Reliability and Reporting Biases for Perceived Parental History of Alcohol-Related Problems: Agreement between Twins and Differences between Discordant Pairs*

WENDY S. SLUTSKE, PH.D., ANDREW C. HEATH, D.PHIL., PAMELA A. F. MADDEN, PH.D., KATHLEEN K. BUCHOLZ, PH.D., STEPHEN H. DINWIDDIE, M.D., MICHAEL P. DUNNE, PH.D., DIXIE J. STATHAM, B.A., AND NICHOLAS G. MARTIN, PH.D.*

Department of Psychiatry, Washington University School of Medicine, 4940 Children’s Place, St. Louis, Missouri 63110

ABSTRACT. Objective: Previous research suggests that family history of alcoholism assessments may be biased by characteristics of the informant. In this report, the reliability and potential biases in offspring reports of paternal and maternal alcohol-related problems were examined in a large community sample of adult twins. Method: Subjects were volunteer participants in the Australian NH&MRC twin registry. Agreement between twin pairs on reports of paternal and maternal alcohol problems was assessed in 2,657 twin pairs (1,444 female-female pairs, 626 male-male pairs, and 587 female-male pairs). In addition, to detect systematic reporting biases, like-sex twin pairs whose paternal alcohol problems reports disagreed (n = 164) were contrasted on measures of personality, state anxiety and depression, parental rearing, alcoholism, and alcohol use. Results: Twin agreement for parental alcohol-related problems was good, with overall kappas of .66 for paternal and .58 for maternal alcohol problems. When discordant twin pairs were compared, we found that women who reported that their father had alcohol problems were significantly lower on EPQ-R Social Conformity than their twin sister who denied paternal alcohol problems; and there was a trend for men who reported that their father had alcohol problems to be higher in negative perceived parenting from father than their twin brother who denied paternal alcohol problems. Twins discordant for reporting paternal alcohol problems did not, however, differ on the major dimensions of personality, state anxiety and depression, alcoholism, or current alcohol use. Conclusions: The results of this study bolster our confidence in using the family history method to examine characteristics of offspring of alcoholics versus offspring of nonalcoholics on self-reported measures of personality and psychopathology, but suggest that some caution should be exercised when using this method to examine differences in offspring-reported perceptions of parental rearing practices. (J. Stud. Alcohol 57: 387-395, 1996)

HIGH-RISK STUDIES have been valuable in establishing the familiality of alcoholism and identifying possible risk factors and precursors of the disorder. Many such studies have used the family history method in which individuals self-identify themselves as offspring of alcoholics or nonalcoholics. These family history studies have found differences between self-identified offspring of alcoholics and offspring of nonalcoholics in several important self-reported domains of functioning, including rates of psychiatric disorders (Dawson et al., 1992; Mathew et al., 1993; Sher et al., 1991), measures of personality (Knowles and Schroeder, 1990; Martin and Sher, 1994; Sher et al., 1991), alcohol expectancies (Mann et al., 1987; Sher et al., 1991), interpersonal functioning (Domenico and Windle, 1993; Greenfield et al., 1993; Johnson and Pandina, 1991), perceptions of parental rearing practices, quality of the parent-offspring relationship, and the childhood family environment (Benson and Heller, 1987; Jarmas and Kazak, 1992; Johnson and Pandina, 1991). Studies in which both familial alcoholism risk and the dependent variables of interest are assessed by self-report of the same individual may, however, be prone to reporting biases that would exaggerate or create artificial associations between reported parental alcoholism history and these dependent variables. It is important to know to what extent the family history method, based on individuals identifying themselves as offspring of alcoholics or nonalcoholics, is accurate and unbiased in order to evaluate such studies.

Investigations of the psychometric properties of family history assessments of alcoholism have consistently demonstrated high interrater reliability (i.e., kappas > .80) (Andreaesen et al., 1977; Crews and Sher, 1992; Zimmerman et al., 1988) and high test-retest reliability (kappas > .70) (Crews and Sher, 1992; Schuckit et al., 1995). The evidence for the validity of assessments of family history of alcoholism is less positive. When family history assessments are compared to direct interviews of target family members, the specificity is consistently high (i.e., > .90), but sensitivities are in the range of about .50; in other words, only about one-half of those meeting alcoholism criteria by direct assessment are identified as alcoholic by reports of their family members (Andreaesen et al., 1986; Crews and Sher, 1992;
O'Malley et al., 1986; Roosa et al., 1993; Thompson et al., 1982), although the sensitivity may be somewhat higher than this when only offspring ratings of their parents are considered (Kosten et al., 1992; Roy et al., 1994). Nonetheless, even with a sample of offspring of women who earlier had received inpatient alcoholism treatment, only 76% of the interviewed offspring correctly identified their mother as having a history of alcoholism (Smith et al., 1994). Thus, some misclassification of family risk status inevitably will occur when the family history method is used.

