Genetic Contribution to Risk of Smoking Initiation: Comparisons Across Birth Cohorts and Across Cultures

Andrew C. Heath

Departments of Psychiatry, Psychology and Genetics,
Washington University School of Medicine,
St. Louis

Randall Cates

Department of Psychiatry,
Washington University

Nichol G. Martin

Queensland Institute of Medical Research,
Herston, Brisbane,
Queensland, Australia

Joanne Meyer
John K. Hewitt
Michael C. Neale
Lindon J. Eaves

Department of Human Genetics,
Medical College of Virginia,
Richmond

Self-report data on smoking initiation (whether the respondent admitted ever having smoked) were obtained from three large adult twin samples (Australia, $N = 3,808$ pairs; Virginia, $N = 2,145$ pairs; AARP, $N = 3,620$ pairs). Data were broken down into birth cohorts, and genetic models were fitted to test whether the decline, in more recent birth cohorts, in the percentage of individuals becoming smokers has

---

led to a change in the relative contributions of genes and environment to risk of becoming a smoker. Despite a marked change in the proportion of male respondents who reported ever having smoked, we found no evidence for cohort differences in genetic and environmental effects (no Genotype × Cohort interaction). Significant differences in genetic and environmental parameters were found between sexes, and between the Australian and the two U.S. samples. In the U.S. samples, estimates of the genetic contribution to risk of becoming a smoker were 60% in men, 51% in women. In the Australian sample, heritability estimates were 33% in men, but 67% in women. Significant shared environmental effects on smoking initiation also were found, accounting for 23% of the variance in U.S. men, 28% of the variance in U.S. women, 39% of the variance in Australian men, and 15% of the variance in Australian women. In models that allowed for the environmental impact of cotwin smoking on a twin's risk of smoking initiation, estimates of the direct genetic contribution to risk of smoking initiation were comparable or higher (49–58% in U.S. women and 71% in Australian women; 58–61% in U.S. men, and 37% in Australian men).

Smokers, compared to nonsmokers, are more likely to report smoking by other family members, including parents (Chassin, Presson, Sherman, Cortz, & Olshavskyl, 1984; Murphy & Price, 1988; Ogawa, Tominaga, Gellert, & Aoki, 1988; Pederson & Lefcoe, 1986; Swan, Creeser, & Murray, 1990) and siblings (Mittelmark et al., 1987; Ogawa et al., 1988; Swan, Carmelli, Rosenman, Fabsitz, & Christian, 1990). Also they are more likely to report smoking by peers (Chassin et al., 1984; McNeill et al., 1988; Mittelmark et al., 1987; Ogawa et al., 1988; Swan, Creeser, & Murray, 1990). These associations generally have been interpreted using a social learning theory approach (Bandura, 1969; Hunter, Baugh, Webber, Sklov, & Berenson, 1982; USDHEW, 1979) on the assumption that the smoking habit is acquired by imitation, and that individuals with similar histories of exposure to smoking by family members and friends all are equally likely to become smokers. However, because parents and siblings are genetically related to the respondents, we cannot use the familial aggregation of smoking to infer **environmental** causation. A series of studies using twin and adoption data have suggested a substantial genetic contribution to smoking behavior, including the initiation of smoking (Carmelli, Swan, Robinette, & Fabsitz, 1990, 1992; Eaves & Eysenck, 1980; Hannah, Hopper, & Mathews, 1985; Heath & Martin, 1993; Hughes, 1986; Kaprio, Koskenvuo, & Sarna, 1981; Pedersen, 1981; Swan, Carmelli, et al., 1990). These studies suggest that even with identical environmental exposure histories, individuals will differ in their risk of becoming smokers because of underlying genetic differences. The relationship between smoking and peer smoking likewise need not imply any direct causal influence: Individuals who are at increased risk of becoming smokers may choose to associate with others who are similarly predisposed, perhaps because they share similar temperaments that are associated with increased risk of smoking initiation (Cherry & Kiernan, 1976; Eaves & Eysenck, 1980; Zuckerman, Ball, & Black, 1990).

Studying the contribution of genetic variables to initiation and progression of smoking provides a useful model for studying genetic contributions to risk of

illicit substance use and abuse. As in the case of illicit substance use, there are marked cohort effects on smoking initiation, with declining numbers in more recent cohorts reporting ever having smoked (O'Malley, Bachman, & Johnston, 1984). Since the 1960s, there also have been important changes in awareness of the health risks of smoking (NHMRC, 1962; Royal College of Physicians, 1962; USPHS, 1964). These changes might be expected to modify differentially the determinants of smoking initiation in younger compared to older cohorts, perhaps increasing the role of risk taking, impulsivity, and similar heritable personality traits (Eaves, Eysenck, & Martin, 1989; Loehlin, 1992). The classical twin design, by comparing pairs of individuals of the same age, reared in the same home (and therefore with the same exposure to smoking by parents or older siblings), who may either be genetically related as ordinary full siblings (dizygotic or DZ twin pairs) or genetically identical (monozygotic or MZ twin pairs), provides a very sensitive test for changes in the genetic and environmental determinants of smoking initiation across cohorts (Heath et al., 1985; Heath, Neale, Hewitt, Eaves, & Fulker, 1989). In this article we test for the interaction of genetic and environmental effects with birth cohort (Genotype × Cohort interaction), using data from twin samples from Australia (Hannah et al., 1985; Jardine & Martin, 1984) and from the United States (Meyer, Heath, & Eaves, 1992; Meyer, Heath, Martin, & Eaves, 1993).

METHOD

Samples

We present analyses of data from three large twin samples, which we shall refer to as the Australian, Virginia, and American Association of Retired Persons (AARP) samples. The Australian sample derives from a mailed questionnaire survey, in 1980–1982, of 5,967 adult twin pairs on the volunteer National Health and Medical Research Council (NHMRC) twin register (Jardine, 1985; Jardine, Martin, & Henderson, 1984). Completed questionnaires were received from both members of 3,808 twin pairs (64% pairwise response rate), and from only one twin of 576 twin pairs (69% individual response rate). Data from incomplete twin pairs are not used here. These data previously were analyzed by Hannah et al. (1985) and Heath and Martin (1993), who found evidence for a genetic effect on smoking initiation, but did not consider whether there were cohort differences in the magnitude of genetic and environmental influences. The breakdown of sample sizes by zygosity and birth cohort is included in Table 1.[1]

The Virginia and AARP samples provide new data on the smoking habits of twins from the population-based Virginia twin register, and from the volunteer AARP twin panel. Between 1985–1987, questionnaires were mailed to 7,858

---

[1] In some publications, the total sample size for the Australian twin pairs is listed as 3,810 twin pairs. In subsequent mailings to this sample, however, it was discovered that 2 twin pairs had been doubly ascertained, and data from these pairs are included only once in the analyses here.

twins on the Virginia twin register, a register of birth-certificate-confirmed mul-
tiple births in Virginia from 1915 onwards. Completed questionnaires were
returned by 5,339 individuals, a 68% response rate, including 2,179 complete
twin pairs and 981 pairs where only one twin responded. Usable data for this
study were available on 2,145 twin pairs. The AARP sample is a volunteer sam-
ple of twin pairs ascertained through advertisements in AARP publications.
The vast majority of twin pairs on the AARP twin panel were members of the
AARP, and therefore aged 50 or older at the time of the survey. A small number
of twin pairs were younger than 50, most of whom are relatives of AARP mem-
bers. Questionnaires were mailed to 12,118 individuals on the volunteer AARP
twin panel between 1985–1989. Completed questionnaires were returned by
9,423 individuals, a 78% response rate, including 3,957 complete twin pairs and
1,509 twins whose cotwin did not respond. For the purpose of this article, usable
data were available from 3,620 twin pairs of known birth cohort and zygosity
from the AARP twin panel. The breakdown of number of twin pairs by birth
cohort and zygosity group in the Virginia and AARP samples is also given in
Table 1.

Measures

Our measures of smoking initiation, that is, of whether or not a respondent
reported ever having been a smoker, were necessarily based on self-report data:
use of biochemical data to validate current reports of smoking was infeasible in
view of the very large sample sizes required for adequate statistical power to
resolve genetic and nongenetic hypotheses (Martin, Eaves, Kearsey, & Davies,
1978). In the Australian sample, twins were classified as never-smokers if they
reported that they had never smoked cigarettes and gave no answers to questions
about age of onset of smoking and average daily consumption of cigarettes,
cigars, and pipes, but otherwise, they were classified as present or former smok-
ers. The Virginia and AARP surveys shared a single questionnaire, which was
somewhat different in its wording from the Australian questionnaire (Meyer et
al., 1992; Meyer et al., 1993). Respondents were classified as never-smokers if
they reported that they had never smoked and reported no age of onset and no
average daily cigarette consumption (or equivalent in pipefuls or cigars).

