE IS LESS BIASED EST. AND SE IS STD. ERR.

BE	LBE	SE
0.2370199	0.2506297	0.1532083

PB IS PROPORTION BIAS AND PCTSCV IS PERCENT SAMPLING CV

PB	PCTSCV
0.0888322	61.129359

Part II. Model Reformulation

In the first part of this report a statistical analysis was designed and carried out on a body of data on twins. In this second part we wish to extend that analysis and to question it, and we would also like to comment briefly on statistical methods and findings in the study of inheritance of human traits from looking at data on twins. We'll begin by offering our assessment of the statistical methods used in three reviews or meta-analyses of studies on the inheritance of IQ. [These are among the reviews that underlay the findings of the currently controversial publication, *The Bell Curve (1994)*.] Then we'll discuss in the same spirit the statistical problems we encountered fitting to twin data and give some more results. By editorial standards this discussion of general topics should perhaps have preceeded the earlier analyses but, in fact, I knew hardly any of the literature then and not too much more now, so this order of topics is more faithful to the way the ideas were developed.

Our interest in our "old" twin data was renewed when Professor Olkin gave a series of seminars on meta-analysis (at NCSU, April, 1995) and cited, as an example, the Erlenmeyer-Kimling and Jarvik (1963) review of intraclass correlations on intelligence tests for twins and for other kinds of pairs. We had found a way of doing the types of linear modelling used by Hedges and Olkin (1985) by using the FREQ statement in PROC GLM of SAS, so we were eager to apply it to our "old" data.

Since the twins were identified by birth order we used, at first, the ordinary ("interclass") correlation coefficient. In retrospect, we should have continued thus, since that approach adjusts for the order of birth effects

which appear to be real. A quick glance into Fisher (1958, pp. 211 ff) gave a formula for r, the intraclass correlation, and a variance for the transform $Z = 1/2[\log(i+r)-\log(1-r)]$ of $(n-1.5)^{-1}$ where n is the number of pairs. His formula (due to "Harris") uses the between family and within family sums of squares as $r_1 = (BSS-WSS)/(BSS + WSS)$.

This formula goes back to Pearson in 1901 and in those days no great effort was made to separate sample statistics from population parameters. More recently such data are viewed as generated by adding a family variance component to an individual component. In this view the population correlation becomes $\sigma_B^2/(\sigma_B^2 + \sigma_W^2)$ with σ_B^2 the between families and σ_W^2 the within families variance components. Using ANOVA estimates for σ_B^2 and σ_W^2 suggests estimating the ratio as $r_2 = (BMS - WMS)/(BMS + WMS)$. [The reader may wish to verify that with 3 pairs of twins having observations of (1,2), (2,3) and (3,5) one finds the quite different values of $r_1 = 9/19$ and $r_2 = 16/27$.]

The results from use of the r_2 formula appear in Table 6. Notice how well the female dizygous twins correlation holds up for R and V (reasoning and verbal abilities). Do girls communicate better?

Table 6. Intraclass Correlations (Using the r_2 Formula) for Four Subsamples and Eight Study Variables

		M	R	V	L	B	S	W	T
Dizygous	Females	.62	.79	.86	.38	.38	.64	.66	.87
	Males	.28	.65	.59	.60	.73	.55	.61	.66
Monozygous	Females	.56	.82	.91	.78	.79	.92	.77	.89
	Males	.55	.77	.89	.83	.87	.97	.95	.88

Faced with the "data" in Table 6 one makes the transform to Z and arrives at the Z_{ijk}, say, where i = 1, 2 for dizygous and monozygous, j = 1, 2 for the sexes and k = 1, 2, ..., 8 indexes the variables. In the data step new values are defined as $Z_{ijk} + 1$ and $Z_{ijk} - 1$ each with an associated frequency of [(n-1.5)/2]. Finally one fits a full factorial model with zygosity, sex and variable as factors using the frequencies as defined. This is the computational scheme we found capable of doing the analyses of means or other statistics, when one is given their sample variances, without having to know the individual observations. The method "reconstitutes" individual values having the same mean and variance properties as the original observations had.

Although we know (from the factor analyses in the firt part) that there are correlations among the variables it may be the case that, after fitting to all the interactions in this model, the residuals are not so much correlated. In any event the P-values suggest that in our data there are strong "zygosity" and "variables" main effects with some zygosity-by-variables interaction and there was also a hint of zygosity-by-sex interaction.

