Behavior Genetics, Vol. 24, No. 1, 1994

A Model System for Analysis of Family Resemblance in Extended Kinships of Twins

K. R. Truett,[1] L. J. Eaves,[1,4,5] E. E. Walters,[4] A. C. Heath,[2] J. K. Hewitt,[1] J. M. Meyer,[1] J. Silberg,[1] M. C. Neale,[1], N. G. Martin,[3] and K. S. Kendler[1,4]

Received 1 May 1992—Final 28 May 1993

The "Virginia 30,000" comprise 29,698 subjects from the extended kinships of 5670 twin pairs. Over 80 unique correlations between relatives can be derived from these kinships, comprised of monozygotic (MZ) and dizygotic (DZ) twins and their spouses, parents, siblings, and children. This paper describes the first application of a fairly general model for family resemblance to data from the Virginia 30,000. The model assesses the contributions of additive and dominant genetic effects in the presence of vertical cultural inheritance, phenotypic assortative mating, shared twin and sibling environments, and within-family environment. The genetic and environmental effects can be dependent on sex. Assortment and cultural inheritance may be based either on the phenotype as measured or on a latent trait of which the measured phenotype is an unreliable index. The model was applied to church attendance data from this study. The results show that the contributions of genes, vertical cultural inheritance, and genotype–environment covariance are all important, but their contributions are significantly heterogeneous over sexes. Phenotypic assortative mating has a major impact on family resemblance in church attendance.

KEY WORDS: Twins; twin kinships; cultural inheritance; assortative mating; religion; maternal effects; twin environment.

INTRODUCTION

If the history of human behavioral genetics before 1970 can be characterized as the attempt to confirm genetic responsibility for variation in behavior, the history of the succeeding two decades can be characterized by a growing desire to detect and analyze the effects of family environment. Prior to 1970 there *had* been significant efforts to complement the mathematical clarity of Fisher's (1918) model for polygenic inheritance with models for the environment, notably Cattell's Multiple Abstract Variance Analysis (MAVA; Cattell, 1960) and Wright's (1921) path analytic formulation of nongenetic inheritance. However, such models either involved unresolved conceptual and mathematical inconsistencies or simply did not attract the attention they deserved.

Theoretical work in the 1970s and 1980s, initially by Cavalli-Sforza and Feldman (1973a, b, 1981) began to develop mathematical treatments for cultural inheritance which rivaled the traditional genetic model of Fisher (1918) and his intellectual descendants (Mather, 1949; Burt and Howard, 1965; Jinks and Fulker, 1970; Mather and Jinks, 1982; Eaves *et al.*, 1975). By generalizing the basic algebra of Mendelian population genetics, Cavalli-Sforza and Feldman extended single-major locus models into mathematical treatments of cultural transmission. Although this contributed important

[1] Department of Human Genetics, Medical College of Virginia, P.O. Box 3, M.C.V. Station, Richmond, Virginia 23298-0003.
[2] Department of Psychiatry, Washington University Medical School, St. Louis, Missouri 63110.
[3] Queensland Institute for Medical Research, Herston, Brisbane, Queensland, Australia, 4006.
[4] Department of Psychiatry, Virginia Commonwealth University, Richmond, Virginia 23298.
[5] To whom correspondence should be addressed.

0001-8244/94/0100-0035$07.00/0

qualitative insights concerning evolution and family resemblance, it was not applicable to the continuous traits that behavior geneticists were attempting to study.

Eaves (1976a, b) showed that the biometrical genetic models of Mather (1949) which had been used by Jinks and Fulker (1970) in their critical evaluation of MAVA could be extended to allow for vertical cultural inheritance (i.e., from parent to offspring) and horizontal nongenetic transmission between siblings. The essence of these models lay in their recognition that there might be a "genetic environment" (Darlington, 1971) so that some genes might be expressed primarily through their impact on the environment of other members of the population. The primary weakness of these models was their restriction to randomly mating populations.

Meanwhile, Morton (1974) and others (Rice *et al.*, 1978; Cloninger *et al.*, 1979) used Wright's (1921) work on path analysis to model genetic and environmental influences within families. These studies laid the foundation for current studies of the inheritance of behavior in humans.

In addition to theoretical advancements, advances in computer technology have greatly simplified genetic studies of behavior. Efficient numerical optimization methods, such as those developed by the Numerical Algorithms Group (NAG, 1991), allow rapid exploration and modification of hypotheses. This is vital because of the increasing complexity of modeling and because data sets have reached proportions unmanageable without computers.

Although there is still no "panacea" for complete resolution of individual differences in humans, large studies of twins and their relatives have become increasingly attractive in contrast to other designs, implemented with smaller samples, which lack the power to test more subtle hypotheses (Heath and Eaves, 1985; Heath *et al.*, 1985).

The "Virginia 30,000" study is perhaps the first attempt to exploit all the collateral two-generational relationships identified in the study of the kinships of twins to estimate the sex-dependent contributions of genes and environment to complex traits in the presence of assortative mating. This paper describes the sample and illustrates the application of one model for biological and cultural effects on a variable for which there are good reasons, a priori, to suspect major genetic effects, namely, church attendance.

ASCERTAINMENT AND STRUCTURE OF THE VIRGINIA 30,000 SAMPLE

Figure 1 shows the idealized pedigree around which the "Virginia 30,000" was designed. The Virginia 30,000 contains data from 14,763 twins, ascertained from two sources. Public birth records in the Commonwealth of Virginia were matched with other public records, such as those held by the Division of Motor Vehicles, to obtain current addresses for twins born in Virginia between 1915 and 1971. Questionnaires were mailed to twins who had returned at least one questionnaire in previous surveys, with complete returns from 5287 families. The remainder of the twins ($N = 9476$ individuals) responded to a letter published in the newletter of the American Association of Retired Persons (AARP). The first group of twins is referred to as the "Virginia cohort" and the second as the "AARP cohort." The twins were not selected for any specific outcome apart from willingness to complete an extensive mailed questionnaire survey dealing with health and lifestyles. The sample is almost exclusively Caucasian (99.8%) because funding was originally available to study Caucasians.

All pedigrees included one of the following: a male or female monozygotic (MZ), or a male, female, or unlike-sex dizygotic (DZ) twin pair, and, ideally, all available parents, siblings, spouses, and children of the twins. This provides a rich combination of 80 sex-specific two-generation relationships (parent/offspring, spousal, twin, cousins, etc.).

After a pilot mailing of the questionnaire, twins were mailed a 16-page "Health and Lifestyles" questionnaire which, in addition to demograhic in-

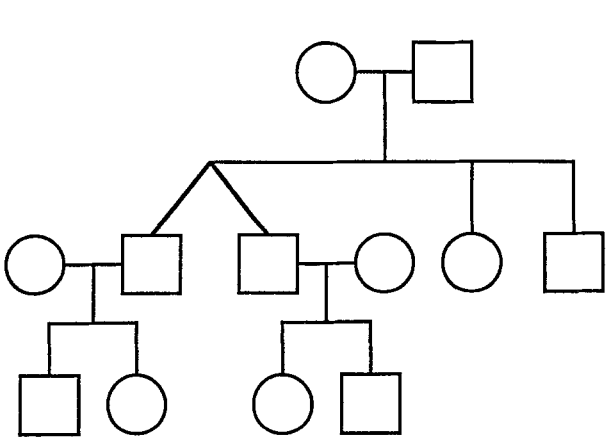

Fig. 1. Idealized pedigree for the extended twin family study.

formation, inquired about health, alcohol and tobacco consumption, passive smoking, life-events, personality, social support, social attidues, psychiatric symptoms, disease history, and family history. The survey also included questions for determination of zygosity which have been shown to be 95% accurate (Eaves *et al.*, 1989) for determination of zygosity. The twins were asked to supply the names and addresses of their spouses, siblings, parents, and children for the follow-up study of relatives of twins. Second and third mailings were sent to nonrespondents who had not specifically indicated that they did not wish to participate. A telephone follow-up was conducted in an attempt to complete pairs where only one twin had responded. Completed questionnaires were received from 69.8% of twins invited to participate.