In the Australian study and the Virginia study, but not the AARP study, test–
retest data were obtained on a small sample of individuals (n = 96 in the Austra-
lian study; n = 504 in the Virginia study). These data allow us to check on the
consistency of respondents' self-descriptions over time. Test–retest tetrachoric
correlations (Joreskog & Sorbom, 1988) were .98 for the Australian sample and
.99 for the Virginia sample. In the Virginia and AARP studies, we also asked
respondents to report on whether or not their cotwins had ever smoked. These
data provide one check on the truthfulness of respondents in reporting whether
they had ever smoked. Agreement between respondent's self-report and rating
of the respondent by his or her cotwin were, again, very high, with tetrachoric
correlations in the range of .96–.99.

Data Summary

Data were summarized as 2 × 2 contingency tables, computed separately for each zygosity group, cross-classifying the smoking status of first twin (i.e., present or former smoker vs. never smoked) by that of the cotwin. Twins from like-sex pairs were assigned as first or second twins on the basis of birth order where this information was available, but otherwise at random. In the case of opposite-sex pairs, twins were reordered so that the first twin was always the female twin from the pair, the second twin the male twin. Each sample was broken down into birth cohorts, and separate contingency tables were computed for each birth cohort. Differences in the age structure of the three samples necessitated the use of somewhat different ranges for year of birth to ensure adequate sample sizes for each sample for each birth cohort.

Twin tetrachoric correlations for smoking initiation were estimated by the method of maximum likelihood (Eaves, Last, Young, & Martin, 1978; Heath & Martin, 1988; Joreskog & Sorbom, 1986; Tallis, 1962). Estimation of the tetrachoric correlation implies the assumption that the observed discontinuous distribution of never-smoked versus smoker (i.e., current or former smoker, combining these two response categories) is determined by an underlying normally distributed latent variable (liability or "risk" of becoming a smoker) with individuals whose risk value exceeds some critical threshold value (to be estimated from the data) becoming smokers. It is further assumed that the joint distribution of twin pairs for this latent variable is bivariate normal. We cannot directly test this distributional assumption with only dichotomous twin data, but note that this is what we would expect to observe if initiation of smoking is determined by multiple genetic or environmental factors, which combine additively, each individual factor having a relatively small effect.

When estimating twin tetrachoric correlations (Eaves et al., 1978), we estimated separate male and female threshold values (equivalent to estimates, in this example, of the proportions of men and women who become smokers) for each birth cohort, constraining these estimates to be the same in first and second twins from like-sex pairs, and to be the same in MZ and DZ pairs; we estimated simultaneously five twin correlations, that is, separate MZ female, MZ male, DZ female, DZ male, and DZ opposite-sex twin correlations. We computed a likelihood-ratio chi square to assess the goodness of fit of each particular model (e.g., Eaves et al., 1978). We also compared the chi squares obtained when we estimated separate twin correlations for each birth cohort (separate analyses) versus a single set of twin correlations for all cohorts (joint analysis of all the contingency tables). The sum of the chi squares obtained in the separate analyses of each birth cohort was subtracted from the chi square obtained in the joint analysis of all birth cohorts to yield a statistic which is itself distributed as chi square, with number of degrees of freedom equal to the sum of the degrees of freedom for the separate analyses, minus the degrees of freedom for the joint analysis. This chi square statistic was used to test whether there was significant heterogeneity of twin tetrachoric correlations across birth cohorts (heterogeneity

chi square). In some analyses, we also analyzed jointly both Virginia and AARP twin contingency tables, or tables from all three samples, to test for heterogeneity of twin correlations across samples.

Fitting Genetic Models

Genetic and environmental models were fitted to the set of twin contingency tables, separately for each sample, by the method of maximum likelihood (Eaves et al., 1978; Heath & Martin, 1988) using the same assumptions that were used in the estimation of tetrachoric correlations. In model fitting, we compared the fit of three basic models: an additive genetic model, allowing for the additive effects of multiple genes on liability (i.e., on risk of becoming a smoker); a shared environmental model, which assumed that environmental influences on risk of smoking initiation were no more highly correlated in MZ than in DZ twin pairs; and a full model allowing for both additive genetic and shared environmental effects. All models allowed for within-family environmental effects, that is, for those differences in environmental experience that make even an identical twin differ from his or her cotwin, because even MZ twin pairs were imperfectly correlated for their smoking status. Under the additive genetic model, it may be shown (e.g., Mather & Jinks, 1971) that the MZ twin correlations are predicted to be exactly twice the corresponding same-sex DZ correlations. Under the shared environment model, MZ and same-sex DZ correlations are predicted to be equal. Under the full model, the MZ correlations are predicted to be greater than the same-sex DZ correlations, but less than twice the value of these latter correlations.

We also tested for sex differences in the magnitude of genetic and environmental effects. We compared the goodness of fit of (1) additive genetic, shared environmental, or full models that assumed no heterogeneity of genetic or environmental parameters across sexes; (2) corresponding models that allowed for sex differences in the magnitude of genetic and environmental parameters, but assumed that the same genetic effects were influencing smoking initiation in both sexes, or that the same shared environmental effects were influencing smoking in both sexes; and (3) models with sex-dependent genetic or environmental parameters that also allowed for the possibility that some genetic effects, or shared environmental effects, were specific to one sex. This third class of model, which allows for sex-specific genetic or environmental effects, estimates the correlation between shared environmental effects in opposite-sex pairs as a free parameter (rather than fixing this to unity, as in other models), or the correlation between genetic effects in opposite-sex pairs as a free parameter (rather than .5); and predicts that opposite sex twin pair correlations will be lower than both the male like-sex and the female like-sex DZ correlations. Because it is not possible to allow for both sex-specific genetic influences and sex-specific shared environmental influences using only data on twin pairs (Neale & Cardon, 1992) when both genetic and shared environmental parameters were included in a model, we allowed for sex-specific shared environmental effects. The assumptions implied by these models are discussed further in Heath et al. (1989) and Neale and Cardon (1992), and will not be repeated here.

In all model-fitting analyses, variance components were constrained to be non-negative, and in all analyses models were fitted to the raw twin pair contingency tables for each birth cohort, rather than to summary correlations or pooled contingency tables. The overall goodness of fit of each model was assessed by chi-square test, as in the case of analyses estimating polychoric correlations. Likelihood-ratio (chi-square difference) tests were used to determine whether the addition of a parameter to a model led to a significant improvement in goodness of fit, in order to identify the simplest model that was consistent with the observed data. (For technical details of fitting genetic models to twin pair contingency tables, see Eaves et al., 1978.)

Fitting Sibling-Effects Models

A potential criticism of the standard genetic models, when they are applied to such behavioral variables as initiation of smoking, is that they ignore the environmental influence of smoking initiation by one twin on risk of smoking initiation by the cotwin. Therefore, we also examined the goodness of fit of sibling-effect genetic models (Carey, 1986; Eaves, 1976; Heath et al., 1989). In addition to the parameters of the traditional genetic model, these models include parameters representing the environmental impact of smoking by one twin on risk of smoking initiation by the cotwin, and vice versa. If there are both genetic effects and reciprocal sibling effects on risk of smoking initiation, differences in incidence of smoking will be predicted between MZ versus DZ twin pairs (e.g., Carey, 1992). This prediction is a particularly powerful one when, as in the case of smoking initiation, the incidence is changing markedly across cohorts. Technical details of the specification of the sibling-effects model in these analyses are included in the Appendix (see, also, Carey, 1992).

RESULTS

Figure 1 summarizes the proportion of respondents who reported ever having smoked, as a function of sample, birth cohort, and sex. In women, cohort differences were rather modest, with approximately 41–47% of women reporting that they were present or former smokers in all birth cohorts except the oldest Australian (34.1%) and AARP (32.8%) cohorts, and the 1955–1959 Virginia cohort (49.5%). There has, however, been a very pronounced decline in the proportion of male respondents who admit to ever having smoked, from older cohorts to more recent birth cohorts. This decline was especially apparent in the two U.S. samples and, in particular, the Virginia sample, where 78% of men born before 1940 reported being current or former smokers, compared with 42% of men from the youngest cohort (1960–1967).