With less than perfect classification of subjects into high-risk and low-risk groups, true differences between groups may be attenuated (Schuckit et al., 1995). Alternatively, the family history method, compared to the family study method (in which relatives are directly interviewed), may exaggerate true differences between groups. Kendler et al. (1991) drew attention to this possibility in a study of female twin pairs discordant for lifetime psychopathology. They showed that women with a history of major depression were significantly more likely than their nondepressed co-twin to report that their mother had a history of major depression as well. In other words, classification into high-risk and low-risk groups on the basis of self-report family history would have led to women with a personal history of problems (major depression, in this example) being disproportionately represented in the high-risk group. Thus, an artifactual association between psychopathology and family risk status would have been inferred even though none existed (since each depressed twin had a nondepressed co-twin). In the same study (Kendler et al., 1991), the results for alcoholism were in the same direction; that is, women with a history of alcoholism were more likely than their nonalcoholic co-twin to report a history of alcoholism in their mother and father, but the differences were not statistically significant.

Attempts to determine the relative veracity of family history reports of alcoholic versus nonalcoholic informants by comparing their agreement with direct interview or best estimate diagnosis of alcoholism in a target family member have not led to a straightforward answer. Two studies suggest that when the family history reports of alcoholic and nonalcoholic siblings are discrepant, the alcoholic informant is more likely to be correct (Chapman et al., 1994; Rice et al., 1995), but two other studies suggest the opposite (Roy et al., 1994; Smith et al., 1994). All of the studies suggest that the family history method may have systematic biases.

In this study, we use a large community sample of adult twin pairs to investigate: (1) the reliability of the family history method for diagnosing parental alcohol-related problems, and (2) potential biases that the family history method may introduce into studies comparing “high-risk” and “low-risk” groups, classified solely on the basis of self-report family history, on self-reported putative risk factors for alcoholism. First, we examine twin agreement for reports of paternal and maternal alcohol problems to extend previous findings using sibling pairs (Crews and Sher, 1992; Rhea et al., 1993; Sher and Descutner, 1986). Second, we explore potential bias in studies of family history of alcoholism by taking advantage of the less-than-perfect agreement between twins. We sought to determine, using twins who disagreed about a parent’s alcoholism status, whether there are systematic differences between individuals who do and do not report that their parent has had alcohol problems, controlling for the parent’s true alcoholism status. Although we did not have information about the true alcoholism status of the parents, we knew that both twins from a pair had the same parent, with the same alcoholism history. Thus, twins who disagreed about their father’s alcoholism status were contrasted on self-report measures of personality, state anxiety and depression, parental rearing, alcoholism, and alcohol use (there were too few twin pairs discordant for reporting maternal alcohol problems for similar comparisons). Given the (albeit equivocal) literature on bias in family history of alcoholism assessments and the literature on recall biases of self-report measures in general (Collins et al., 1985; Giovanucci et al., 1993; Weinstock et al., 1991), we predicted that twins who report that their father had alcohol-related problems would be more “deviant,” i.e., more likely to have a history of alcoholism, lower on measures of behavioral control, higher on measures of neuroticism, anxiety, depression and alcohol use, and more likely to give negative reports about their parents’ rearing styles compared to their co-twins who denied paternal alcohol problems. In other words, we anticipated that any reporting bias would be in the direction that would inflate differences between high-risk and low-risk groups in conventional family history comparisons.

Method

Subjects

The subjects were participants in the Australian National Health and Medical Research Council (NH&MRC) twin panel, a volunteer twin registry recruited through the media, schools and a variety of other sources (Jardine and Martin, 1984; Kendler et al., 1986). Three major alcohol-related surveys have been conducted to date with this sample: a mailed questionnaire survey in 1980-81 (Jardine and Martin, 1984), a follow-up mailed questionnaire survey in 1988-89 (Heath et al., 1994b; Heath and Martin, 1994) and a telephone interview survey conducted in 1992-93 (Heath et al., 1994a). Data obtained from the telephone interview survey and the 1988-89 questionnaire survey were analyzed for the present study.