The original twin questionnaire was modified slightly to provide two additional forms, one appropriate for the parents of twins and another for the spouses, children, and siblings of twins. The modifications affected those aspects which specifically related to twinning. The aim was to obtain self-report data on all subjects for the outcome variables. These questionnaires were sent to the relatives whose twins gave permission for this and mailed follow-up was conducted as before. The response rate (44.7%) was much lower from the relatives than from the twins.

The demographics of the sample are as follows: 59.7% female; 50% under 50 years in age; 74% married or living with someone and 13% widowed, separated, or divorced; 12% lacking a high-school diploma and 33% college graduates; and 65.8% Protestant, 15.5% Catholic, 3.9% Jewish, and 10.3% another (unspecified) religion. This may be compared to the U.S. population, where whites are 51.2% female and 90% Catholic or Protestant, 3% Jewish, and 7% "other" or "no" religion. The median family income in this study is $34,000, compared to the estimated $30,260 average family income for Caucasians in 1985 (New York Public Library, 1989).

RESOLVING BIOLOGICAL AND CULTURAL EFFECTS ON CHURCH ATTENDANCE

The self-report data on church attendance in the "VA 30,000" are used to illustrate the potential value of the data set. The data were obtained using a question which asked the respondents to indicate the number corresponding to the frequency at which they attend church services. The responses were scored on a 6-point scale: don't know = 1 (these were treated as missing values); more than once a week = 2; once a week = 3; once or twice a month = 4; a few times a year = 5; rarely = 6; and never = 7. The question also asks twins to report on their cotwins, parents, and spouse relatives. The item included on surveys that were sent to nontwins was similar except that parents were asked to report on both twins, and other relatives were asked to report only on themselves, their spouses, and their parents.

Because twins scored their relatives as well as themselves for frequency of church attendance, the data can be examined for a response bias by comparing differences in the twins' reporting of the frequency of church attendance for nonresponders versus responders. The frequencies from this analysis are presented in Table I. Relatives reported as deceased were omitted from this analysis. A slight bias was noted for husband and cotwin responders to be more frequent church attenders but this bias was not present in parents or wives. Although the differences in frequency distribution are statistically significant, the differences in mean rating are small. For twins with relatives who did respond, the correlation between the twins' estimation of their relatives' church attending habits and the relatives' self-report was .86, .87, .87 and 0.88 for mother, father, spouse, and cotwin, respectively. Table I also presents self-report church attendance frequencies for the VA 30,000 by sex.

CALCULATING CORRELATIONS BETWEEN RELATIVES

The model presented in this paper (see below) utilizes the self-report data from all respondents in the study. Biological inheritance and cultural inheritance yield algebraically distinct contributions to some *80* unique collateral and two generational correlations from these kinships when sex differences are permitted in genetic and environmental effects. Computing the summary correlations between relatives for these extended kinships of twins, therefore, becomes a formidable task. Because of the difficulty of obtaining self-reports from adults in three generations within a single family, most families involve only two generations—either twins and their parents or twins and their children. Fewer than 100 (of 8644) families have responses from at

Table I. Twin's Reports on Relatives by Response Status of Relative and Frequency Distribution (%) of Self-Report Church Attendance

Relation to twin	Status	2+ times/wk	1 time/wk	1–2 times/wk	Few times/yr	Rarely	Never	Total
Mother	Respondents	16.0	36.3	11.8	11.8	15.6	8.5	7,865
	Nonrespondents	15.6	36.0	10.9	13.5	15.8	8.2	2,186
Father	Respondents	12.2	31.9	10.5	11.5	20.0	13.9	7,895
	Nonrespondents	11.7	29.0	8.4	13.0	21.3	16.7	1,415
Wife	Respondents	12.6	35.0	13.8	16.1	16.7	5.9	1,416
	Nonrespondents	16.4	29.3	12.1	16.7	17.5	8.1	1,473
Husband	Respondents	9.1	34.1	9.4	14.0	21.6	11.7	2,668
	Nonrespondents	11.1	23.1	8.8	14.6	26.6	15.7	2,058
Cotwin[a]	Respondents	16.4	31.8	10.9	14.6	18.2	8.1	2,123
	Nonrespondents	12.8	26.6	10.2	16.8	22.8	10.7	10,735
Self report								
All females		16.2	31.3	11.0	17.0	17.4	7.1	17,349
All males		13.1	25.2	11.0	18.5	23.0	9.2	11,714

[a] Cotwins cannot be divided by gender since the gender of nonresponding twins is unknown.

least one parent and one offspring of the twins. Thus, the effort of specifying the expected correlations for these relationships greatly exceeds their contribution to the parameter estimates, so grandparental correlations are omitted from the analysis. Data from half-siblings and step- and adoptive relationships are omitted for identical reasons.

The required correlations are computed in three stages using a SAS program. First, the entire data set is corrected for the linear and quadratic effects of age, sex, twin status, and source of ascertainment (Virginia vs. AARP) and interactions between these effects. Subsequent analyses are based on the residuals from this regression analysis. Optimally, the regression model (and the residual phenotypic variances) would be simultaneously estimated in the genetic analysis. We chose not to do this because of the computational burden, but experience suggests that the results differ little under the two approaches. All family members are defined by their relationship to the twins within the family and Pearson product–moment correlation coefficients calculated for every possible pair. Some of the expected correlations should be algebraically identical under the most simple model of genetic and cultural inheritance (i.e., the father–son relationship (correlation) should not differ between a male twin/father pair and a male twin/son pair). An efficient algorithm for maximum-likelihood estimation of the correlations has not yet been formulated for this complex family structure and large sample sizes. Therefore, the correlations are pooled into groups (retaining sex specificity) as defined by the pair's familial relationship using the procedure outlined by Snedecor and Cochran (1980). For example, the multiple estimates of the father–son correlation (fathers of twins with first or second twin, father of twins with male siblings of twins, male twins with their male children, and husbands of female twins with their male children) can now be pooled into a single father–son estimate. For sibling and cousin correlations, several possible pairings exist within a family. For these groups, therefore, each possible pairing is used in calculating the correlation. This means that the statistics are not independent and their precision may be somewhat overestimated; however, they are unbiased (McGue *et al.*, 1984).

Regression analysis produced an overall multiple r^2 of .0545. The residuals were correlated across relative pairs and then pooled to obtain 80 correlations. These and the number of pairs contributing to each correlation for self-report church attendance are presented in Table II. The amount of information about the correlations will be strictly correct only if the raw observations are normal. In our case, with categorical data, they will be an approximation at best.