Table 1 gives the proportion of twin pairs who reported being concordant smokers, that is, where both twins reported having smoked at some stage in their lives, and the proportion of pairs who were discordant, that is, where one twin had never smoked and the other was a present or former smoker. In each sample, in every birth cohort, the proportion of discordant pairs was higher in

AUSTRALIA   VIRGINIA   AARP

S
M
O
K
E
R
S
(%)

80

60

40

20

0

60-64   50-54   pre 36   60-67   50-54   15-39   37-67   16-25
   55-59   36-49   55-59   40-49   26-36   pre 16
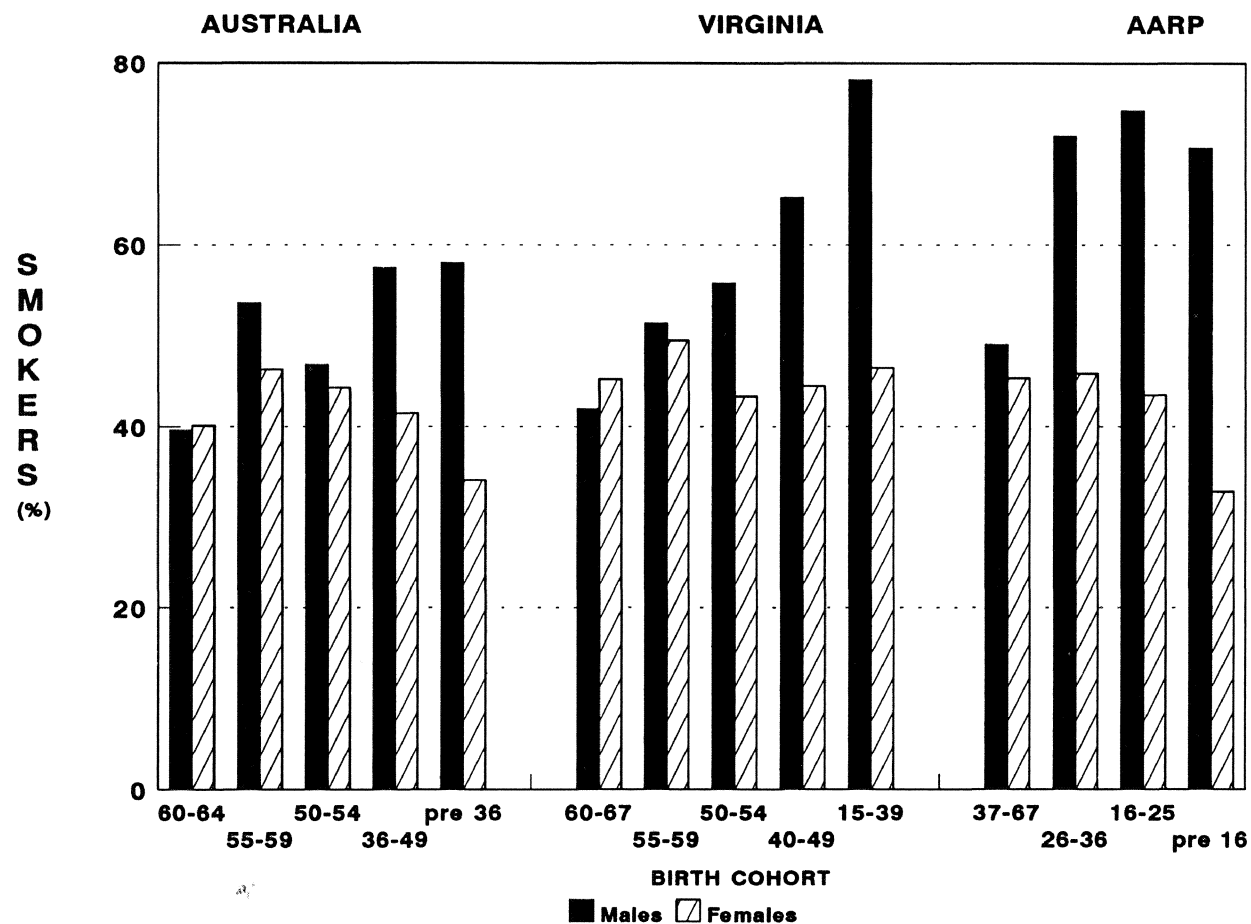
BIRTH COHORT

■ Males ▨ Females

Figure 1.   Proportion of respondents who reported ever having smoked, by birth cohort and sample. (*Note:* 60–64 denotes 1960–1964 birth cohort, etc.)

Table 1. Breakdown of Sample Sizes, and Proportions of Twin Pairs Who Are Concordant Smokers (CC), and Discordant for Ever Having Smoked (CX), by Birth Cohort, Sex, and Zygosity Group

Sample and Birth	MZ Female Pairs			MZ Male Pairs			DZ Female Pairs			DZ Male Pairs			DZ Opposite Sex Pairs		
	n	%CC	%CX	n	%CC	%CX	n	%CC	%CX	n	%CC	%CX	n	%CC	%CX
Australia															
1960–1964	189	27.5	17.5	101	26.7	18.8	127	29.1	27.6	87	21.8	24.1	208	28.4	34.6
1955–1959	220	38.6	15.5	97	36.1	20.6	128	32.8	30.4	72	37.5	33.3	183	33.3	41.0
1950–1954	190	34.7	17.9	103	33.0	27.2	118	29.7	33.9	59	25.4	40.6	140	25.7	39.3
1936–1949	347	30.8	21.0	151	39.7	26.5	206	22.8	37.9	74	44.6	29.7	209	33.5	43.5
pre-1936	286	20.6	21.0	115	46.1	26.1	172	18.6	33.1	60	51.7	25.0	166	21.7	50.6
Virginia															
1960–1967	218	29.8	17.0	100	28.0	13.0	157	33.1	29.9	92	23.9	27.2	189	25.9	34.9
1955–1959	105	36.2	21.0	53	39.6	26.4	73	37.0	34.3	58	32.8	34.5	122	30.3	34.4
1950–1954	75	29.3	18.7	57	49.1	17.5	67	31.3	29.9	54	29.6	33.3	93	28.0	43.0
1940–1949	46	32.6	19.6	62	46.8	19.4	71	25.4	33.8	83	49.4	32.5	150	30.7	46.7
1915–1939	15	20.0	6.7	33	63.6	18.2	15	33.3	33.3	67	65.7	25.4	90	41.1	41.1
AARP															
1937–1967	230	31.3	18.7	100	37.0	20.0	94	31.9	24.5	49	22.5	34.7	116	37.9	33.6
1926–1936	442	33.3	22.6	134	53.0	17.2	216	29.6	34.7	49	63.3	24.5	203	37.9	43.8
1916–1925	502	33.9	21.3	158	65.8	17.1	327	30.0	33.3	81	59.3	24.7	270	35.6	48.5
pre-1916	224	25.5	20.5	86	60.5	14.0	162	19.1	26.5	53	50.9	28.3	124	29.0	52.4

like-sex DZ than in MZ pairs of the same sex, with the single exception of Australian male like-sex twin pairs born prior to 1936. This is consistent with a genetic influence on smoking initiation, with the single discrepant result perhaps being due to sampling variation. The proportion of discordant twin pairs was also generally higher in opposite-sex DZ twin pairs than in same-sex DZ twin pairs. In male like-sex twin pairs, there were pronounced changes in the proportion of twin pairs who were concordant for ever having smoked, the decline in this proportion being particularly pronounced when we compared older versus younger birth cohorts from the two U.S. samples (e.g., 28.0% of Virginia MZ male pairs born 1960–1967 compared to 63.6% of pairs born 1915–1939 are concordant smokers). However, because the proportion of twins who became smokers was also declining, and this decline would itself directly influence the proportion of pairs who are concordant smokers, we must not infer from this trend anything about changes over time in the genetic and environmental influences on risk of becoming a smoker.

Twin Correlations

Tetrachoric correlations for smoking initiation for male MZ and like-sex DZ twin pairs, computed separately for each birth cohort from each sample, are summarized in Figure 2. In estimating the tetrachoric correlations, we allowed for the cohort differences in the incidence of smoking. The consistency of these correlations, despite the pronounced changes in the proportion of men becoming smokers, was striking. A similar consistency was observed for female twin pairs (Figure 3), and for opposite-sex twin pairs (not shown). Only in the case of the Australian sample do we see any consistent trend from older to more recent cohorts, both MZ and DZ twin correlations increasing in more recent cohorts.

