Long-Term Stability and Heritability of Telephone Interview Measures of Alcohol Consumption and Dependence

Narelle K. Hansell,[1] Arpana Agrawal,[2] John B. Whitfield,[1] Katherine I. Morley,[3] Gu Zhu,[1] Penelope A. Lind,[1] Michele L. Pergadia,[2] Pamela A. F. Madden,[2] Richard D. Todd,[2] Andrew C. Heath,[2] and Nicholas G. Martin[1]

[1] Genetic Epidemiology, Queensland Institute of Medical Research, Brisbane, Australia

[2] Department of Psychiatry, School of Medicine, Washington University, United States of America

[3] Centre for Molecular, Environmental, Genetic and Analytic Epidemiology, School of Population Health and Applied Genetics Unit, ORYGEN Research Centre, Department of Psychiatry, University of Melbourne, Australia

Alcohol dependence symptoms and consumption measures were examined for stability and heritability. Data were collected from 12,045 individuals (5376 twin pairs, 1293 single twins) aged 19 to 90 years in telephone interviews conducted in three collection phases. Phases 1 and 2 were independent samples, but Phase 3 targeted families of smokers and drinkers from the Phase 1 and 2 samples. The stability of dependence symptoms and consumption was examined for 1158 individuals interviewed in both Phases 1 and 3 (mean interval = 11.0 years). For 1818 individuals interviewed in Phases 2 and 3 (mean interval = 5.5 years) the stability of consumption was examined. Heritability was examined for each collection phase and retest samples from the selected Phase 3 collection. The measures examined were a dependence score, based on DSM-IIIR and DSM-IV criteria for substance dependence, and a quantity × frequency measure. Measures were moderately stable, with test–retest correlations ranging from .58 to .61 for dependence and from .55 to .64 for consumption. However, the pattern of changes over time for dependence suggested that the measure may more strongly reflect recent than lifetime experience. Similar to previous findings, heritabilities ranged from .42 to .51 for dependence and from .31 to .51 for consumption. Consumption was significantly less heritable in the younger Phase 2 cohort (23–39 years) compared to the older Phase 1 cohort (28–90 years).

Genetic research on alcohol dependence, whether using genetic linkage or association or latent variable behavioral genetic approaches, has conventionally relied upon the analysis of categorical diagnoses, operationalized according to one or more diagnostic criteria sets (e.g., Heath, Bucholz et al., 1997; Kendler et al., 2006; Reich et al., 1998). In a contrast of clinically ascertained cases and controls, categorical diagnoses and the underlying criterion items, as assessed by structured interview, show acceptable reliability (Bucholz et al., 1994; Bucholz et al., 1995) and validity (Hesselbrock, Easton, Bucholz, Schuckit, & Hesselbrock, 1999). However, such case series disproportionately represent severely affected individuals, in whom difficulties in the discrimination between unaffected and affected are minimized (Bucholz et al., 1996). In contrast, in general community samples, the majority of individuals reporting a history of alcohol dependence will be individuals with milder disorder (Heath, Bucholz et al., 1994; Lynskey et al., 2005). A growing literature, initially using latent class analysis (Bucholz et al., 1996; Heath, Bucholz et al., 1994; Lynskey et al., 2005), and more recently Item Response Theory (Saha, Chou, & Grant, 2006) suggests that alcoholism may best be viewed as at least a semi-continuous trait, avoiding an arbitrary dichotomization into unaffected and affected individuals. Such an approach would have the potential to increase the power of genetic research approaches.

A second literature has for many decades documented significant genetic variance in alcohol consumption patterns (Heath, 1995). Short-term test-retest reliability of consumption measures in adults, typically with reassessment at 1 to 6 weeks, is modest to good for a typical quantity measure (drinks per drinking day: retest correlations .48-.88), better for frequency of consumption (number of days used alcohol: retest correlations .64–.85), and highest for quantity × frequency (i.e., approximating total drinks per week: .73–.91; (Friesema et al., 2004; Grant, Harford, Dawson, Chou, & Pickering, 1995; Gruenewald & Johnson, 2006; Webb, Redman, Gibberd, & Sanson-Fisher, 1991). Test–retest correlations are notably lower over longer follow-up periods (see Table 1), though still

substantial, particularly for cohorts that were older at baseline assessment and using frequency or quantity × frequency measures. There is evidence for strong genetic overlap between heaviness of consumption and alcohol dependence risk (Heath & Martin, 1994; Whitfield et al., 2004), suggesting that it may be useful to take advantage of quantitative consumption measures as covariates in genetic research on alcoholism.