A preliminary examination of the correlations provides some expectations concerning the struc-

Table II. Church Attendance: Correlations for Biological Relationships

	Male–male		Female–female		Male–female		Female–male	
	r	n^a	*r*	n^a	*r*	n^a	*r*	n^a
Nuclear families								
Siblings	.312	1538	.331	3630	.314	4370	—	
DZ twins	.387	575	.402	1165	.296	1310	—	
MZ twins	.502	772	.600	1863	—		—	
Parent–child (parent's sex 1st)	.342	2160	.401	4541	.361	3004	.372	3052
Avuncular via[b]								
Father's MZ cotwin	.262	221	—		.223	341	—	
Mother's MZ cotwin	—		.263	1046	—		.234	674
Father's DZ cotwin	−.056	107	.178	192	.227	147	.109	116
Mother's DZ cotwin	.130	161	.266	520	.226	202	.204	334
Father's sibling	.222	96	.173	198	.112	155	.211	137
Mother's sibling	.257	234	.129	548	.158	294	.146	397
Cousins via								
Opposite-sex DZ twins[c]	.078	37	−.053	67	−.058	51	−.110	72
MZ father	.252	37	.178	95	.230	100	—	
MZ mother	.200	155	.252	347	.170	449	—	
DZ father	−.319	18	.108	40	.004	50	—	
DZ mother	−.084	53	.124	145	.064	160	—	
Spouses	—		—		.707	4837	—	
Spouse of twin with[d]								
MZ cotwin	—		—		.424	1148	.383	614
DZ cotwin	.126	357	.243	448	.370	608	.321	417
Sibling of twin	.179	431	.266	476	.258	752	.192	363
Parent of twin	.248	194	.309	303	.281	349	.168	215
Spouse of MZ cotwin	.375	300	.294	182	—		—	
Spouse of DZ cotwin	.314	122	.454	105	.137	163	—	
Affine avuncular via[e]								
Father's MZ cotwin	—		.257	221	—		.287	130
Mother's MZ cotwin	.208	347	—		.233	508	—	
Father's DZ cotwin	.130	34	.281	82	.248	61	.090	59
Mother's DZ cotwin	.249	123	.099	97	.331	171	.020	72

[a] Number of pairs.
[b] Aunt/uncle's sex listed first; niece/nephew's sex listed second.
[c] First sex listed is sex of male twin's child.
[d] First sex listed is spouse's sex.
[e] Aunt/uncle's sex listed first; niece/nephew's sex listed second.

ture of the model necessary to account for the correlations. The large spousal resemblance indicates that assortative mating or marital interaction will be required. The correlations between siblings and parents and offspring are quite large compared with correlations in the personality domain (Eaves *et al.*, 1989) and comparable with many in the cognitive and physiological domains (Fuller and Thompson, 1978), indicating substantial involvement of genetic or environmental factors in family resemblance. Virtually all the correlations involving the MZ relationship are greater than the corresponding correlations involving DZ twins. Not only do the MZ twin correlations exceed the DZ correlations, but the correlations between cousins related through MZ twins exceed those related through DZ twins, etc. This indicates that genes may play a substantial role in behaviors such as church attendance. Correlations between female pairs generally exceed correlations between males, indicating that genetic and environmental effects may depend on the sex of the individual. The resemblance of like-sex DZ twins is slightly greater than that for like-sex siblings, which may indicate a small special twin environment or an interaction between genotype and age or parity. These considerations are formalized

in a model for family resemblance in the Virginia 30,000.

METHODS: A MODEL FOR FAMILY RESEMBLANCE

The model is presented in three stages:

(1) the model without sex differences, assuming that mate selection and cultural inheritance are both based on the measured phenotype;
(2) the model without sex differences, including a latent variable for which assortment and nongenetic inheritance occur; and
(3) the model allowing for sex differences in all genetic and environmental factors and containing a latent variable for assortment and cultural inheritance.

The Basic Model Without Sex Differences (Fig. 2)

Figure 2 summarizes the elements of this model. The lowercase letters denote the path coefficients and correlations as follows:

(1) additive genetic effects, h;
(2) genetic dominance, d;
(3) path from environment to phenotype, e;
(4) path from parental phenotype to offspring environment, w;
(5) phenotypic correlation between mates, μ;
(6) path from residual sibling shared environment to phenotype, s;
(7) path from additional twin shared environment to phenotype, t; and
(8) correlation between genotype and environment, ρ.

The model assumes *primary phenotypic* assortative mating (No. 5 above) and "P-to-E" vertical cultural inheritance (No. 4 above). The *genotype–environment correlation*, ρ, occurs when the parental phenotype, which can contribute to the offspring's environment through parent–offspring transmission, is partly genetic in origin. This results in a correlation between the offspring's environment and genes. The process of transmission and assortment is assumed to be in equilibrium, and thus, ρ is constant between generations. That is, ρ_{parent} is constrained to be equal to ρ_{child}. Since models are fitted to correlations, the scale of measurement has unit variance; therefore, we impose the further constraint that the sum of all sources of variance for individuals equals one.

Assortment and Cultural Transmission Based on a Latent Variable

The measured trait may not correlate perfectly with the trait for which mate selection and cultural transmission are actually occurring. Morton (1974) argued for a model of "social homogamy" in which assortment and cultural transmission are based on a correlated latent variable to which genes make no contribution. Another mechanism of assortment (proposed by Heath and Eaves, 1985) presents a model for mixed homogamy in which mate selection is based on both the social background of the spouses and the phenotype of the mate. We have used "phenotypic assortment plus error" (Heath, 1983), in which the actual measurement is considered a more or less unreliable index of the latent score on which assortment is based. In this model, all expected correlations were multiplied by the square of the path from "true" (or latent) score to "observed" score [the reliability (rel)]. When there is significant assortative mating or cultural inheritance, there is sufficient information to estimate the reliability without repeated measurement (Heath, 1983).

Allowing for Sex Differences in Model Parameters (Fig. 3)

It is possible that the genetic and environmental factors causing variation differ between the sexes. For the simple case of randomly mating populations, a model for sex differences in gene action was specified by Eaves (1977), which allowed for the same genes to have different magnitudes of effects on males and females. This model allows for estimation of separate genetic variances for males and females and a correlation between gene effects in males and females. The genetic correlation between the sexes will be unity if the effects of all autosomal loci on one sex are constant multiples of their effects on the other sex. In this case, we speak of "scalar sex limitation of the gene effects." Analogous definitions may be given for the "sex-limited" effects of the shared environment.