Table 2 summarizes goodness-of-fit chi squares obtained when a single set of twin correlations was estimated for all cohorts (joint analysis), separately for each sample, and when separate correlations were estimated for each cohort (separate analyses). Also given are the heterogeneity chi squares that test whether the twin correlations were significantly heterogeneous across birth cohorts. In none of the samples did we find significant heterogeneity of twin correlations across birth cohorts. Also shown in Table 2 are the chi squares obtained when we analyzed the twin pair contingency tables from the two U.S. samples in a single joint analysis, estimating a single set of twin correlations. Subtracting the chi squares obtained under the separate joint analyses of the Virginia and AARP data sets from the chi square under the combined joint analysis yielded a chi-square value, on 5 degrees of freedom, to test whether twin correlations in the two U.S. data sets were significantly heterogeneous. This chi-square value was nonsignificant, indicating no significant differences in the twin correlations between the two samples, $\chi^2(5) = 2.25$, $p = .81$. However, when we performed a joint analysis of all three samples, estimating a single set of correlations, the goodness-of-fit chi square was extremely poor, $\chi^2(177) = 229.62$, $p = .005$; and further analysis confirmed that the twin correlations in the Australian sample did differ significantly from those in the U.S. samples, $\chi^2(5) = 42.24$, $p < .001$.
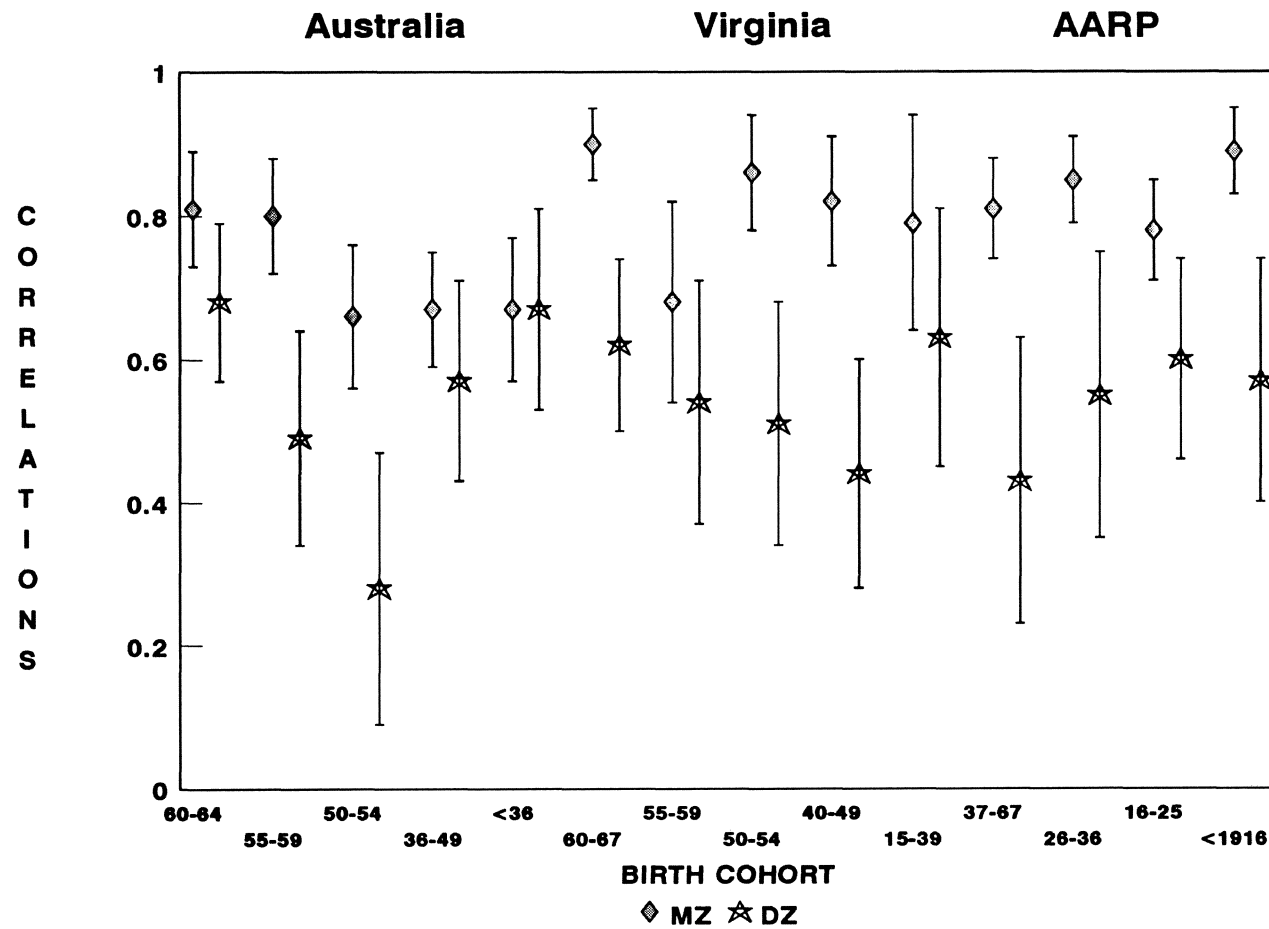
Figure 2. Male like-sex DZ and MZ twin correlations (± SE), by birth cohort and sample. (*Note:* 60–64 denotes 1960–1964 birth cohort, etc.)
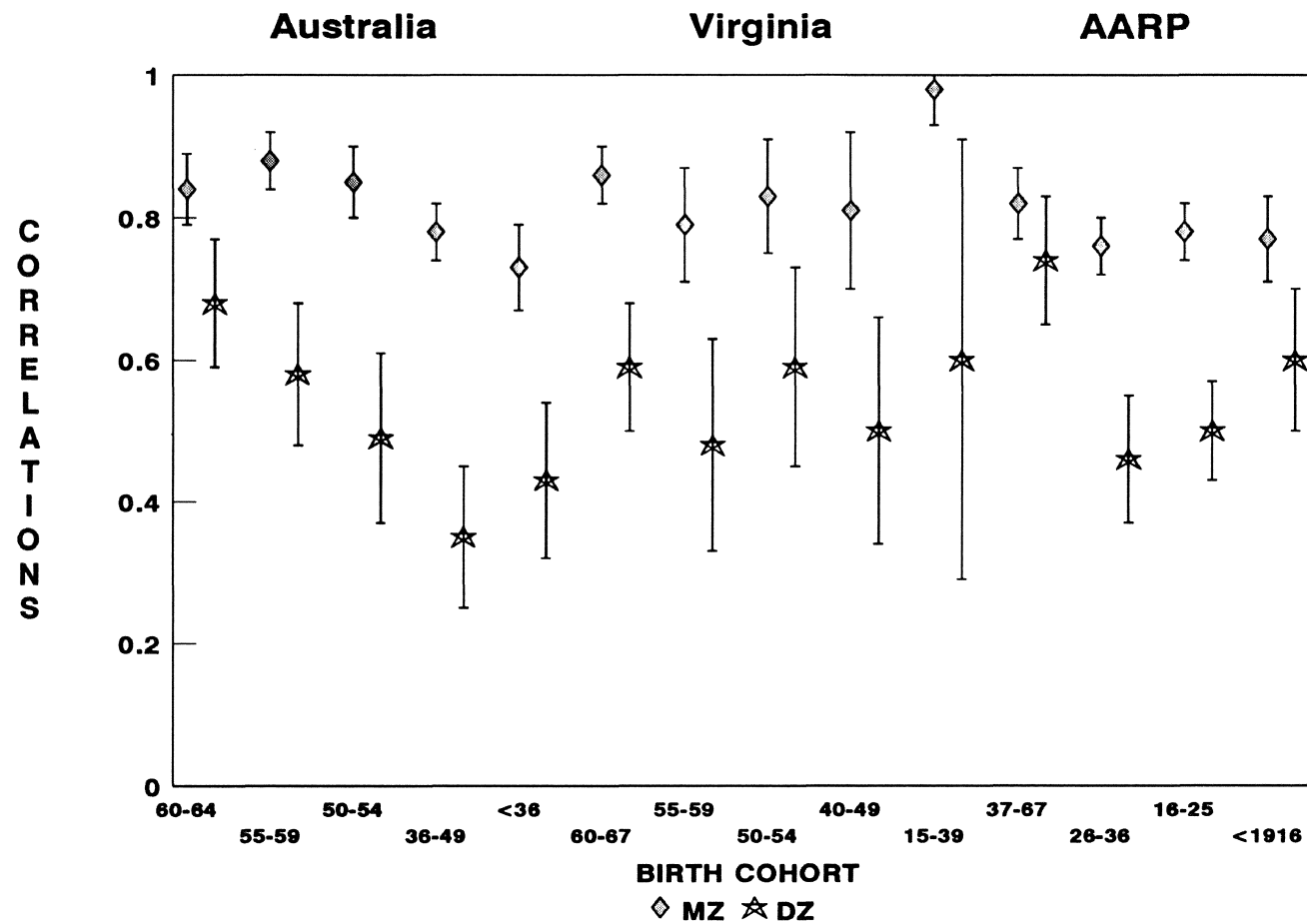
Figure 3.   Female like-sex DZ and MZ twin correlations (± *SE*), by birth cohort and sample. (*Note:* 60–64 denotes 1960–1964 birth cohort, etc.)

Table 2. Heterogeneity of Twin Correlations Across Birth Cohorts

	Separate Analyses			Joint Analyses			Heterogeneity		
	df	χ^2	*p*	*df*	χ^2	*p*	*df*	χ^2	*p*
Australia	40	48.50	.17	60	69.01	.20	20	20.51	.43
Virginia	40	55.68	.05	60	66.22	.27	20	10.54	.12
AARP	32	40.51	.14	47	49.90	.36	15	9.39	.86
Combined U.S. samples	72	96.19	.03	112	118.37	.32	40	22.18	.99

Table 3 summarizes twin tetrachoric correlations, and their standard errors, estimated in the joint analysis of all birth cohorts, separately for each sample, and for the combined analysis of the two U.S. samples. In each sample, we observed strong resemblance of MZ twin pairs (correlations of .72–.83), implying that within-family environmental effects (i.e., those differences in environmental experiences that make even an identical twin differ from his or her cotwin) were having a relatively minor influence on risk of becoming a smoker. MZ twin correlations were substantially higher than the corresponding like-sex DZ twin correlations (correlations of .48–.55), but the latter were greater than one half the MZ correlations, implying that both genotype, and shared environment, were contributing to risk of becoming a smoker. Finally, opposite-sex DZ twin correlations (.35–.37) were lower than both male like-sex and female like-sex DZ correlations in all three samples implying that some environmental influences shared by same-sex pairs might only be influencing smoking initiation in one sex.

Models Ignoring Sibling Effects

Table 4 compares the results of fitting genetic and environmental models to the twin contingency tables in a joint analysis of all birth cohorts. These models ignore sibling effects. Once again, separate analyses are reported for the Australian, Virginia, and AARP samples, as well as a joint analysis of the two U.S.