Telephone interviews were attempted with any twin pairs where at least one twin either responded to the 1988-89 survey or had participated in an alcohol challenge study conducted in 1978-79 (Heath and Martin, 1992; Martin et al., 1985a,b). Telephone interviews were completed with 5,995 individuals: 2,087 men (mean ±SD age = 42.7 ± 11.8 years, range = 28-89) and 3,908 women (mean age =
44.7 ± 12.6 years, range = 27-90); 87.5% of men and 90.6% of women respondents to the 1988-89 mailed questionnaire were interviewed by telephone in 1992-93.

The final sample sizes for the analyses included 2,657 complete twin pairs (1,444 female-female twin pairs, 626 male-male twin pairs, and 587 female-male twin pairs) who provided information on maternal history of alcohol problems, and 2,589 complete twin pairs (1,405 female-female twin pairs, 611 male-male twin pairs, and 573 female-male twin pairs) who provided information on paternal history of alcohol problems (see Table 1). Table 1 presents the numbers of twin pairs concordant and discordant for perceived paternal and maternal alcohol problems. There were 122 female-female and 66 male-male twin pairs discordant for perceived paternal alcohol problems. Of these, there were 91-110 female-female pairs and 43-54 male-male pairs who provided usable information on the mailed questionnaire measures of personality, state anxiety and depression, and perceived negative parenting, and 122 female-female and 64 male-male pairs who provided enough information at interview to make a positive or negative diagnosis of alcohol dependence. Due to the low base rate of maternal alcohol-related problems, the number of twin pairs discordant for perceived maternal alcohol problems was too few to conduct further analyses of discordant pairs.

Measures

Measures used in the present study were obtained from two surveys of the Australian NH&MRC twin registry. Perceived family history of alcohol-related problems and personal history of alcohol dependence and alcohol use were assessed in the telephone interview survey conducted in 1992-93. Telephone interviews have become increasingly common in psychiatric research, particularly in family studies (Weissman et al., 1986) or in surveys of large representative national samples (e.g., Lyons et al., 1995). Several studies suggest that telephone interviews yield information comparable to that obtained by face-to-face interview (e.g., Fenig et al., 1993, Siemiatycki, 1979; Wells et al., 1988).

The remaining measures of personality, state anxiety and depression, and perceived negative parenting were assessed in the mailed questionnaire survey conducted in 1988-89. These measures were assessed about 4 years prior to the assessment of perceived family history of alcohol-related problems. We report relatively high test-retest reliabilities of these measures over periods of 2 to 8 years. The high degree of longitudinal stability of these measures suggests that the results of this study would not be substantially altered had these measures been obtained concurrently with the interview assessments. Nonetheless, we note this as a potential study limitation.

Perceived family history of alcohol-related problems. The FHAM interview (Family History Assessment Module; Rice et al., 1995), a family history assessment originally developed for the Collaborative Study on the Genetics of Alcoholism (COGA), was modified for use as a telephone interview in Australia. Paternal and maternal history of alcohol problems, for the purposes of this study, were each assessed by a single item that was asked for both the respondent’s father and mother: “Has drinking ever caused your (natural) father/mother to have problems with health, family, job or police, or other problems?”

Alcohol dependence and use. The SSAGA interview (Semi-Structured Assessment for the Genetics of Alcoholism; Bucholz et al., 1994), also originally developed for the COGA study, was modified for use as a telephone interview in Australia. Individual alcohol dependence symptoms were aggregated into a lifetime DSM-III-R alcohol dependence diagnosis (American Psychiatric Association, 1987) by computer algorithm. Bucholz et al. (1994) report high one week interrater test-retest reliabilities of alcohol dependence diagnoses using the SSAGA in face-to-face interviews (Cohen’s kappa = .84-90, Yule’s Y = .87-.90). Previous studies have found no difference in the rates of alcohol dependence assessed by in-person versus telephone interview (Kendler et al., 1992), and high levels of diagnostic agreement for alcohol dependence assessed by the two interviewing modes (kappas ≥ .84) (Paulsen et al., 1988; Watson et al., 1992), which are comparable to the reliability of alcohol dependence diagnoses assessed twice using the same mode of interviewing (Bucholz et al., 1994). Thus, the use of telephone interviews seems to be justified for the assessment of alcohol dependence in this large national sample of twins.