If the magnitudes of the loci or, by analogy,

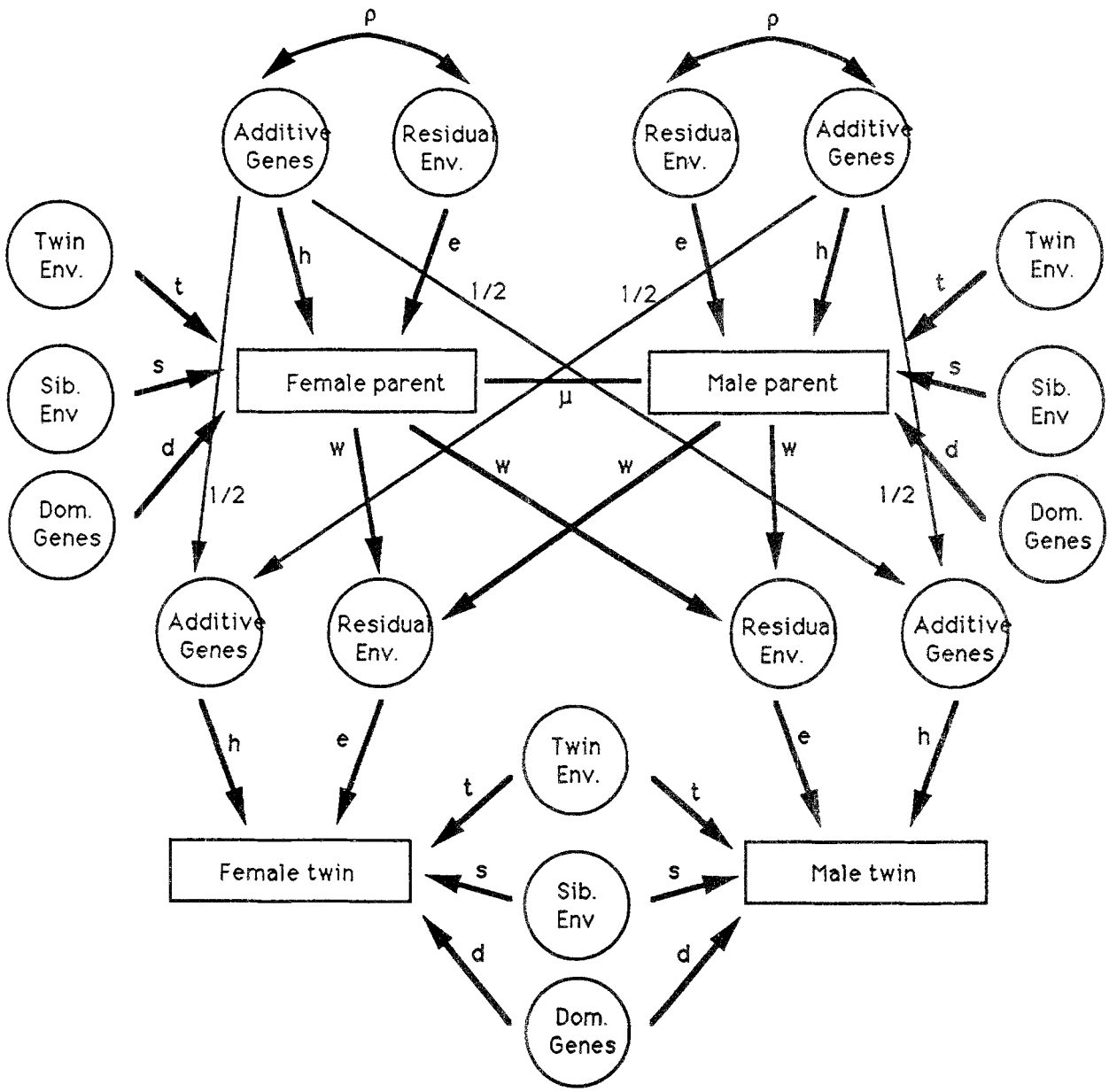

Fig. 2. The extended family model without sex differences for opposite-sex DZ, twins and their parents. Additive genes and environment are correlated in the children as shown for the parents.

"environmental effects" on one sex are not constant multiples of their effects on the other sex, then we speak of nonscalar sex limitation of genetic (or environmental) effects.

In the present model, we employ the following notation for the effects of dominance, sibling environment, and special twin environment: d_m, s_m, and t_m, respectively, for males; d_f, s_f, and t_f for females; and r_d, r_s, and r_t for the correlations across sexes of the dominant, sibling environmental, and twin environmental effects. Our model for the twin environment assumes that the same basic environmental factors influence twins and nontwins. Hence, the parameter t^2 contributes to the total variance of all individuals in the study. However, only in twins are these environmental effects correlated. We note that genotype × age interaction

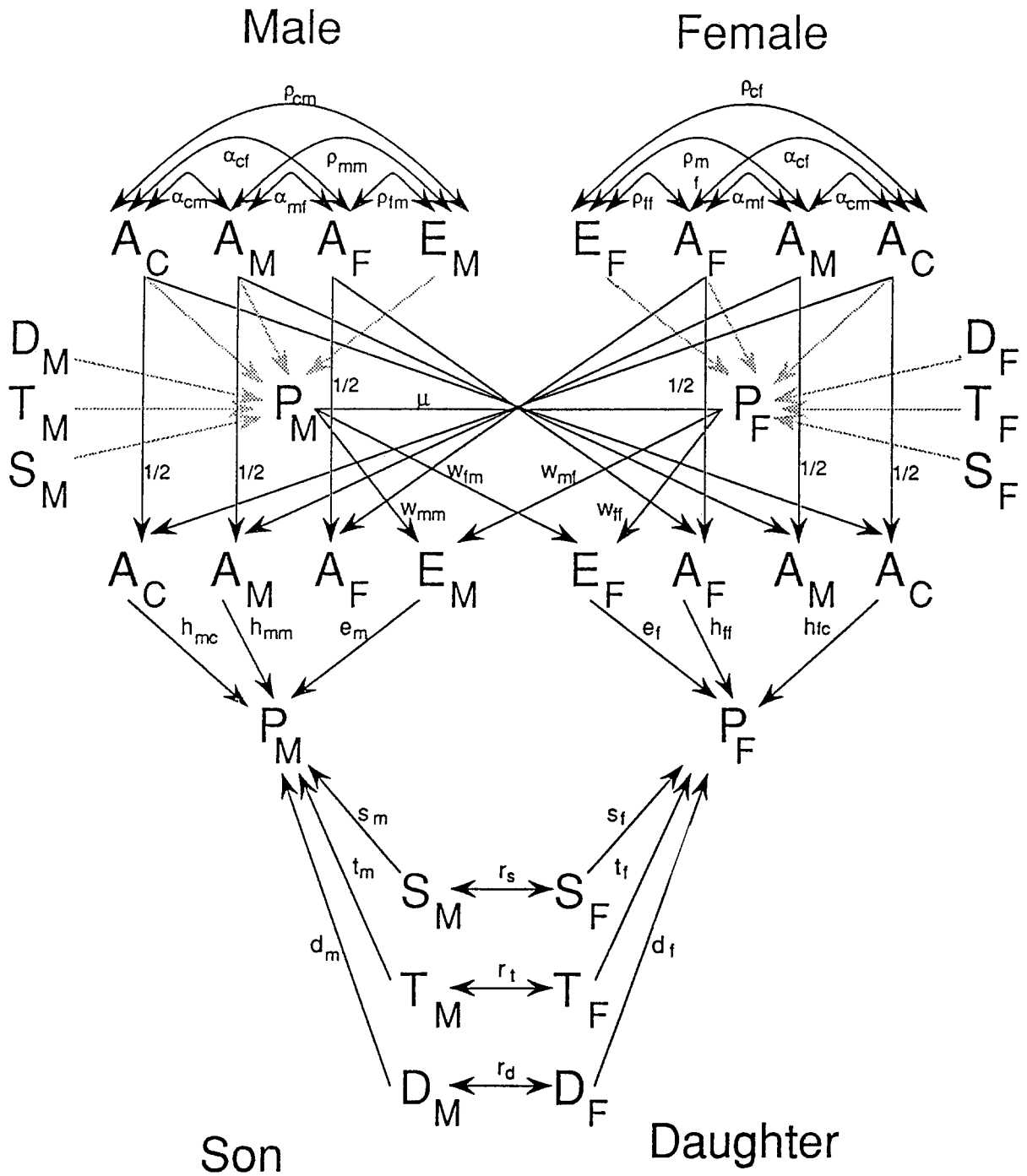

Fig. 3. The full sex-dependent extended family model for opposite-sex DZ twins and their parents. Path coefficients are the same in both generations. Gene–gene and gene–environment correlations occur in both generations. Space limitations made labeling all paths impossible. The parameters are explained in the text.

effects may create the semblance of a shared twin environment in data of this type (Eaves, cited by Eysenck, 1980).