Table 3. Twin Tetrachoric Correlations for Smoking Initiation: Joint Analysis of Birth Cohorts

	Australian Sample			Virginia Sample			AARP Sample			Combined U.S. Samples		
	n	ρ	*SE*	*n*	ρ	*SE*	*n*	ρ	*SE*	*n*	ρ	*SE*
MZ female pairs	1232	.81	.02	459	.83	.03	1398	.78	.02	1857	.79	.02
MZ male pairs	567	.72	.04	305	.83	.04	478	.83	.03	783	.83	.03
DZ female pairs	751	.48	.05	383	.55	.06	799	.53	.04	1182	.54	.04
DZ male pairs	352	.56	.07	354	.51	.07	232	.55	.09	586	.53	.05
DZ opposite-sex pairs	906	.35	.05	644	.37	.06	713	.35	.06	1357	.36	.04

Table 4. Goodness of Fit of Genetic and Environmental Models in Joint Analysis of All Birth Cohorts

Model	Australia			Virginia			AARP			Combined U.S. Sample		
	df	χ^2	p	df	χ^2	p	df	χ^2	p	df	χ^2	p
1. Genetic	64	82.53	.06	64	<u>72.01</u>	.23	51	63.07	.12	116	137.08	.09
2. Genetic, sex-dependent	63	78.56	.09	63	71.98	.21	50	61.96	.12	115	136.13	.09
3. Genetic, sex-dependent ($r_G < .5$)	62	77.93	.08	62	71.28	.20	49	61.09	.12	114	134.79	.09
4. Shared environment	64	170.09	<.001	64	132.96	<.001	51	133.09	<.001	116	268.18	<.001
5. Shared environment, sex-dependent	63	164.55	<.001	63	131.05	<.001	50	132.73	<.001	115	265.51	<.001
6. Shared environment, sex dependent ($r_c < 1$)	62	121.29	<.001	62	101.07	.001	49	89.26	<.001	114	192.48	<.001
7. Full	63	80.57	.07	63	70.81	.23	50	58.36	.19	115	131.01	.15
8. Full, sex-dependent	61	<u>70.29</u>	.19	61	67.69	.26	48	<u>52.20</u>	.31	113	<u>121.90</u>	.27
9. Full ($r_c < 1$)	60	69.01	.20	60	66.22	.27	47	49.90	.36	112	118.37	.32

Note. r_G denotes correlation between genetic effects in opposite-sex pairs; r_c denotes correlation between shared environmental effects in opposite-sex pairs. The full models allow for both genetic and shared environmental effects on smoking initiation. For each sample, the chi square for the best-fitting model is underlined.

samples. In each sample, models that ignored genetic effects were rejected by a chi-square goodness-of-fit test, but a simple additive genetic model with no sex differences in the magnitude of the genetic influence gave an adequate, albeit somewhat marginal, fit to the data (i.e., $p > .05$). Finally, a model that allowed for the magnitude of genetic and shared environmental influences on smoking initiation to depend upon sex gave a significant improvement in goodness of fit over the additive genetic model by likelihood-ratio chi-square test in all samples except the Virginia sample: Australia, $\chi^2(3) = 12.23$, $p < .001$; Virginia, $\chi^2(3) = 4.32$, $p = .23$; AARP, $\chi^2(3) = 8.87$, $p = .03$; and combined U.S., $\chi^2(3) = 15.18$, $p < .001$. Allowing for sex-specific shared environmental effects gave a nonsignificant improvement in fit in each sample. However, in the joint analysis of the two U.S. samples, this improvement in fit was almost significant, $\chi^2(1) = 3.53$, $p = .08$.

Estimates of genetic and environmental variance components, which give the proportions of the variance in liability (i.e., risk of becoming a smoker) accounted for by genetic, shared environmental, and within-family environmental variables, are summarized in Table 5. Variance components, and their standard errors, are reported both under the best-fitting model, that is, the most parsimonious model consistent with the observed data (Neale & Cardon, 1992) and also under the full model (allowing for sex-dependent genetic and environmental effects with some sex-specific shared environmental influences). Under the full model, additive genetic factors accounted for 60% of the variance in liability in U.S. men, 51% of the variance in U.S. women, 33% of the variance in Australian men, and 67% of the variance in Australian women. Shared environmental influences accounted for 23% of the variance in liability in U.S. men, 28% in U.S. women, 39% in Australian men, and 15% in Australian women. Within-family environmental effects accounted for 28% of the variance in Australian men, but for only 17–21% in U.S. men and women, and Australian women. The estimate of the correlation between shared environmental effects in opposite-sex pairs was .34 in the U.S. sample and .49 in the Australian sample, but in neither case did this estimate differ significantly from unity by likelihood-ratio chi-square tests (compare Models 8 and 9 in Table 4). Under the most parsimonious model, the estimate of the additive genetic variance was increased to 83% in U.S. men, and the estimate of the shared environmental variance in U.S. men was effectively zero, but variance component estimates for U.S. women, and for Australian men and women, changed only modestly.

Sibling-Effects Models

Table 6 summarizes the results of fitting genetic and full models, allowing for a sibling effect (i.e., an environmental influence of smoking by the cotwin), on risk of smoking initiation. In the Australian sample, a genetic model (Model 3) which allowed for sex-dependent genetic effects with some genetic effects also being sex-specific, plus sex-dependent sibling effects, but no shared environmental effects, was the simplest sibling-effects model consistent with the data. The full model, which allowed also for sex-dependent shared environmental

Table 5. Variance Component Estimates (%) Under Best-Fitting Model and Full Models from Joint Analyses of All Birth Cohorts

	Additive Genetic		Shared Environment			Within-Family Environment	
	Males	Females	Males	Females	r_c	Males	Females
Best-Fitting Model							
Australia	28 (15)	77 (05)	43 (13)	4 (04)	1.00	29	19
Virginia	84 (04)	84 (03)	— —	— —	1.00	16	16
AARP	83 (04)	48 (10)	1 (02)	30 (09)	1.00	16	22
Combined U.S.	83 (03)	50 (08)	1 (02)	29 (07)	1.00	16	21
Full Model							
Australia	33 (15)	67 (11)	39 (14)	15 (10)	.49 (.27)	28	18
Virginia	64 (16)	58 (14)	19 (15)	26 (13)	.32 (.28)	17	16
AARP	54 (19)	49 (10)	28 (18)	29 (09)	.32 (.21)	18	22
Combined U.S.	60 (12)	51 (08)	23 (11)	28 (08)	.34 (.17)	17	21

Note. r_c denotes the correlation between family environmental effects in men and women. Standard errors of parameter estimates are given in parentheses. Variance components for within-family environment are obtained by subtracting the additive genetic and shared environmental components from unity.

effects, did not give a significantly better fit than Model 3, $\chi^2(2) = .36$, $p = .84$. However, this full model also did not give a significant improvement in fit compared to the best-fitting model, which ignored sibling effects (Model 8 in Table 4), $\chi^2(3) = 2.54$, $p = .47$. On the basis of these likelihood-ratio chi-square comparisons, therefore, the full model with no sibling effects and the sex-dependent genetic model with sibling effects were equally consistent with our observed

Table 6. Goodness of Fit of Genetic Models Incorporating Sibling Effects

Model	Australia			Combined U.S. Sample		
	df	χ^2	p	df	χ^2	p
1. Genetic	63	80.49	.07	115	130.45	.15
2. Genetic, sex-dependent	61	74.96	.11	113	128.65	.15
3. Genetic, sex-dependent ($r_G < .5$)	60	68.11	.22	112	116.43	.37
4. Full	62	80.37	.06	114	130.45	.14
5. Full, sex-dependent	59	67.75	.20	111	116.10	.35
6. Full, sex-dependent ($r_c < 1$)	58	67.75	.18	110	116.10	.33

Note. r_G denotes correlation between genetic effects in opposite-sex pairs; r_c denotes correlation between shared environmental effects in opposite-sex pairs. The full models allow for both genetic and shared environmental effects on smoking initiation. For each sample, the chi square for the best-fitting model is underlined. In sex-dependent models, the magnitude of the sibling environmental effect, as well as genetic and shared environmental parameters, are allowed to vary as a function of sex.

Table 7.   Parameter Estimates Under Best-Fitting Genetic Models, Allowing for Sibling Effects

	Australia		Combined U.S. Samples			
			(Model 3)		(Model 5)	
Additive genetic variance						
Men	.46	(.17)	.73	(.07)	.70	(.08)
Women	.77	(.05)	.68	(.05)	.45	(.07)
r_G	.17	(.15)	.18	(.10)	1.00	
Shared environmental variance						
Men	.00		.00		.00	
Women	.00		.00		.38	(.09)
r_c	.00		.00		.00	
Within-family environmental variance						
Men	.54		.27		.27	
Women	.23		.32		.17	
Sibling effect						
Men	.20	(.08)	.12	(.05)	.14	(.05)
Women	.06	(.04)	.12	(.04)	−.05	(.04)

Note. Standard errors of parameter estimates are given in parentheses.

data. Table 7 summarizes the parameter estimates under the best-fitting sibling-effect model. The estimated sibling effect was substantial in men (.2) and smaller in women (.06). The additive genetic and within-family environmental variances reported in Table 7 are proportions of the total variance prior to any reciprocal sibling effects (i.e., h^2 = 46% in men, 77% in women). The direct genetic contribution to the total variance (including reciprocal sibling effects, see Appendix) was estimated as 37% in Australian male MZ twins and 71% in Australian female MZ twins (or 39% and 71% if indirect genetic defects also were included), values that were very comparable to the heritability estimates under the full model with no sibling effects (33% and 67%, respectively).