Assessments of alcohol use were also obtained from the SSAGA interview. These included the frequency of alcohol consumption in the last 12 months and the quantity of alcohol

<table>
<thead>
<tr>
<th>Group</th>
<th>Paternal alcohol problems</th>
<th>Maternal alcohol problems</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-/-</td>
<td>+/-</td>
</tr>
<tr>
<td>Female-female pairs</td>
<td>1,094</td>
<td>122</td>
</tr>
<tr>
<td>Male-male pairs</td>
<td>477</td>
<td>66</td>
</tr>
<tr>
<td>Female-male pairs</td>
<td>441</td>
<td>F+:+26</td>
</tr>
</tbody>
</table>

Note: -/- = pairs concordant for reporting no paternal/maternal alcohol problems; +/- = pairs discordant for reporting paternal/maternal alcohol problems; +/- = pairs discordant for reporting paternal/maternal alcohol problems; F+ = females reporting a paternal/maternal alcohol problem; M+ = males reporting a paternal/maternal alcohol problem. Table entries are numbers of twin pairs in each category.
consumed per drinking occasion (typical amount in the last 12 months, maximum amount in a single day in the last 12 months, and maximum amount ever in a single day). Quantity of alcohol consumption measures were log-transformed prior to statistical analysis to reduce the positive skew of the distributions.

**Personality.** A 54-item short form of the Tridimensional Personality Questionnaire (TPQ; Cloninger et al., 1991), a 48-item short-form of the Eysenck Personality Questionnaire (EPQ-R; Eysenck et al., 1985), and a 50-item measure of social attitudes (adapted from Wilson and Patterson, 1968) were assessed by mailed questionnaire in 1988-89. The TPQ responses yielded scores on the three scales of Novelty-seeking, Harm Avoidance and Reward Dependence (Cloninger, 1987a,b), and the responses to the EPQ-R items yielded scores for Extraversion, Neuroticism, Toughmindedness ("Psychoticism" scale), and Social Conformity ("Lie" scale). Treloar (1992) reports high 8-year test-retest reliabilities for EPQ-R Neuroticism ($r = .67$) and Extraversion ($r = .73$) in Australian women, and Heath et al. (1994b) report high 2-year longitudinal stabilities for all of the TPQ and EPQ-R scales, except for the EPQ-R Toughmindedness scale, in Australian men and women ($r = .62-.84$ for TPQ scales; $r = .69-.87$ for EPQ-R Extraversion, Neuroticism and Social Conformity; and $r = .36-.72$ for EPQ-R Toughmindedness) (see Heath et al., 1994b, for further details about these measures).

On the measure of social attitudes, subjects indicated whether they agreed or disagreed with a variety of topics. For the purposes of this study, the responses were scored for a social conservatism scale (e.g., items such as opposition to abortion and to divorce, belief in the truth of the bible and church authority) and a political conservatism scale (e.g., opposition to multiculturalism, government welfare, and immigration, belief in the death penalty and stiffer jail terms). These conservatism scales, previously identified through factor analysis (Heath and Martin, 1993) are only modestly intercorrelated in Australian subjects ($r = .20$) and exhibit high long-term stability (8-year test-retest reliability of .63-.76 for social conservatism and .49-.74 for political conservatism) (Heath and Martin, 1993). Because of their high longitudinal stability, the social and political conservatism scales may be best characterized as measures of enduring personality traits (Heath and Martin, 1993). We anticipated that social conservatism would be protective against alcohol-related problems and other high-risk behaviors such as smoking, whereas political conservatism might be positively associated with risky behaviors because of opposition to state attempts to influence individual health-related habits (Heath et al., 1995).

**State anxiety and depression.** Subjects completed the 7-item State Anxiety and State Depression scales from the Delusions-Symptoms-States Inventory (DSSI; Bedford and Foulds, 1977). The scales contain items such as "Recently I have worried about every little thing" (State Anxiety) or "Recently I have been so miserable that I have had difficulty with my sleep" (State Depression). Subjects chose a response on a 1-4 scale which indicated for each statement whether they felt that way "not at all," "a little," "a lot" or "unbearably." Responses for each scale were summed across items, yielding scores ranging from 7 to 28. Because the distribution of scores on these scales were positively skewed, analyses were performed on log-transformed state anxiety and depression scores.