Since vertical cultural transmission is assumed under this model to be based on the parental phenotype for the trait under investigation,

the question of "non scalar" vertical cultural transmission does not apply. However, the cultural impact of mothers may differ from that of fathers and may further depend on the sex of the offspring. In the model for sex differences, therefore, we require *four* cultural parameters: w_{mm}, w_{mf}, w_{fm}, and w_{ff}. The first subscript denotes the sex (m = male) of the offspring, and the second denotes the sex of the parent.

Specification of sex-limited additive genetic effects is more difficult when there is assortative mating, which induces correlations between loci that would otherwise be independent (see, e.g., Fisher, 1918). We have adopted one of several, formally equivalent, ways of parameterizing the additive sex-limited effects. Recognizing that the additive genetic variances in the two sexes and the genetic covariance between them require three free parameters for their complete specification, we assume that one set of genes explains *all* the genetic variance in females and the genetic covariance between the sexes. The paths from this "common" set of genes to the male and female phenotypes are denoted h_{mc} and h_{fc}, respectively. A second set of genes has effects which are specific only to males, and the path from these genes to the male phenotype is specified by h_{mm}. Although the "male-specific" genes are not expressed in females, they are still present in females and correlated, through phenotypic assortment, with the "common genes." We denote the induced correlation between the two sets of additive genetic effects α_{cm}.

The joint effects of assortment and vertical cultural transmission induce *four* genotype–environment correlations: two between the "common" additive genetic effects and the environments of males and females, ρ_{cm} and ρ_{cf}, respectively; and two between the "male-specific" additive genetic effects and the environments of males and females, ρ_{mm} and ρ_{mf}, respectively. These genotype–environment correlations are estimated as constrained parameters when fitting the model (i.e., they are functions of other parameters). Separate parameters are required to specify the path from male environment to phenotype (e_m) and female environment to phenotype (e_f). Under the simple model for "phenotypic assortment with error," the paths from true score to observed score, rel_m and rel_f, may differ between males and females.

Since the total phenotypic variance is standardized to unity in both sexes, two further constraints are required to enforce these conditions. Thus, seven constraints are imposed on parameter values under the full model. The full model for sex-limited effects is given for pairs of unlike-sex DZ twins in Fig. 3.

FITTING MODELS TO THE OBSERVED CORRELATIONS

The rules of path analysis can be used to derive algebraic expectations for the 80 correlations between relatives under the above model. An annotated FORTRAN subroutine coding these expectations is available from the authors.

Initially, two classes of models can be fit to the observed correlations for the measured variable. The first set constrains the genetic and environmental effects to be identical in males and females and is referred to as "sex-independent models." The "full" form of this model (Model 1) is specified by constraining all parameters to be identical across the sexes and by fixing the cross-sex correlations to unity: $h_{mm} = 0$, $h_{mc} = h_{fc}$, $e_m = e_f$, $w_{mm} = w_{mf} = w_{fm} = w_{ff}$, $d_m = d_f$, $s_m = s_f$, $t_m = t_f$, $\rho_{cm} = \rho_{cf}$, $\alpha_{cm} = 0$, $\rho_{mm} = \rho_{mf} = 0$, and $r_d = r_s = r_t = 1$. Eight submodels, each of which tests the significance of a particular parameter or group of parameters by comparison of goodness-of-fit statistics to the full model, can be fit to the data. The second principal class of models allows for sex differences between parameters of the model. The full model shown in Fig. 3 (Model 6) allows all effects in the model to differ in magnitude between the sexes and allows for all effects in the model to be partially sex specific ("nonscalar sex limitation"). Within each principal set of models, submodels can again be fit to test specific hypotheses. The submodels, hypotheses tested, and corresponding degrees of freedom are detailed in Table III.

The models are fitted to the z-transformed observed correlations by iterative diagonal weighted least squares (WSL) using the amount of information about each transformed product–moment correlation ($N-3$) as weights. Minimization of the loss function $s^2 = \sum_i w_i(z_i - Ez_i)^2$ was performed using a FORTRAN program which

Table III. Models Tested with Goodness-of-Fit Statistics

No.	Hypothesis tested[a]	Parameters involved	χ^2	df	p	AIC
		Models without sex differences				
1	Full model		117.76	73	.00	−28.24
2	No latent phenotype	rel = 1.00	122.60	74	.00	−25.40
3	NSE: assortative mating[c]		—[c]	74	.00	—
4	NSE: common family environment	t, s, w	135.56	76	.00	−16.44
4A	NSE: twin environment	t	125.99	74	.00	−22.01
4B	NSE: sibling environment	s	124.24	74	.00	−23.76
4C	NSE: cultural transmission	w	124.45	74	.00	−23.55
5	NSE: genes	h, d	272.12	75	.00	122.12
5A	NSE: dominant genes	d	134.62	74	.00	−13.38
		Models with sex differences				
6	Full sex-dependent model[b]		65.98	61	.31	−56.02
7	No latent phenotype	rel = 1.00	74.13	63	.16	−51.87
8	NSE: assortative mating[c]		—	62	.00	—
9	NSE: common family environment	t_m, t_f, s_m, s_f, w_m, w_f	108.40	71	.00	−33.60
9a	NSE: twin environment	t_m, t_f	90.26	64	.02	−37.74
9b	NSE: sibling environment	s_m, s_f	84.79	64	.04	−43.21
9c	NSE: cultural transmission	w_m, w_f	77.43	65	.14	−52.57
10	NSE genes	h_m, h_f, d_m	132.20	67	.00	−1.80
10a	NSE: dominant genes	d_f, d_m, d_f	71.13	64	.25	−56.87
11a	NSD: latent trait	$rel_m = rel_f$	67.04	62	.31	−56.96
11b	NSD: cultural inheritance	w_{xx}	76.36	64	.14	−51.64
11ca	NSD: twin environment	$r_t = 1$	69.92	62	.23	−54.08
11cb	NSD: effects of twin environment	$t_m = t_f$	81.73	63	.06	−44.27
11da	NSD: sibling environment	$r_s = 1$	65.98	62	.34	−58.02
11db	NSD: effects of sibling environment	$s_m = s_f$	66.34	63	.36	−59.66
11ea	NSD: dominant genetic effects	$r_d = 1$	65.98	62	.34	−58.02
11eb	NSD: effects of dominant genes	$d_m = d_f$	67.61	63	.32	−58.39
11f	NSD: residual environment	$e_m = e_f$	69.23	63	.28	−56.77
12	NSD: additive genetic effects	$h_{mm} = 0$	66.08	62	.34	−57.92
13	Proposed model (see Table V for details)		75.89	72	.35	−68.11

[a] NSE, effects of parameter listed are not significant; NSD, effects of parameter stated are not sex dependent.
[b] Seven parameters (e_m, e_f, a_{cm}, a_{cf}, r_{mm}, r_{cf}, r_{cm}) can be expressed as functions of the other parameters of the model. Hence, there are only 19 free parameters in the full model, yielding 80 − 19 = 61 df for the goodness-of-fit chi-square.
[c] Model did not converge to stable solution.

used the NAG routine E04UCF for constrained nonlinear optimization. A copy of the code (without the associated copyright routines) may be obtained from L.J.E.