The zygosity-by-variables interaction (at the .03 level) suggests that heritability (the drop in correlation from monozygous to dizygous) differs from one variable to the next, while the zygosity-by-sex interaction (at the one-tailed .04 level) suggests that environmental influences are greater for females (which is what we noted earlier in connection with the r values of Table 6).

This approach is meta-analytic in style only. We lack the many studies, each with their r_2's and sample sizes for the four groups and several variables. This lack is our own fault since Erlenmeyer-Kimling and Jarvik (1963) have presumably published their raw data. Examining Figure 1 of

Erlenmeyer-Kimling and Jarvik (1963) I find that for monozygous twins that r is about .85 as compared to .60 for dizygous with notably greater variability from study-to-study among the dizygous twins.

These results are most convincing (even without a full-blown meta-analysis calculation) that genetic and environmental effects are both present. However, being expressed as differences between intraclass correlations they lack the full impact that a single summarizing quantity such as "heritability" carries. This lack has been corrected or, at least, attempts have been made to supply the needed results. One review we looked at was by Christopher Jencks (1972) and the other was done by Raymond B. Cattell (1982). Both are monumental and complicated and, to a statistician (to myself at any rate), unsettling.

Both reviews covered twin data but were mainly concerned with studies having data on parents, other siblings and pairs living apart as well. Both gave complex expressions for the numerous correlations in terms of numerous parameters and then proceeded to plug in observed correlations and solve for these parameters. This is the "path-analysis" approach which has roots in genetics [see Wright, S. (1978)], but in neither review did I find the model equations ("accounting equations") underlying the path diagrams.

With trepidation we offer an overly general characterization of such a modelling style. When an investigator includes latent (unobserved) variables into path diagrams some reification is bound to occur. In particular, when a variable G (genetic effect) and a variable E (environmental effect) are penned in, some reality seems immediately to attach to them, so that it is quite natural to allow for their correlation. If the investigator is the creator of the variable (as one is when the variable is unobserved) then one "should"

assure that the variable carry as little baggage by way of additional parameters as possible. That is, in general, latent variables should have unit variance and be uncorrelated with other latent variables.

The correlation, denoted "s", introduced by Jencks between genetic effect and environmental effect is, we believe, unnecessary. We suspect that the phenomena which s is taken to represent is the very real interplay between evidence of mental ability or its lack at time t by a child, and responses of discouragement or encouragement from the family at time t+1. However, data collected at one point in time (as all data of the type under review were) can only with difficulty be used to estimate such a relationship, when no clean measures exist for genetic effect distinct from environment effect.

One encounters the parameter s early in Jencks' (1972) review so we early lost interest in the details of his analysis, but the general lines were apparent and the substantive results seem quite reasonable. He corrects for attenuation which is welcome. There is, however, a correction for "cultural differences" that tends to sharpen the results when merging several studies. This may be due to treating cultural differences as a fixed factor rather than a random one and thus ignoring the interaction between culture and heritability.

In the other review Cattell (1982), one quickly finds (p. 68) an ambiguity in his formula for the calculation of the intraclass correlation coefficient. Was it r_1 or r_2? His derivations started with variances written as σ^2's, which were population quantities, expressed in terms of numerous parameters but the parameters were estimated from sample correlations. He gives excellent references, but we had such difficulties tracking down the origin of his

expressions that we gave up. Needless to say he provided no model equations for the raw data.

Among Cattell's (1982) references we did find one that arrived at essentially the same expressions for expectations of mean squares as we had derived. This was a paper by Haseman and Elston (1970). Their paper also described the weighted least squares method of estimating parameters, all done for a single variable. In tracing the more recent literature to see if they or others have added a second variable within this same components model we are uncertain. Early on Kempthorne and Osborne (1961) clearly advocated computing the between family and within family mean squares and mean cross products but I don't find formulas for estimates of the parameters. They compiled a list of eleven "forces," in the cautious style of overkill by confounding, but seemed unwilling to reduce it to a list of estimable variance components.

We did locate a monumental collection of mean squares and mean cross products for 14 mental abilities variables on 123 monozygous and 75 dizygous twins. See Loehlin and Vandenberg (1968). The aim of these researchers was to provide separate correlation matrices for genetic and for environmental components. Notice the close parallel to our results, and also notice that Stephen Vandenberg was behind both studies. Even the same twins are on the data sets.