Results of fitting sibling-effect models jointly to the two U.S. samples, shown in Tables 6 and 7, also gave only suggestive evidence for a sibling effect on risk of smoking initiation. As with the Australian data, the simplest sibling-effect model consistent with the data was the model allowing for sex-dependent genetic effects with some sex-specific genetic effects (Model 3 in Table 6): This model gave a significantly better fit than the corresponding model ignoring sibling effects, $\chi^2(2)$ = 18.36, p < .001, and also gave better fits than simpler sibling-effects models; the addition of two shared environmental parameters (the full sibling-effects model, Model 6 in Table 6) did not give a further significant improvement in fit, $\chi^2(2)$ = 0.33, p = .85. However, once again, the full model ignoring sibling effects (Model 9 in Table 5) also did not give a significantly worse fit than the full sibling-effects model, $\chi^2(2)$ = 2.27, p = .32. Compared to the most parsimonious model ignoring sibling effects (Model 8 in Table 5), sibling-effects Model 5 (allowing for sex-dependent genetic and shared environmental effects with r_c = 1) gave an improvement in fit that was just short of significance,

$\chi^2(2) = 5.71$, $p = .06$, and yielded a nonsignificant estimate, very close to zero, for shared environmental effects in men, $\chi^2(1) = 0.19$, $p = .66$. Parameter estimates for the U.S. data are therefore presented in Table 7 for two sibling-effects models: the sex-dependent genetic model (Model 3 in Table 6) and the full sex-dependent model (Model 5, but with shared environmental effects in men fixed to zero). In men, estimates of the genetic contribution to risk of smoking initiation were very comparable: The proportion of the variance excluding sibling effects accounted for by genetic factors (h^2) was 73% under Model 3 and 70% under Model 5; the estimated proportion of the total variance in MZ men accounted for by direct genetic effects was, respectively, 61% and 58% (increased to 62% and 59% when we allowed for indirect genetic effects). In women, under Model 3, corresponding estimates were 68%, 58%, and 59%, respectively. Under Model 5, our estimate of h^2 was 45% in women, with shared environmental factors accounting for an additional 38% of the variance excluding sibling effects. However, our estimate of the sibling effect on risk of smoking initiation ($-.05$) was slightly negative. These values implied that direct genetic effects (and also direct plus indirect genetic effects) were accounting for 49% of the total variance in MZ female pairs.

DISCUSSION

Determinants of Smoking Initiation

We analyzed twin data on initiation of smoking, assessed by retrospective report, on two large adult twin samples from the U.S., the Virginia and AARP samples, and a large Australian adult twin sample. The Virginia sample was a population-based sample ascertained from birth records, the AARP and Australian samples were derived from samples of volunteer twins, recruited through the media and other sources. The AARP sample comprised a high proportion of older twins, whereas the U.S. and Australian samples included a much higher proportion of young adult twins. Despite these differences in sampling, our analyses provided strikingly consistent evidence for a genetic influence on risk of becoming a smoker, confirming previous reports by other researchers (Eaves & Eysenck, 1980; Hannah et al., 1985; Hughes, 1986). The evidence for a genetic influence remained, and estimates of the magnitude of the genetic contribution (heritability) were even slightly increased when we allowed for the environmental influence of smoking by one twin on risk of smoking initiation by the cotwin (sibling effects).

Our confirmation of a genetic influence on risk of smoking initiation was not unexpected. However, the magnitude of that genetic influence was considerably greater than might have been anticipated from some previous reports (Hughes, 1986). Although estimates of the proportion of the variance in risk attributable to genetic factors varied markedly as a function of genetic model, country, and gender, the lowest estimate that we obtained for any group was 28% (Australian men). Under different models, making different assumptions about the presence or absence of sibling effects or other shared environmental effects, heritability estimates ranged from 28–37% for Australian men, 67–77% for Australian women, 45–59% for U.S. women, and 58–83% for U.S. men. It

remains to be determined whether this genetic influence arises because of an association between personality, which is known to be subject to strong genetic influence (Eaves et al., 1989; Loehlin, 1992), and risk of becoming a smoker (Cherry & Kiernan, 1976; Eysenck, 1980; Zuckerman et al., 1990); or because of genetically determined differences in reactivity to nicotine, with some individuals who gave up smoking after very limited experience (e.g., after smoking only a single cigarette) describing themselves as never having smoked, rather than as former smokers; or because of the influence of other heritable factors. It clearly cannot be assumed, as a strict interpretation of social learning theory (Bandura, 1969) would imply, that all individuals with the same exposure to smoking by parents, siblings, or peers are equally at risk of becoming smokers.

In contrast to findings for the inheritance of personality (Bouchard, Lykken, McGue, Segal, & Tellegen, 1990; Eaves et al., 1989; Loehlin, 1992; Pedersen, Plomin, McClearn, & Friberg, 1988) and many other behavioral traits (Plomin & Daniels, 1987) where there is little evidence for any influence of shared environment, we did find a significant shared environmental influence on smoking initiation when we fitted models ignoring sibling effects. Shared environmental effects (which would include the environmental influences of parental smoking and smoking by other siblings, as well as of shared peers, attending the same school, growing up in the same neighborhood, etc.) accounted for a modest 33% of the variance in risk of smoking initiation in U.S. women, 15% of the variance in Australian women, 23% of the variance in U.S. men, and 39% of the variance in Australian men under the full model. The estimate of the correlation between shared environmental effects in opposite-sex twin pairs under the full model ignoring sibling effects, was only .33 in the combined U.S. analyses, and .49 in the Australian analyses, consistent with the possibility that one important influence on smoking initiation is that of same-sex peers, or of similar variables that are more likely to be shared by same-sex than opposite-sex siblings. However, we should note that this parameter did not differ significantly from unity, that is, the full model did not quite give a significant improvement in fit over a model that fixed this correlation to unity in either sample.

When we fitted sibling-effects models, which allowed for the environmental impact of smoking by one twin on risk of smoking initiation by the cotwin, a sex-dependent genetic model (which allowed for no shared environmental influences on risk other than those associated with the reciprocal sibling environmental influences) gave an excellent fit to both Australian and U.S. data sets. As others have commented (e.g., Eaves, 1976), the classical twin design comparing MZ and DZ twin pairs reared together provides a powerful means of detecting shared environmental influences on behavior (Martin et al., 1978) but provides very poor resolution of the extent to which such shared environmental influences are determined by cotwin's behavior versus other features of family background and shared experience. Extension of the classical twin design by collecting data on other relationships, for example, singletons who grow up without any siblings in the family, is needed to improve this resolution (Eaves, 1976). We shall address this issue in future articles on the determinants of smoking initiation, using extended twin-family data.

In both U.S. and Australian samples, the effects of within-family environment (those differences in experience that make even an identical twin differ from his or her cotwin) were remarkably modest, accounting for less than 30% of the variance under the models ignoring reciprocal sibling effects. This figure may be contrasted with findings for the inheritance of personality and other behavioral traits, where within-family environmental effects account for as much as 50% of the variance (Eaves et al., 1989; Loehlin, 1992; Plomin & Daniels, 1987).

Birth Cohort, Sex, and Cross-Cultural Effects

Our primary concern in this article was to examine whether changes in the incidence of smoking, occurring in association with such societal changes as the increase in awareness of the health hazards of smoking, have modified the relative contributions of genes and environment to risk of becoming a smoker. We were able to confirm differences between birth cohorts in the proportion of respondents who admitted ever having smoked, though this cohort effect was most pronounced in male respondents. When we tested for heterogeneity of the relative magnitudes of genetic and environmental influences on smoking initiation across birth cohorts, however, we found no evidence for such cohort differences in any of the three samples. Indeed, we found no significant heterogeneity of genetic and environmental parameters between the AARP and Virginia samples, despite the wide age range covered by these two samples, which included twins born at the end of the 19th century (AARP) as well as twins born in the late 1960s (Virginia). These finding suggest the possibility that the same variables, be they personality traits or other biological effects, have continued to determine which individuals will become smokers, but that in more recent years only individuals with more extreme values on those traits are becoming smokers.

Our analyses did demonstrate an important sex difference in the relative contributions, to risk of becoming a smoker, of genes, shared environment, and within-family environment. In all analyses, models allowing for sex-dependent parameters gave a significantly better fit to the data than models assuming no heterogeneity of genetic and environmental effects. In the two U.S. samples, genetic variables had a somewhat greater impact on risk of becoming a smoker in men than in women. Estimates of the direct genetic contribution in the joint analyses of the two U.S. samples ranged, under different models making different assumptions about the presence or absence of sibling effects or shared environmental effects, from 58%–61% in men (or even as high as 83% if we use the parameter estimates from the best-fitting model ignoring sibling effects), and from 45%–59% in women.