It should be noted that these measures of state anxiety and depression were collected about 4 years prior to the interview in which family history of alcohol-related problems was assessed. State anxiety and depression assessments were not made at the 1992-93 interview. Thus, we are unable to say to what extent current anxious or depressed mood affects family history reports of alcohol problems, but only the extent to which previously measured anxious or depressed mood affects family history reports of alcohol problems. Nonetheless, Treloar (1992) reports high longitudinal stabilities for DSSI State Anxiety and Depression in Australian women (2-year test-retest: $r = .73$ for anxiety and .67 for depression; 8-year test-retest: $r = .50$ for anxiety and .48 for depression); thus these measures may also tap enduring trait-like qualities of an individual.

**Perceived negative parenting.** Subjects completed a 14-item short form of the Parental Bonding Instrument (Parker, 1990; Todd et al., 1994). For this measure, subjects reported on a 1-4 scale how much ("not at all," "a little," "somewhat" and "a lot") their mother and father engaged in certain parenting behaviors when the respondent was growing up. Subjects rated their mother and father separately with identical 7-item scales. For the purposes of this study, six of the seven items with clearly positive or negative valences were chosen and summed to form perceived negative parenting from mother and perceived negative parenting from father scales. Items included "let me do those things I liked doing" (reverse scored), "seemed emotionally cold to me," "appeared to understand my problems and worries" (reverse scored), "made me feel I wasn't wanted," "tried to make me dependent on him/her" and "was overprotective of me." Todd et al. (1994) report 1-year test-retest reliabilities of .55-.66 for 3- and 4-item subscales of the short form of the Parental Bonding Instrument. Prior to statistical analyses, scores were log-transformed to reduce the positive skew of the distributions of the perceived negative parenting scales.

**Data analysis**

Cohen's kappa (see Fleiss, 1981) was calculated as an index of twin agreement for perceived parental alcohol problems. The kappa statistic is widely used to assess interrater reliability, but it has the undesirable property of dependence on base rate (Grove et al., 1981; Spitznagel and Helzer, 1985). We present another agreement statistic that does not have this undesirable property: Yule's $Y$. In population-based studies, in which the base rate of a diagnosis will be low, or when comparing two diagnoses of differing base rates (such as paternal versus maternal alcohol problems), Yule's $Y$ is preferred to kappa as an index of reliability (Spitznagel and Helzer, 1985).
Two independent sample Yule’s Ys, for example, calculated in men versus women, were compared using the Breslow-Day test for the homogeneity of odds ratios. Because Yule’s Y is a monotonic transformation of the odds ratio, statistical tests of the difference between odds ratios can also be used to test the difference between Yule’s Ys. Two non-independent sample Yule’s Ys, for example, calculated for paternal versus maternal alcohol problems in the total sample, cannot be compared using standard statistical tests because standard errors are underestimated. For comparisons of non-independent Yule’s Ys, bootstrapped standard errors were computed (see Efron and Tibshirani, 1986). Ten thousand replicate samples were drawn from the original dataset, using sampling with replacement. For each replicate sample, odds ratios were computed for twin agreement for paternal and maternal alcohol problems, and the difference between the logarithm of the odds ratios was calculated. The mean and standard deviation of the difference between the log odds ratios over the 10,000 replicate samples provided a bootstrapped estimate of the population mean and standard error of the log odds ratio difference. An odds ratio difference significantly different from zero indicates that the two odds ratios, and thus the two Yule’s Ys, differ significantly from each other. A similar procedure was used to obtain bootstrapped estimates of the standard error of the difference between two non-independent kappas.

Twin pairs discordant for perceived paternal alcohol-related problems were compared in a matched pairs case-control analysis. The rate of lifetime DSM-III-R alcohol dependence in twins who reported a positive paternal history of alcohol problems was compared with the rate in their co-twins who did not report paternal alcohol problems using McNemar’s test. The rate of frequent alcohol use (more than once a week) in the last 12 months was similarly compared in twins discordant for perceived paternal alcohol problems. For the remaining continuous variables (quantity of alcohol use, TPQ and EPQ-R, social and political conservatism, State Anxiety and Depression, and perceived negative parenting from mother and father scale scores), the case and co-twin-control groups were compared with the matched pair t test. T tests were evaluated at two significance levels: \( p < .05 \) and \( p < .001 \). The latter is a conservative significance level which is adjusted for multiple tests using the Bonferroni correction. When hypothesis tests are not independent, for example, when comparing group differences on a number of correlated measures (as in the present study), this procedure will result in an overcorrection, or an overly conservative criterion for significance. Because the magnitude of bias in the family history method was of interest, the effect size is also presented. The effect size was calculated as the standardized mean difference of the score in the twin who reported paternal alcohol problems minus the score in the co-twin who did not report paternal alcohol problems.