Since the pairs of relatives are not independent in our sample, the weight matrix is not strictly diagonal and the resulting loss function may be somewhat greater than the chi-square which would result from application of strict maximum likelihood. While theoretically more efficient, maximum likelihood adds an extreme computational burden when calculated for each of the 8644 pedigrees in this large sample. However, simulation studies have been conducted (McGue *et al.*, 1984) which suggest that although inferences assuming diagonal WLS based on the chi-square test goodness of fit are overpowered, they are not seriously misleading.

Test of nested hypotheses were conducted by treating the difference between the loss function under the more general and restricted models as a chi-square with degrees of freedom equal to the difference in the number of free parameters under the two models. Choice between alternative models is a decision process subject to error, however, we seek the best compromise between parsimony and fit. Akaike's information criterion (AIC; Akaike, 1970) (obtained by subtracting

twice the residual degrees of freedom from the goodness-of-fit chi-square) is one quantitative index reflecting these two criteria. Alternatively we may select the simplest model which fits as well as the full model by likelihood-ratio criteria. In our case, both criteria lead to the same conclusion.

RESULTS: APPLYING THE MODEL TO THESE DATA

Because the full model without sex differences fit significantly worse than the full model with sex differences ($\chi^2_{73} = 117.67, p = .00$, vs. $\chi^2_{61} = 65.98, p = .31$), sex-dependent parameters must be included to explain the patterns seen in our correlations. Therfore, only results from sex-dependent model testing are presented. The models tested and their corresponding goodness-of-fit statistics are presented in Table III. A summary of the results from sex-dependent model testing follows.

Reliabilities Cannot Be Jointly Set to Unity ($\chi^2_2 = 8.15, p < .01$; Model 7)

This indicates that church attendance is actually a measure of a latent variable on which assortative mating is based and for which cultural transmission is occurring. The correlation between the latent variable and the measured phenotype (church attendance) is .94 in males and .97 in females. These terms can be squared to provide a measure of the reliability, with the results of .88 in males and .94 in females. Reliability can also be estimated by calculating the correlation between a twin's self-report and the cotwin's estimation of his/her frequency of church attendance. For male twins ($N = 2273$) this correlation was .88, and for female twins ($N = 3441$) the correlation was .90. The similarity of these two calculations implies that both methods of calculating reliability are fairly accurate. Constraining the male and female reliabilities to be equal (Model 11a) did not significantly worsen the fit of model, indicating that correlation between the measured phenotype and the latent variable for which assortment occurs is not significantly different in males and females.

All Sources of Family Environment Cannot Be Jointly Eliminated from the Model ($\chi^2_{10} = 42.42, p < .01$; Model 9)

Individual tests indicate that environments shared by twins (Model 9a) and by siblings (Model 9b) cannot be removed from the data ($\chi^2_3 = 24.28$ and 18.81; $p < .01$ for both). However, the correlation between sibling environment in males and sibling environment in females approached unity, and the path coefficients between the environments and the phenotype were nearly equivalent between the sexes, indicating sex-independent effects, an observation confirmed in Model 11db ($\chi^2_2 = .36, p = .84$). Models 11ca and 11cb confirm that the twin environments are sex specific and uncorrelated. Removal of all parental effects on the child's environment (Model 9c) significantly worsens the model's fit ($\chi^2_4 = 11.45, p = .00$); however, individual tests of significance indicate that only maternal effects, the magnitude of which are four times the magnitude of paternal effects in the full model, should be retained. This would mean that mothers influence their child's church attendance (independent of the child's sex), while fathers do not.

All Sources of Genetic Variation Cannot Be Jointly Removed from the Model ($\chi^2_6 = 66.22, p < .01$; Model 10) Without Worsening the Fit

The genetic components are divided into additive and dominant genetic effects; however, the latter do not contribute significantly to the model ($\chi^2_3 = 5.15, p = .16$; Model 10a). Goodness-of-fit parameters indicate that, for church attendance, sex-specific additive genetic effects are not significant. The genetic effects on church attendance can be completely attributed to common genes which have a greater impact on females.

Taking all these and other results into account, we proposed a reduced form of the "full" model which omitted sex-specific additive genetic effects, dominance, and the effects of the paternal environment. The effects of the shared sibling environment were assumed to be perfectly correlated across sexes ($r_s = 1$) and the twin environments appeared to be entirely sex specific ($r_t = 0$). The parameter estimates of fitting this model (and the full model) are presented in Table IV. The AIC under this model is -68.11, com-

Table IV. Full and Reduced Model Parameter Estimates for "Frequency of Church Attendance" Data

Genetic parameter			Environmental parameters			Other parameters		
	Full	Reduced		Full	Reduced		Full	Reduced
h_{tc}	.437	.579	t_m	.323	.365	μ	.782	.756
h_{mc}	.426	.511	t_f	.329	.373	r_{cf}	.171	.092
h_{mm}	.159	—	r_t	−.287	.000[a]	r_{cm}	.137	—[b]
a_{cm}	.053	—	s_m	.295	.247	r_{mf}	.038	—
d_m	.196	—	s_f	.247	—[b]	r_{mm}	.030	—
d_f	.356	—	r_s	1.000	1.000[a]	rel_m	.935	.968
r_d	1.000	—	w_{ff}	.284	.172	rel_f	.966	—[b]
			w_{mf}	.235	—[b]			
			w_{fm}	.072	—			
			w_{mm}	.051	—			
			e_f	.645	.630			
			e_m	.685	.692			

[a] Parameter is equated to the parameter directly above it (such as equating the effects of twin environment across sexes).
[b] Parameter is fixed to the value shown.

pared with a value of −56.02 under the full model, suggesting a gain in information. However, the decision about the form of the reduced model was based on post hoc inspection of the model-fitting results and so should be accorded little more than exploratory status.

In Table IV, we present the proportions of the total *true score* variance in each sex attributed to the principal sources of variation under the full sex-limited model and the proposed reduced form. This helps to clarify the overall inheritance pattern. The proportions of variance are calculated using the equations presented in Table V.