In this article, the long-term test-retest reliability and heritability of a dependence symptom count measure and a quantity × frequency consumption measure, will be examined in studies in the Australian twin panel. Their potential informativeness for gene-mapping studies of alcohol use disorder will be evaluated.

Materials and Method

Participants

Data collected from 12,045 individuals (5376 twin pairs and 1293 single twins) were examined. For 2976 individuals (1099 twin pairs and 778 single twins), data were collected on two occasions. The sample was drawn in large part from earlier studies, as shown in Figure 1, with the current report focused on data collected in three phases of telephone interviews. Phase 1 data were collected from 1992 to1993, Phase 2 data were collected

from 1996 to 2000 and Phase 3 data were collected from 2001 to 2005. These studies are described in Table 2.

Sample sizes for each collection phase and retest are shown in Table 4, in addition to the age of participants, the proportion of each sample that is female, and the proportion that is monozygotic (MZ). For the Phase 1/Phase 3 retest sample, the mean time interval between data collection phases was 11.0 years (SD = 0.9, range 8.7–13.1). For the Phase 2/Phase 3 retest, the mean time interval was 5.5 years (SD = 1.6, range 1.1–9.6).

Ethics approval was received from the institutional review boards (Queensland Institute of Medical Research and Washington University School of Medicine) appropriate to each study. In addition, informed verbal consent was received from participants in the Phase 1 and Phase 2 studies, and informed verbal and written consent was received from Phase 3 study participants. Abstainers (individuals who had never tried alcohol — 3%, 1%, and 4% of individuals at Phases 1, 2, and 3 respectively) were excluded from all analyses and are not included in sample numbers.

Zygosity

Zygosity was initially determined by self-report questionnaire with standard questions regarding physical

---

Table 1

Long-Term Stability Studies of Self-Reported Alcohol Consumption in General Population Samples

Study	N (Age)	Collection method	Interval	Findings
Kaprio et al. (1992)	13,404 on two occasions (at time 1: older group 24–43 yrs, younger group 18–23 yrs)	Questionnaire	6 years	Older Group Quantity: $r = .64 – .70$ Frequency, beer: $r = .52–.53$ Quantity: $r = .39–.51$ Frequency, beer: $r = .30–.32$
Pape & Hammer (1996)	1084 on 3 occasions (19–22 years at time 1)	Questionnaire	2 years 4 years 6 years	Quantity × Frequency: $r = .54$ (2 years) $r = .37$ (4 years) $r = .35$ (6 years) Frequency: $r = .72$ (2 years) $r = .64$ (4 years) $r = .58$ (6 years)
Wennberg et al. (2000)	80 males on 2 occasions (18 years at time 1, 34–37 years at time 2)	Interview	~18 years	Quantity: $r = .24$ Frequency: $r = .17$
Kerr et al. (2002)	(a) 7127 on 3 occasions (25–74 yrs at time 1) (b) 3113 on 4 occasions (17–62 yrs at time 1)	Questionnaire	(a) 5 years ~10 yrs (b) 5 years 10 yrs 15 yrs	(a) Quantity × frequency: $r = > .50$ (5 years) $r = .26 –.36$ (~10 years) (b) Quantity: $r = > .50$ (5 and 10 years) $r = < .50$ (15 years)
Whitfield et al. (2004)[†]	8184 in 1980 Study (18–88 years) 6570 in 1989 Study 5996 in 1993 Study	Questionnaire in 1980, 1989 and telephone interview (SSAGA[a]) in 1993	~4 years ~9 years ~13 yrs	Quantity × frequency: $r = .66–.69$ (~4 years) $r = .57–.60$ (~9 years) $r = .54–.56$ (~13 years)

Note: [†]includes a subsample of data examined in the current study plus data from earlier collection phases.