We also found evidence for a significant cross-cultural difference in the magnitude of genetic and environmental influences on risk of becoming a smoker. The sex difference in the magnitude of the genetic influence on smoking initiation was reversed in the Australian sample, with direct genetic influences accounting for only 28–37% of the variance in men, but for as much as 67%–77% of the variance in women. We had no priori expectation that such sex differences in the magnitude of genetic effects would be found. A comparable sex-difference was found for patterns of alcohol use in the Australian sample, with a

much stronger genetic influence in women than in men (Heath, Meyer, Jardine, & Martin, 1991). Until this finding is replicated in a second sample from the Australian population, we cannot exclude the possibility that the relatively weak genetic influence that we observe in Australian men merely reflects an accident of sampling rather than a genuine cross-cultural difference.

A variety of plausible hypotheses, both psychosocial and psychobiological, can be advanced for the existence of sex differences in tobacco use (Grunberg, Winders, & Wewers, 1991), and this applies equally to our finding of sex differences in the genetic influence on risk of smoking initiation. If personality variables play an important role as mediators of the genetic influence on smoking behavior, as some have hypothesized (e.g., Eysenck, 1980), a stronger influence of personality variables in one sex might explain this finding. Little is known about the genetics of nicotine sensitivity in humans, so that the possibility of a sex difference in the magnitude of any genetic influence on nicotine sensitivity cannot be discounted. Clearly, if the reversal of the direction of the sex difference in genetic effects in Australia, compared to the two U.S. samples, is a replicable phenomenon, then this implies that either psychobiological or psychosocial explanations must account for interactions with culture.

It is possible, because our model fitting assumes the environments to which MZ twins are exposed are no more highly correlated than those of DZ twins (except insofar as these environments are shaped by the behaviors of the twins themselves; Plomin & Bergeman, 1991), that we have overestimated the importance of the genetic contribution to risk of smoking initiation. A critical test of this possibility is the extent to which our genetic and environmental parameter estimates are able to predict the correlations between other types of relatives for smoking initiation. As data on the smoking habits of the adult relatives of the three twin samples—their parents, siblings, spouses, and adult children— become available, we shall address this issue further.

When interpreting the findings reported here, it is important to remember that our data consist of the retrospective reports of adult twin pairs, and we cannot be completely certain about how respondents are defining a "smoker." We cannot exclude the possibility that individuals who have experimented with one or two cigarettes have nonetheless classified themselves as nonsmokers. The good stability of respondents' reports over time and the strong agreement between respondents' self-reports and ratings by their cotwin, give us no reason to doubt the accuracy of reports of whether an individual ever smoked. However, it remains to be determined whether the same strong genetic influence on risk of becoming a smoker will be found in cohorts of twins now in adolescence, or whether the pattern of inheritance of smoking initiation has been modified in this younger group.

REFERENCES

Bandura, A. (1969). *Principles of behavior modification.* New York: Holt, Rinehart & Winston.
Bouchard, T.J., Lykken, D.T., McGue, M., Segal, N.L., & Tellegen, A. (1990). Sources of human psychological differences: The Minnesota Study of twins reared apart. *Science, 250,* 223–228.
Carey, G. (1986). Sibling imitation and contrast effects. *Behavior Genetics, 16,* 319–341.

Carey, G. (1992). Twin imitation for antisocial behavior: Implications for genetic and family environment research. *Journal of Abnormal Psychology, 101,* 18–25.

Carmelli, D., Swan, G.E., Robinette, D., & Fabsitz, R.R. (1990). Heritability of substance use in the NAS-NRC twin registry. *Acta Geneticae Medicae et Gemellologiae, 39,* 91–98.

Carmelli, D., Swan, G.E., Robinette, D., & Fabsitz, R. (1992). Genetic influence on smoking—A study of male twins. *New England Journal of Medicine, 327,* 829–833.

Chassin, L., Presson, C., Sherman, S.J., Cortz, E., & Olshavsky, R.W. (1984). Predicting the onset of cigarette smoking in adolescents: A longitudinal study. *Journal of Applied Social Psychology, 14,* 225–243.

Cherry, N., & Kiernan, K. (1976). Personality scores and smoking behaviour: A longitudinal study. *British Journal of Preventive Social Medicine, 30,* 357–379.

Eaves, L.J. (1976). A model for sibling effects in man. *Heredity, 36,* 205–214.

Eaves, L.J., & Eysenck, H.J. (1980). The genetics of smoking. In H.J. Eysenck (Ed.), *The causes and effects of smoking.* London: Maurice Temple Smith.

Eaves, L.J., Eysenck, H.J., & Martin, N.G. (Eds.). (1989). *Genes, culture and personality: An empirical approach.* London: Academic.

Eaves, L.J., Heath, A.C., Neale, M.C., Hewitt, J.K., & Martin, N.G. (1993). Sex differences and non-additivity in the effects of genes on personality. Unpublished manuscript.

Eaves, L.J., Last, K., Young, P.A., & Martin, N.G. (1978). Model-fitting approaches to the analysis of human behavior. *Heredity, 41,* 249–320.

Eysenck, H.J. (Ed.). (1980). *The causes and effects of smoking.* London: Maurice Temple Smith.

Grunberg, N.E., Winders, S.Z., & Wewers, M.E. (1991). Gender differences in tobacco use. *Health Psychology, 10,* 143–153.

Hannah, M.C., Hopper, J.L., & Mathews, J.D. (1985). Twin concordance for a binary trait: 2. Nested analysis of ever-smoking and ex-smoking traits and unnested analysis of a "committed smoking" trait. *American Journal of Human Genetics, 37,* 153–165.

Heath, A.C., Berg, K., Eaves, L.J., Solaas, M.H., Corey, L.A., Sundet H.M., Magnus, P., & Nance, W.E. (1985). Education policy and the heritability of educational attainment. *Nature, 314,* 734–736.

Heath, A.C., & Martin, N.G. (1988). Teenage alcohol use in the Australian twin register: Genetic and social determinants of starting to drink. *Alcoholism: Clinical and Experimental Research, 12,* 735–741.

Heath, A.C., & Martin, N.G. (1993). Genetic models for the natural history of smoking: Evidence for a genetic influence on smoking persistence. *Addictive Behaviors, 18,* 19–34.

Heath, A.C., Meyer, J., Jardine, R., & Martin, N.G. (1991). The inheritance of alcohol consumption patterns in a general population twin sample: 2. Determinants of consumption frequency and quantity consumed. *Journal of Studies on Alcohol, 52,* 425–433.

Heath, A.C., Neale, M.C., Hewitt, J.K., Eaves, L.J., & Fulker, D.W. (1989). Testing structural equation models for twin data using LISREL. *Behavior Genetics, 19,* 9–35.

Hughes, J.R. (1986). Genetics of smoking: A brief review. *Behavior Therapy, 17,* 335–345.

Hunter, S.M., Baugh, J.G., Webber, L.S., Sklov, M.C., & Berenson, G.S. (1982). Social learning effects on trial and adoption of cigarette smoking in children: The Bogalusa Heart Study. *Preventive Medicine, 11,* 29–42.

Jardine, R. (1985). *A twin study of personality, social attitudes and drinking behavior.* Unpublished doctoral thesis, Australian National University, Canberra.

Jardine, R., & Martin, N.G. (1984). Causes of variation in drinking habits in a large twin sample. *Acta Geneticae Medicae et Gemellologiae, 33,* 435–450.

Jardine, R., Martin, N.G., & Henderson, A.S. (1984). Genetic covariation between neuroticism and the symptoms of anxiety and depression. *Genetic Epidemiology, 1,* 89–107.

Joreskog, K., & Sorbom, D. (1986). *PRELIS: A preprocessor for LISREL.* Mooresville, IN: Scientific Software.

Joreskog, K., & Sorbom, D. (1988). *LISREL VII: Analysis of linear structural relationships.* Mooresville, IN: Scientific Software.

Kaprio, J., Koskenvuo, M., & Sarna, S. (1981). Cigarette smoking, use of alcohol and leisure-time physical activity among same-sexed adult twins. In L. Gedda, P. Parisi, & W.E. Nance (Eds.), *Twin research: Vol 3. Epidemiological and clinical studies.* New York: Alan R. Liss.

Loehlin, J.C. (1992). *Genes and environment in personality development.* Newbury Park, CA: Sage.

Martin, N.G., Eaves, L.J., Kearsey, M.J., & Davies, P. (1978). The power of the classical twin study. *Heredity, 40,* 97–116.

Mather, K., & Jinks, J.L. (1971). *Biometrical genetics: The study of continuous variation.* London: Chapman & Hall.

McNeill, A.D., Jarvis, J.J., Stapleton, J.A., Russell, M.A.H., Eiser, J.R., Gammage, P., & Gray, E.M. (1988). Prospective study of factors predicting uptake of smoking in adolescents. *Journal of Epidemiology and Community Health, 43,* 72–78.

Meyer, J.M., Heath, A.C., & Eaves, L.J. (1992). Using non-metric multidimensional scaling to explore the dimensionality of categorical multifactorial traits. *Genetic Epidemiology, 9,* 87–107.

Meyer, J.M., Heath, A.C., Martin, N.G., & Eaves, L.J. (1993). *Genetic and environmental influences on onset of the smoking habit and quantity smoked.* Unpublished manuscript.