This estimation would be straightforward were no genotype–environment covariance or phenotypic assortment present in the model, because each proportion would be the path from the source to the phenotype squared. For covariates, the product of the path coefficients from the phenotype through the two sources of variation back to the phenotype equals the proportion of variance resulting from covariation. The proportion of genetic variance resulting from assortative mating is most easily obtainable as the difference between the total genetic variance and the proportion due to non-assortative (random) genetic effects. The genetic parameter values presented in Table IV assume an assortatively mating population. The proportion of random genetic effects in an assortatively mating population is $h^2\ (2-2\alpha)$, where α is the proportion of genes shared by siblings (not necessarily equal to .5 if parents did not randomly mate). The difference between this and the total additive genetic component equals the genetic variance resulting from assortment.

For the reduced model for church attendance, in males genes contribute 26.1% of the variance, family environment contributes 20.9%, and unique environment accounts for 46.5%; and in females the percentages are 33.5 (genes), 21.2 (family environment), and 38.6 (unique environment). Gene–environment covariance explains the remaining variance. There is substantial assortative mating ($\mu = .756$); however, only 7.22% (males) and 9.26% (females) of the phenotypic variance results from the consequent increase in additive genetic variance.

DISCUSSION

The detailed analysis of the sources of variation (Table VI) indicates the power of the extended twin-kinship design to test hypotheses which are beyond the scope of designs more limited in sample sizes and in their set of relationships. Our results, together with the weight of previous data (e.g., Eaves *et al.*, 1989), strongly suggest that studies which lack either the size or the structure to analyze sex differences in the causes of variation are potentially misleading.

Parameters which were not retained in the reduced model explained little of the variance in the full model. The removal of these parameters

Table V. Formulae for Computing the Components of Variance

Males

Genes $= h_{mc}^2 + h_{mm}^2 + 2h_{mc}h_{mm}a_{cm} + d_m^2$
  Additive $= h_{mc}^2 + h_{mm}^2 + 2h_{mc}h_{mm}a_{cm}$
    Random $= h_{mc}^2(2 - 2a_{cc}) + h_{mm}^2(2 - 2a_{mm})$
    Assortative $= (h_{mc}^2 + h_{mm}^2 + 2h_{mc}h_{mm}a_{cm}) - [(h_{mc}^2(2 - 2a_{cc}) + h_{mm}^2(2 - 2a_{mm})]$
  Dominant $= d_m^2$
Environment $= t_m^2 + e_m^2 + s_m^2$
  Family $= t_m^2 + s_m^2 + e_m^2(w_{mm}^2 + w_{mf}^2 + 2w_{mm}w_{mf}\mu)$
    Maternal $= e_m^2(w_{mf}^2 + w_{mm}w_{mf}\mu)$
    Paternal $= e_m^2(w_{mm}^2 + w_{mm}w_{mf}\mu)$
    Twin $= t_m^2$
    Sibling $= s_m^2$
  Residual $= e_m^2[1 - (w_{mm}^2 + w_{mf}^2 + 2w_{mm}w_{mf}\mu)]$
G–E covariance $= 2e_mh_{mc}r_{cm} + 2e_mh_{mm}r_{mm}$
  Total $= t_m^2 + e_m^2 + s_m^2 + d_{mc}^2 + h_{mc}^2 + h_{mm}^2 + 2r_{cm}e_mh_{mc} + 2h_{mc}h_{mm}a_{cm} + 2e_mh_{mm}r_{mm}$

Females

Genes $= h_{fc}^2 + h_{ff}^2 + 2h_{fc}h_{ff}a_{cf} + d_f^2$
  Additive $= h_{fc}^2 + h_{ff}^2 + 2h_{fc}h_{ff}a_{cf}$
    Random $= h_{fc}^2(2 - 2a_{cc}) + h_{ff}^2(2 - 2a_{ff})$
    Assortative $= (h_{fc}^2 + h_{ff}^2 + 2h_{fc}h_{ff}a_{cf}) - [h_{fc}^2(2 - 2a_{cc}) + h_{ff}^2(2 - 2a_{ff})]$
  Dominant $= d_f^2$
Environment $= t_f^2 + e_f^2 + s_f^2$
  Family $= t_f^2 + s_f^2 + e_f^2(w_{ff}^2 + w_{fm}^2 + 2w_{ff}w_{fm}\mu)$
    Maternal $= e_f^2(w_{ff}^2 + w_{ff}w_{fm}\mu)$
    Paternal $= e_f^2(w_{fm}^2 + w_{ff}w_{fm}\mu)$
    Twin $= t_f^2$
    Sibling $= s_f^2$
  Residual $= e_f^2[1 - (w_{ff}^2 + w_{fm}^2 + 2w_{ff}w_{fm}\mu)]$
G–E covariance $= 2e_fh_{fc}r_{cf} + 2e_fh_{ff}r_{ff}$
  Total $= t_f^2 + e_f^2 + s_f^2 + + d_f^2 + h_{fc}^2 + h_{ff}^2 + 2r_{cf}e_fh_{fc} + 2h_{fc}h_{ff}a_{cf} + 2e_fh_{ff}r_{ff}$

Table VI. Estimates of the Components of Variance for Church Attendance for the Full and Reduced Models.

	Males		Females	
Component	Full	Reduced	Full	Reduced
Total genes	25.23	26.08	31.78	33.48
Additive genes	21.37	26.08	19.08	33.48
Variance under random mating	16.54	18.87	14.77	24.22
Variance due to assortment	4.83	7.22	4.31	9.26
Dominate genes	3.85	.00	12.70	.00
Total environment	66.12	67.40	58.58	59.79
Family environments	22.76	20.89	21.83	21.22
Twin	10.43	13.35	10.81	13.93
Sibling	8.73	6.12	6.11	6.12
Maternal cultural	3.03	1.42	4.04	1.18
Paternal cultural	.56	.00	.89	.00
Unique environment	43.36	46.51	36.74	38.57
Gene–environment covariance	8.66	6.52	9.64	6.72
Totals	100.0	100.0	100.0	100.0

resulted in changes of only .85, 1.28, and 2.14% in the genetic, environmental, and gene–environment covariance contributions to the variance, respectively. In females, the changes were equally small: 1.70, 1.21, and 2.92%, respectively. Thus, reduction to the final model clarifies rather than changes the inheritance pattern in any substantial way.

It is expected, a priori, that a variable which seeks to index a characteristically *human* trait, such as religious behavior, would maximize our chances of detecting *nongenetic* inheritance. Although it is clear that there are significant shared environmental effects on twins and siblings living together, only a small fraction of these influences derives directly from the parental phenotype, and virtually none from the father. Most of these effects are uncorrelated with the parental phenotype for church attendance and may arise from extrafamilial sources such as peers and teachers. As with almost every other human behavioral trait studied so far, the largest fraction of the environmental variation arises from those specific environmental influences which do not correlate even between identical twins. These analyses are based on cross-sectional data from individuals of widely differing ages. Any tendency for genetic or cultural factors to interact with age would reduce, for example, intergenerational correlations, compared with intragenerational correlations. This effect, in term, would tend to inflate estimates of nonadditive genetic variance and reduce estimates of intergenerational transmission. Further analysis is required to resolve these more complex effects.