When I modified my computer program to accommodate more than two variables and to fit to these mean squares, I chose to introduce a separate variance parameter for each variable rather than the one base variance with a ratio parameter for other variables as before. I also used a more accurate weighting for the mean cross product and the product of the two mean squares divided by

degrees of freedom. In the earlier calculations we had dropped all observations of any cases with missing values and presently they are back in. However, we still have not taken account of the correlations among the mean squares.

I believe that due to a combination of these reasons (reparameterizing, reweighting and ignoring covariances among data points) the new model seems more difficult to fit. I found it necessary to restart the PROC NLIN with different methods alternating between MARQUARD and GRADIENT and then using GAUSS in later stages, plus reordering the mean squares on the input data set. In view of these problems we are deferring the fitting of that body of mean squares and cross products. Let's just look at results for the same bivariate cases as were treated above and then we'll make a few summary comments.

A comparison between Tables 3 and 7 gives some cause for concern and some encouragement. Reparameterizing the scale factors into G1 and G2 from the original σ^2_{Ti} and γ gave somewhat different results. We should have found $G1^2 = \sigma^2_{Ti}$ and $G2/G1 = \gamma$ and, although the first equality seems reasonable, the G2/G1 ratio stayed too close to 1. The correlations also seemed to wander quite a bit, but this is to be expected from the small smaple sizes. The values for the heritabilities agree quite well between Table 3 and Table 7. This is most encouraging. The physical sizes have higher heritabilities than the mental tests.

Table 7. Parameter Estimates in Fitting (Using Corrected Weights and Two Scale Parameters) to Mean Squares for Four Pairs of Variables. Heritability is Estimated as $H' = H/(1 + \alpha - .05)$.

	Variables Playing Roles of *ith* and *i'th*			
Parameter	Reasoning & Verbal	Head Length & Breadth	Stature & Weight	Raven's Matrices & Ave. of R and V
α	.26	.22	.06	.17
H	.27	.64	.51	.46
γ_1	15.0	6.6	70.0	8.2
γ_2	15.9	4.9	20.0	14.5
$\rho_{MI,ii'}$	.36	-.25	.54	1.02
$\rho_{G,ii'}$	.68	.31	.58	.31
$\rho_{F,ii'}$	1.34	.08	.75	1.05
H'	.22	.55	.50	.41

In view of the sex differences that appeared in Table 6 we ran the fitting program separately for males and females and the results are seen in Table 8. Although no differences for estimates in Table 7 had been observed when outliers were omitted, the sample sizes for these comparisons by sex in Table 8 are getting quite small. However, when the putative outliers were dropped the only difference(in all of Table 8) was an increase from .10 to .19 in the estimate of H for Females using the variables M and TT (Raven's Matrices and the average of R and V). The basic findings in Table 8 are then that on the mental tests the males show more heritability than the females, while on the head sizes the females show higher heritability.

Table 8. Fitting Separately for Males and Females Using the Same Model as for Table 7.

	Reasoning & Verbal		Raven's Matrices & R & V Average		Stature & Weight		Head Length & Breadth	
	Males	Females	Males	Females	Males	Females	Males	Females
α	.28	.22	.15	.11	.05	.07	.20	.26
H	.31	-.03	.51	.10	.50	.49	.46	.93
γ_1	15.0	15.0	8.2	8.2	70.0	70.0	6.9	6.2
γ_2	16.0	16.1	14.5	14.8	20.0	20.0	4.7	4.3
$\rho_{MI,ii'}$	.36	.26	1.00	.99	.50	.51	-.59	.19
$\rho_{G,ii'}$	.68	.67	.42	.36	.59	.31	.34	.14
$\rho_{F,ii'}$	.95	1.39	1.25	1.50	.68	.40	.25	-.07
H'	.25	-.03	.46	.09	.50	.48	.40	.77

Having indulged in criticism of others' methods let's mention a criticism we have received. Since we have given a model equation and seem to have verified normality of distributions, one could ask that we write the likelihood and maximize it. This is generally the approach we find in recent pedigree studies. We claim that the within family and between family mean squares and mean cross products are sufficient statistics and thus modelling these sample moments enjoys the same properties as any procedure that uses weighted least squares in place of maximizing the Wishart likelihood. This claim needs to be verified. In more detail it is claimed that the within are independent of the between mean squares. If the claim is true we must still take account of the correlations among some of the mean squares and cross products.