### Results

**Twin agreement for perceived paternal and maternal alcohol problems**

Overall, 17.4% of twins reported that their father and 3.4% reported that their mother had alcohol-related problems. Rates of perceived paternal history of alcohol problems did not vary by the sex of the respondent (17.5% among women versus 17.2% among men; \( \chi^2 = 0.11, 1 \text{ df}, p = .74 \)), but women were more likely than men to report that their mother had alcohol problems (3.9% of women versus 2.6% of men; \( \chi^2 = 5.64, 1 \text{ df}, p = .02 \)).

Table 2 presents indices of agreement between twins for paternal and maternal alcohol-related problems. On the whole, twin agreement for parental alcohol problems was good. Although inspection of the kappas suggested that the level of agreement was lower for maternal alcohol problems than paternal alcohol problems, none of the comparisons of the kappas yielded statistically significant differences. In fact, the overall level of agreement as indexed by Yule’s Y was significantly higher for maternal alcohol problems than paternal alcohol problems (0.82 versus 0.73; \( t = 2.95, 2.573 \text{ df}, p = .003 \)), and these overall differences were mainly due to higher levels of agreement for maternal versus paternal alcohol problems among twin daughters (0.85 versus 0.77; \( t = 2.41, 1.396 \text{ df}, p = .016 \)). Other Yule’s Y differences in offspring agreement for paternal versus maternal alcohol problems (e.g., for male-male and female-male twin pairs), although in the same direction, were not statistically significant.

Inspection of Table 2 suggests that female-female twin pairs exhibited better agreement than either male-male pairs or female-male pairs for both paternal and maternal alcohol problems. Nonetheless, the only Yule’s Y difference that was statistically significant was agreement for maternal alcohol problems in female-female versus female-male pairs (0.85

### Table 2

| Table 2. Agreement for perceived paternal and maternal alcohol-related problems between adult Australian twin pairs |
|---------------------------------|---------------------------------|
| **Paternal alcohol problems**   | **Maternal alcohol problems**   |
| \( n \)                           | \( \kappa \) (95% CI)           | \( \kappa \) (95% CI)           |
| Female-female twin pairs         | \( 1,405 \) .70 (.65-.76)       | \( 1,444 \) .65 (.60-.70)       |
| Male-male twin pairs             | \( 611 \) .61 (.53-.69)         | \( 626 \) .54 (.46-.62)         |
| Female-male twin pairs           | \( 573 \) .61 (.52-.69)         | \( 587 \) .39 (.31-.47)         |
| All twin pairs                   | \( 2,589 \) .66 (.62-.70)       | \( 2,657 \) .58 (.54-.62)       |
| \( \text{CI} \) = confidence interval; entries with the same superscripted letter differ from each other at \( p < .05 \). | \( \text{CI} \) = confidence interval; entries with the same superscripted letter differ from each other at \( p < .05 \). |
versus 0.71; \( \chi^2 = 5.61, 1 \text{ df}, p = .018 \), and this difference was confirmed by comparisons of the corresponding kappas. Examination of the composition of discordant unlike-sex twin pairs (see Table 1) indicated that there were no systematic biases in unlike-sex pairs, such as over- or underreporting of paternal or maternal alcohol problems in one sex compared to the other. Although men from unlike-sex pairs were more likely to report that their father had alcohol-related problems than their female co-twin, the numbers of the two types of discordant pairs (39 pairs where only the male twin reported a paternal alcohol problem versus 26 pairs where only the female twin reported a paternal alcohol problem) were not significantly different from each other (matched-pair odds ratio = 1.5; 95% confidence interval = 0.89-2.57). The higher rate of female compared to male offspring-reported maternal alcohol problems in the total sample was not observed within unlike-sex twin pairs; there were nearly equal numbers of the two types of discordant pairs (11 versus 12).