Mittlemark, M.B., Murray, D.M., Luepker, R.V., Pechacek, T.F., Pirie, P.L., & Pallonen, U.E. (1987). Predicting experimentation with cigarettes: The Childhood Antecedents of Smoking Study (CASS). *American Journal of Public Health, 77,* 206–208.

Murphy, N.T., & Price, C.J. (1988). The influence of self-esteem, parental smoking, and living in a tobacco production region on adolescent smoking behaviors. *Journal of School Health, 58,* 401–405.

National Health and Medical Research Council (NHMRC). (1962). *Smoking and lung cancer: Report of 53rd Session of NHMRC.* Canberra, Australia: Commonwealth Government Printer.

Neale, M.C., & Cardon, L.R. (Eds.). (1992). *Methodology for genetic studies of twins and families* (NATO ASI Series D, Behavioral and Social Sciences, Vol. 67). Doordrecht, Netherlands: Kluwer Academic.

Ogawa, H., Tominaga, S., Gellert, G., & Aoki, K. (1988). Smoking among junior high school students in Nagoya, Japan. *International Journal of Epidemiology, 17,* 814–820.

O'Malley, P.R., Bachman, J.G., & Johnston, L.D. (1984). Period, age, and cohort effects on substance use among American youth, 1976–1982. *American Journal of Public Health, 74,* 682–688.

Pedersen, N.L. (1981). Twin similarity for usage of common drugs. In L. Gedda, P. Parisi, W.E. Nance (Eds.), *Twin research: Vol. 3. Epidemiological and clinical studies.* New York: Alan R. Liss.

Pederson, L.L., & Lefcoe, N.M. (1986). Change in smoking status among a cohort of late adolescents: Prediction and explanation of initiation, maintenance and cessation. *International Journal of Epidemiology, 15,* 519–526.

Pedersen, N.L., Plomin, R.L., McClearn, G.E., & Friberg, L. (1988). Neuroticism, extraversion, and related traits in adult twins reared apart and reared together. *Journal of Personality and Social Psychology, 55,* 950–957.

Plomin, R.L., & Bergeman, C.S. (1991). The nature of nurture: Genetic influence on "environmental" measures. *Behavioral and Brain Sciences, 14,* 373–385.

Plomin, R.L., & Daniels, D. (1987). Why are children in the same family so different from one another? *Behavioral and Brain Sciences, 10,* 1–60.

Royal College of Physicians. (1962). *Smoking and health.* New York: Pitman Publishing.

Swan, G.E., Carmelli, D., Rosenman, R.H., Fabsitz, R.F., & Christian, J.C. (1990). Smoking and alcohol consumption in adult male twins: Genetic heritability and shared environmental influences. *Journal of Substance Abuse, 2,* 39–50.

Swan, G.E., Creeser, R., & Murray, M. (1990). When and why children first start to smoke. *International Journal of Epidemiology, 19,* 323–330.

Tallis, G.M. (1962). The maximum-likelihood estimation of correlation from contingency tables. *Biometrics, 18,* 342–353.

United States Department of Health, Education, and Welfare (USDHEW). (1979). *Smoking and health—A report of the Surgeon General* (PHS Publication No. 79-5066). Washington, DC: U.S. Government Printing Office.

United States Public Health Service (USPHS). (1964). *Smoking and health: A report of the Advisory Committee to the Surgeon General of the Public Health Service* (PHS Publication No. 1103). Atlanta, GA: U.S. Department of Health, Education, and Welfare, Center for Disease Control.

Zuckerman, M., Ball, S., & Black, J. (1990). Influences of sensation seeking, gender, risk appraisal and situational motivation on smoking. *Addictive Behaviors, 15,* 209–220.

APPENDIX

In this appendix, we summarize the assumptions that were used, following Carey (1992), when we fitted genetic models allowing for reciprocal sibling environmental effects, that is, the environmental impact (represented by reciprocal paths i in men, i' in women) of smoking by one twin or risk of smoking by the cotwin, and vice versa. (Strictly, because we are using a "threshold" or liability model, the sibling-effects model allows for the environmental influence of having a cotwin who is high on risk of smoking initiation). Under such sibling-effects models, if genetic effects are also present, the expected variances on the underlying scale of liability or "risk," as well as the expected covariances of twin pairs, will vary as a function of zygosity (Carey, 1986; Eaves, 1976). For discontinuous variables such as smoking initiation, differences in variance in liability will lead to differences in incidence between MZ and DZ pairs, provided that the prevalence of smoking initiation differs from 50% (Carey, 1992). We denote by r_j the expected correlation between twin pairs from the jth twin group ($j = 1, \ldots, 45$ for the joint analysis of data from five zygosity groups from nine age cohorts, in the analyses of the U.S. data; $j = 1, \ldots 25$ for the Australian analyses) in the absence of reciprocal sibling environmental effects, expressed as a function of genetic and shared and nonshared environmental parameters; we assume, for convenience, that our latent liability variable, in the absence of reciprocal sibling interaction, is standardized to have a mean of zero and a variance of one. The expected variance of twins from the jth twin group, in the case of female like-sex twin pairs, will be

$$V_{Fj} = (1 - i'^2)^{-2} (1 + i'^2 + 2i' \, r_j),$$

and the corresponding expected covariance will be

$$C_j = (1 - i'^2)^{-2} [(1 + i'^2) \, r_j + 2i']$$

(Carey, 1986), with expected variances and covariances for male like-sex pairs being obtained by substituting i for i'. In the case of opposite-sex twin pairs, we cannot estimate simultaneously both an environmental effect of smoking by a male twin on risk of smoking initiation by the female cotwin, and the environmental effect of smoking by a female twin on risk of smoking initiation by the male cotwin. We have, therefore, assumed that the environmental impact on a female twin of having a smoking cotwin does not depend upon whether the cotwin is female or male, and likewise that the environmental impact on risk of smoking initiation by a male twin of smoking by his cotwin does not depend upon the sex of the cotwin. This is likely, at best, to be a reasonable approximation to reality. We also considered the possibility that reciprocal sibling effects were so small as to be effectively zero in opposite-sex pairs (Carey, 1992), in preliminary analyses of these data, but this model did not, in general, lead to sensible parameter estimates. Assuming the same parameter value (i) for sibling effects on male twins from like-sex and opposite-sex twin pairs, and similarly the

same parameter value (i') for sibling effects on female twins, then the expected variance for male twins from opposite-sex twin pairs will be

$$V_{Mj} = (1 - i\,i')^{-2}\,(1 + i^2 + 2i\,r_j),$$

and that for female twins from opposite-sex pairs will be

$$V_{Fj} = (1 - i\,i')^{-2}\,(1 + i'^2 + 2i'\,r_j),$$

and the covariance of opposite-sex twin pairs will be

$$C_j = (1 - i\,i')^{-2}\,[(1 + ii')\,r_j + i + i'].$$

Because we have assumed standardization to unit variance prior to sibling interaction, we must rescale threshold values (e.g., Carey, 1992) as

$$t'_{Mjk} = t_{Mk}/V_{Mj}{}^{.5},$$

where t'_{Mnk} denotes the standardized threshold (for men) for the nth zygosity group ($n = 1, \ldots 5$), from the kth birth cohort, and t_{Mk} is the estimated threshold value for men from the kth birth cohort, corresponding to the case of no reciprocal sibling interaction. Under nongenetic models, sibling effects will not generate differences in variance as a function of zygosity, and so it is not possible to obtain separate estimates of the contribution of sibling effects, and other shared environmental influences, to risk of smoking initiation. The goodness of fit of nongenetic sibling-effect models will, therefore, be identical to the goodness of fit of nongenetic models that ignore sibling effects. Therefore, we report the goodness of fit only for sibling-effects models that also allow for a genetic contribution to risk of smoking initiation.

The goodness of fit of sibling-effects models was assessed by chi-square tests in the same manner as for conventional genetic models. Likelihood-ratio chi-square tests were used to compare the fit of different sibling-effects models and to compare the fit of the best-fitting sibling-effects model to that of the models ignoring sibling effects. As with the conventional genetic analysis, genetic and environmental variance components were constrained to be non-negative.

When significant sibling effects are found, there are several ways in which we can summarize the genetic contribution to risk of smoking initiation. We have computed the proportion of the variance in liability, excluding that component of the variance which arises through sibling effects, that is explained by genetic effects. This quantity, which we shall denote by h^2, is needed to predict correlations in liability for other relationships where sibling effects are not occurring (e.g., separated twin pairs with no adoptive siblings). We also report the proportion of the total variance (including that component of variance arising through sibling effects) that is attributable to direct genetic effects, computed as

$$h^2/(1 + i^2 + 2i\,r_j)$$

(substituting i' for i in the case of female like-sex twin pairs); we also report the proportion of the total variance that is attributable both to direct genetic effects, and to the indirect effects arising through the reciprocal influence of each twin's liability on the liability of the cotwin, computed as

$$h^2 (1 + i^2)/(1 + i^2 + 2\ i\ r_j).$$

We note that these two latter quantities will be different for different twin groups and report them only for same-sex MZ pairs.