I feel sure that others have faced these fitting problems and indeed there was mention of many standard methods. On the other hand I found very few numerical examples showing which method worked. I suspect no one best method has yet been found. This, I believe, is due to the multiplicity of issues that we are hoping to learn about from twin data. Even in the bivariate case there are seven parameters based on 12 data points. In linear models one is nervous when there are more than 1 in 10 of coefficients relative to observations. Here we have a 7 in 12 ratio.

We noted in many references, too numerous to cite, the increasing use of LISREL and other elaborations on path models fitted to data from extended families or pedigrees. We agree with the following cautions offered by Khoury, *et al* (1993): "The main disadvantage in path analysis is its lack of a clearly specified mathematical form for the wide scope of possible path models." This

quote appears in their section on "Path Analysis" which is preceeded by their section on "Variance Components Models."

The following quote [from a paper by Lange, *et al* (1976) that also mentions Fisher scoring, Newton-Raphson and cites Jenrich & Sampson)] has a ring of realism about it: "Adding layer upon layer of complexity to explain phenotypically plastic behavior traits like intelligence is apt to be futile, both statistically and genetically." On the other hand, twins (particularly in a human culture that has norms designed to rear ordinary people in all cases) offer a unique opportunity to separate the genetic from the environmental influences.

Since so much of the above discussion has dealt with issues of statistical methods we would like now to recapitulate some of the substantive findings. In the first part we noted the effects of birth order, and in Cattell (1982) we came across the statement "among twins the first-born twin tends to be larger than the second born" so our finding is apparently well known. Perhaps the birth order effects on the mental scores are not so well known, and perhaps they are not replicated.

The re-fittings in Table 7 still show (as did Table 3) that the family effects correlation (ρ_F) is higher for S and W than it is for L and B. We still believe nutrition may be responsible for this. The very high levels of ρ_F for the mental scores are certainly due to cultural effects, but may also be increased by nutritional patterns as noted by Lynn (1989).

The empirical finding with the greatest "popular scientific" impact is the level of heritability (the H′ quantity) for mental tests, which ("fit" or "re-fit") ranges in our study from 20% to 40%. In the interpretation of this

quantity we have come to realize that the level of heritability is a result of a particular culture. It depends largely on how children are reared.

That there are neurological and otherwise physical differences relevant to IQ that are governed by genes is not denied. These features are, however, fairly fixed and given, while cultural influences vary both systematically and widely. For example, when data show, as they seem to, higher heritabilities of IQ scores in England than in the U.S. (Jencks, 1972) such differences could arise completely from the cultural and societal variations, even if the genetic causes were essentially the same.

References

Cattell, R. B. (1982). *The Inheritance of Personality and Ability: Research Methods and Findings*. Academic Press, NY.

Erlenmeyer-Kimling, L. and Jarvik, L. F. (1963). "Genetics and intelligence: A review," *Science*, 142:1477-1478.

Fisher, R. A. (1958). *Statistical Methods for Research Workers*. Hafner Pub. Co., Inc., NY.

Haseman, J. K. and Elston, R. C. (1970). "The estimation of genetic variance from twin data," *Behaviour Genetics*, 1:11-14.

Hedges, L. V. and Olkin, Ingram (1985). *Statistical Methods for Meta-analysis*. Academic Press, NY.

Herrnstein, R. J. and Murray, Charles (1994). *The Bell Curve: Intelligence and Class Structure in American Life*. The Free Press, NY.

Jencks, Christopher (1972). *Inequality: A Reassessment of the Effort of Family and Schooling in America*. Basic Books, Inc., NY.

Khoury, M. J., Beaty, T. H., and Cohen, B. H. (1993). *Fundamental of Genetic Epidemiology*. Oxford University Press, NY.

Lange, K., Westlake, J., and Spence, M. A. (1976). "Extensions to pedigree analysis, III. Variance components by the scoring method," *Annals of Human Genetics*, 39:485-419.

Loehlin, J. C. and Vandenberg, S. G. (1968). "Genetic and enviromental components in the covariation of cognitive abilities: An additive model," Chapter 15 in Vandenberg, S. G. (ed.) *Progress in Human Genetics*, Johns Hopkins Press, Baltimore.

Lynn, Richard (1989). "A nutrition theory of the secular increases in intelligence: Positive correlations between height, head lice and IQ," *British Journal of Educational Psychology*, **59**:372-377.

Wright, Sewall (1978). "The application of path anlaysis to etiology," in *Genetic Epidemiology*, eds. Morton, N. E. and Chung, C. S. Academic Press, NY.