**Differences between twins discordant for perceived paternal alcohol-related problems**

Table 3 presents matched-pair odds ratios comparing twin pairs discordant for perceived paternal alcohol-related problems on self-reported lifetime DSM-III-R alcohol dependence. Among female twin pairs discordant for perceived paternal alcohol problems, the rate of DSM-III-R alcohol dependence was slightly lower among twins who reported paternal alcohol problems than that in their co-twins who reported no paternal alcohol problems (6.6% versus 11.5%), but the matched-pair odds ratio of the association between perceived paternal alcohol problems and lifetime alcohol dependence in women was not statistically significant (OR = 0.54; 95% confidence interval = 0.18-1.45). This may be due in part to the low rate of informative female twin pairs doubly discordant for both perceived paternal alcohol problems and alcohol dependence diagnosis (20 of 122 twin pairs). Among male twin pairs discordant for perceived paternal alcohol problems, the rate of DSM-III-R alcohol dependence was slightly but not significantly higher among twins who reported paternal alcohol problems compared to the rate in their co-twins who reported no paternal alcohol problems (34.4% versus 26.6%; matched-pair odds ratio = 1.83; 95% confidence interval = 0.62-5.04). As for women, there were few twin pairs discordant for both paternal history of alcohol problems and a diagnosis of alcohol dependence (17 out of 64).

The quantity and frequency of alcohol use, assessed at the 1992-93 interview, were also examined in the discordant twin pairs. There was no evidence to suggest that twins who reported that their father had a history of alcohol problems drank any more or less in the past year than their co-twin who denied paternal alcohol problems (see Table 3 for results for frequency of alcohol use and Table 4 for results for quantity of alcohol use). In response to a question about the maximum quantity of alcohol ever consumed in a single 24 hour period, men who reported paternal alcohol problems had a slightly lower lifetime maximum quantity of alcohol consumption than their co-twins who denied paternal alcohol problems (\( d = -.30, t = 2.37, 63 \text{ df, } p = .02 \); see Table 4), although this result was not statistically significant when a Bonferroni correction for multiple comparisons was applied.

Table 4 also displays the results of the comparisons of twin pairs discordant for perceived paternal history of alcohol problems on the questionnaire measures. Generally, the differences in personality, state anxiety and depression, and perceived negative parenting between the discordant twin pairs were small and nonsignificant, except for Social Conformity in women and perceived negative parenting from father in men. Women who reported that their father had a history of alcohol problems were lower on EPQ-R Social Conformity.
("Lie" scale) than their twin sisters who denied such a family history of alcohol problems \( (d = -0.33, t = -3.45, 109 \text{ df}, p = .0008) \). Men who reported that their father had a history of alcohol problems were higher on the perceived negative parenting from father scale than their twin brothers who reported a negative paternal history of alcohol problems \( (d = 0.36, t = 2.55, 49 \text{ df}, p = .014) \), although this result was not statistically significant when a Bonferroni correction for multiple comparisons was applied. In sum, there were few statistically significant differences between twins who reported a positive versus a negative paternal history of alcohol problems. Only three out of 36 comparisons were statistically significant at the \( p < .05 \) level—a result not much different than what would be expected by chance. Not only were there few statistically significant differences, but the magnitude of the differences were quite modest; the mean effect size magnitude for the 32 comparisons presented in Table 4 was only 0.11. Furthermore, differences between discordant twins were typically opposite in sign to those predicted by theories of the etiology of alcoholism and the literature on children of alcoholics (Sher, 1991). For example, the family history positive twins were lower on TPQ Novelty-seeking, EPQ-R Neuroticism, and State Anxiety, and higher on social conservatism than their family history negative co-twins.

**Discussion**

In a large community sample of adult Australian twin pairs, the interrater reliability of perceived parental history of alcohol-related problems was examined. Consistent with previous investigations, the reliability of perceived paternal and maternal alcohol problems was quite good (overall kappa = .66 for paternal alcohol problems, .58 for maternal alcohol problems), and this despite the fact that the assessment consisted of a single item. Other studies (e.g., Crews and Sher, 1992) also found that a single item assessment of parental alcohol problems performed well. It appears that, at least with respect to reliability, the use of such family history assessments is justified. Nonetheless, without direct interviews with the parents of these twins, we are unable to say to what extent our single item assessment of parental alcohol problems is valid.

Previous studies that have examined the effect of informant gender on the accuracy of parental history of alcohol problems reports have yielded somewhat mixed results. Searles et al. (1993) did not find statistically significant differences in reported rates of familial alcoholism in male and female college students, whereas Prescott et al. (1994) found that women reported higher rates of parental alcohol problems than did men. Similarly, some family and sibling studies have found no gender of informant effect (Crews and Sher, 1992; Rice et al., 1995; Smith et al., 1994), whereas others found that reports of female informants produced higher sensitivities (Kosten et al., 1992; Roy et al., 1994), or that sibling agreement was higher for reports of opposite-sex parents (Rhea et al., 1993). In the present study, women reported higher rates of maternal alcohol problems than did men, and there was a nonsignificant but consistent trend for agreement between sisters to be higher than agreement between brothers for both paternal and maternal alcohol problems. Agreement for perceived maternal alcohol problems was lower among unlike-sex than like-sex twin pairs, and this did not appear to be due simply to a difference between men and women in their threshold for considering a given behavior or set of behaviors an "alcohol problem." If there is a difference between men and women in their reporting of parental alcohol problems, the reason for it is unclear.