An emerging puzzle within human behavioral genetics is the relatively trivial contribution of vertical cultural transmission to individual differences and family resemblance for human characteristics expected to show mainly nongenetic transmission. Assuming that all the studies of individual differences are not intrinsically biased in favor of genetic interpretation, an integrative theoretical framework is needed which can make

sense of these findings. Such a framework cannot come from within behavioral genetics itself, at least as it is currently practiced, because it emphasizes the analysis of the *immediate* causes of individual differences, trait by trait. There has been little attempt to account for the patterns of individual differences across variables in terms of the comparative evolutionary trajectory of populations subject to different regimes of selection and causation. The only possible exception is the concern for analyzing "genetic architecture" as a key to the evolutionary history of traits. In contrast, human sociobiology is replete with fertile speculation about why things are the way they are but lacks data to match the quality and subtlety of those available to the behavioral geneticist.

The large and informative samples now available to the human behavioral geneticist are making it possible to establish a number of replicable findings about the differential causes of family resemblance. Among the issues which could benefit from an evolutionary theoretical perspective are the differential contribution of assortative mating to different variables, the pattern of sex differences in genetic and environmental control, and the contribution (or lack of it) or vertical cultural inheritance to variation in a wide range of human traits. It has long been recognized that the relative uniformity of any environment will leave only genetic effects and errors of measurement to be expressed. Perhaps the relative lack of vertical cultural inheritance has little explanation beyond the fact that cultural change is so rapid and dependent on extrafamilial information that the only differences left to study are the genetic "noise" which creates relatively minor idiosyncracies of taste. If this is so, then behavioral genetics, whatever it might achieve in terms of practical understanding of human diversity at a particular point in history, will say relatively little about what it is to be human. Cohort and secular effects will emerge as major contributors to variation in comparison to individual differences at a particular moment. Of course, our sample includes only adults and does not exclude a major impact of nongenetic effects between parents and their younger children. Alternatively, the apparent lack of vertical cultural inheritance might be attributable to the fact that parents are relatively unreliable as a source of adaptive information in a rapidly changing environment. That is, the optimal strategy for an individual is to select, from the smorgasbord of opportunities available, those in which his/her individual genotype is best able to prosper. One possible consequence of this mechanism is a progressive increase in the genetic component of variation during development. To be equally sensitive to any parental pressure, or to allow parental influence to extend beyond the period of actual physical and economic dependence, may actually be evolutionarily disadvantageous.

Obviously, speculations of this type have no solid theoretical basis at present, but such a foundation is sorely needed in the future if human behavioral genetics is to go beyond mere empirical description of human differences.

ACKNOWLEDGMENTS

This research was supported in part by NIH Grants GM30250 and AG04954, by ADAMHA Grants AA06781, AA07728, AA07535, and MH40828, and by a gift from R.J.R. Nabisco. We thank Louise Totten for assistance with data management and help with checking the expectations for the correlations between relatives. We would also like to acknowledge the AARP for assistance with twin ascertainment and Walter Nance and and Linda Corey for making twins from the VA twin registry available to us.

REFERENCES

Akaike, H. (1970). Statistical predictor identification. *Ann. Inst. Stat. Math.* **21**:243–247.

Burt, C., and Howard, M. (1965). The multifactorial theory of inheritance and its application to intelligence. *Br. J. Stat. Psychol.* **9**:115–125.

Cattell, R. B. (1960). The multiple abstract variance analysis equations and solutions for nature-nuture research on continuous variation. *Psychol. Rev.* **67**:353–372.

Cavalli-Sforza, L. L., and Feldman, M. W. (1973a). Models for cultural inheritance. 1. Group mean and within group variation. *Theor. Popul. Biol.* **4**:42–55.

Cavalli-Sforza, L. L., and Feldman, M. W. (1973b). Cultural vs biological inheritance: Phenotypic transmission from parents to children (a theory of direct effects of parental phenotype on children's phenotype). *Am. J. Hum. Genet.* **25**:618–637.

Cavalli-Sforza, L. L., and Feldman, M. W. (1981). *Cultural Transmission and Evolution: A Quantitative Approach*, Princeton University Press, Princeton, NJ.

Cloninger, C. R., Rice, J., and Reich, T. (1979). Multifactorial inheritance with cultural transmission and assortative mating. II. A general model of combined polygenic and cultural inheritance. *Am. J. Hum. Genet.* **31**:176–198.

Darlington, C. D. (1971). Axiom and process in genetics. *Nature* **234**:521–525.
Eaves, L. J. (1976a). A model for sibling effects in man. *Heredity* **36**:215.
Eaves, L. J. (1976b). The effects of cultural transmission on continuous variation. *Heredity* **37**:41–57.
Eaves, L. J. (1977). Inferring the causes of human variation. *J. Roy. Stat. Soc. A* **104(3)**:324–355.
Eaves, L. J., Last, K. A., Young, P. A., and Martin, N. G. (1975). Model fitting approaches to the analysis of human behavior. *Heredity* **41(3)**:249–320.
Eaves, L. J., Eysenck, H. J., and Martin, N. G. (1989). *Genes, Culture and Personality*, Academic Press, London.
Eysenck, H. J. (1980). *The Causes and Effects of Smoking*, Sage, Beverly Hills, CA.
Fisher, R. A. (1918). The correlation between relatives on the supposition of Mendelian inheritance. *Trans. Roy. Soc. (Edinburgh)* **52**:399–433.
Fuller, J. L., and Thompson, W. R. (1978). *Foundations of Behavior Genetics,* C. V. Mosby, St. Louis.
Heath, A. C. (1983). *Human Quantitative Genetics: Some Issues and Applications*. D.Phil thesis (unpublished), University of Oxford, Oxford.
Heath, A. C., and Eaves, L. J. (1985). Resolving the effects of phenotype and social background on mate selection. *Behav. Genet.* **15**:15–30.
Heath, A. C., Kendler, K. S., Eaves, L. J., and Markell, D. (1985). The resolution of cultural and biological inheritance: Informativeness of different relationships. *Behav. Genet.* **15**:439–465.
Jinks, J. L., and Fulker, D. W. (1970). Comparison of the biometrical genetical, MAVA, and classical approaches to the analysis of human behavior. *Psychol. Bull.* **73(5)**:311–349.
Mather, K. (1949). *Biometrical Genetics: The Study of Continuous Variation,* Methuen, London.
Mather, K., and Jinks, J. L. (1982). *Biometrical Genetics: The Study of Continuous Variation*, 3rd ed., Chapman and Hall, London.
McGue, R., Wette, R., and Rao, D. C. (1984). Evaluation of path analysis through computer simulation: Effects of incorrectly assuming independent distribution of familial correlations. *Genet. Epidemiol.* **1**:255–269.
Morton, N. E. (1974). Analysis of family resemblance. I. Introduction. *Am. J. Hum. Genet.* **26**:318–330.
NAG (1991). *Fortran Library Manual*, Mark 14, Numerical Algorithms Group; Oxford.
New York Public Library (1989). *Desk Reference*, Simon and Schuster, New York.
Rice, C., Cloninger, C., and Reich, T. (1978). Multifactorial inheritance with cultural transmission and assortative mating. I. Description and basic properties of unitary models. *Am. J. Hum. Genet.* **30**:618–643.
SAS Institute Inc. (1990). *SAS Language: Reference, Version 6,* 1st ed., SAS Institute, Cary, NC.
Snedecor, G. W., and Cochran, W. G. (1980). *Statistical Methods.* 7th ed., Iowa State University Press, Ames.
Wright, S. (1921). Correlation and causation. *J. Agr.* **20**:551.

Edited by David W. Fulker