The results of previous studies (e.g., Roy et al., 1994) suggest that there are biases in family history of alcoholism studies that may exaggerate or attenuate observed differences between offspring of alcoholics and offspring of nonalcoholics. We found little evidence to suggest systematic biases when family history reports are used to compare offspring of alcoholics with offspring of nonalcoholics on measures of personality and psychopathology. Out of 36 comparisons of twin pairs discordant for perceived paternal history of alcohol-related problems, only three were statistically significant at the \( p < .05 \) level, with differences between groups of only about one-third of a standard deviation on these three measures.

In the present study, when female twin pairs discordant for perceived paternal history of alcohol problems were compared on a variety of measures, they differed significantly only on the EPQ-R Lie, or Social Conformity scale. This scale was originally developed to identify respondents who were denying minor, nearly universal faults, and contains items such as, "Are all your habits good and desirable ones?" and "If you say you will do something, do you always keep your promise?" Subsequently, studies have shown that the EPQ-R Lie scale is correlated with other forms of social rebelliousness, so that this scale can also be used as an index of social conformity (Francis et al., 1991, McCrae and Costa, 1985). Thus, the finding of lower Social Conformity scores in women who reported that their father had alcohol problems compared to their twin sisters who denied such problems has at least two interpretations. One interpretation is that the women who reported that their father had alcohol-related problems were generally more willing to admit undesirable qualities in themselves and their father than were their co-twins. Another interpretation is that women who perceived their father as having alcohol problems were more socially rebellious and nonconforming than their sisters who did not perceive their fathers as having alcohol problems. However, this latter interpretation is inconsistent with the fact that along with lower Social Conformity scores, the twins who perceived their father as having alcohol problems were also lower on TPQ Novelty-seeking and higher on social conservatism than their twin sisters. The pattern of results suggests that the Social Conformity difference in female twin pairs discordant for perceived paternal alcohol problems may be an indication of more truthful responding in general in the twins who reported that their father had alcohol problems than in their
co-twins who did not admit such problems. However, without information directly obtained from the parents of these twins, the veracity of these competing hypotheses cannot be tested.

Male twin pairs discordant for perceived paternal history of alcohol-related problems differed on the perceived negative parenting from father scale, but this association was not significant after a conservative Bonferroni correction for multiple significance tests. The relationship between perceived parenting and reports of parental alcohol-related problems, if such exists, is not entirely clear. Perhaps twins whose fathers were less caring and more overprotective towards them in comparison to their co-twin were more likely to identify their father as having alcohol problems because they were also exposed to more alcohol problem behaviors. On the other hand, the actual behavior of the parent may have been fairly similar and members of twin pairs may have different perceptions of the same behaviors: twins with more negative perceptions of their father’s parenting may also be more likely to report that the father has had alcohol problems. Direct interviews of the fathers of these twins are necessary to sort through these alternative explanations. Regardless of the interpretation, we cannot exclude the possibility that measures of offspring-reported perceptions of parental rearing practices, quality of the parent-offspring relationship, and the childhood family environment may be biased in family history studies.

In summary, we found good interrater reliability for a single item measure of parental alcoholism history, and no evidence to suggest that the family history method produced results that were biased with respect to the respondent’s own self-reported history of alcohol dependence, quantity and frequency of current alcohol use, or the major dimensions of personality and affect. The results of this study bolster our confidence in using the family history method to examine characteristics of offspring of alcoholics versus offspring of nonalcoholics on self-reported measures of personality and psychopathology.

**Acknowledgments**

We thank Edward Spitznagel, Ph.D., for his statistical advice, especially in the use of bootstrapping methods. We also acknowledge the assistance of S. Mason, P. Saunders, J. Pearson, O. Zheng, M. Hodge, R. Cates, the interviewers, and the twins from the Australian NH&MRC for their continuing cooperation. Two anonymous reviewers provided helpful comments on an earlier version of the manuscript.

**References**