[a] SSAGA: Semi-Structured Assessment for the Genetics of Alcoholism (Bucholz et al., 1994)
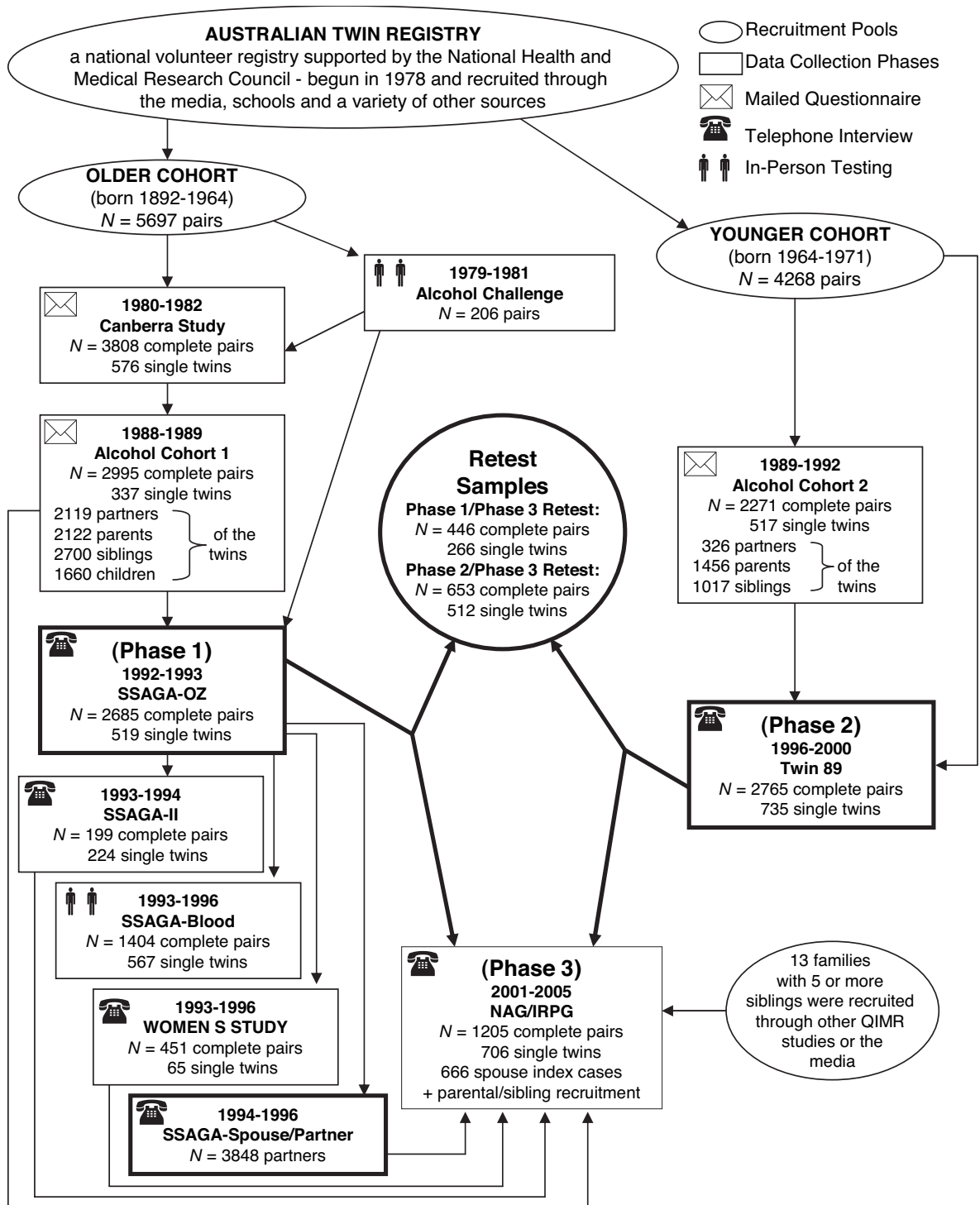
Figure 1

Flowchart showing data sources and data collection phases leading to the present analyses (includes SSAGA-OZ spin-off studies). Twin data collected during Phases 1, 2, and 3 were examined for heritability and retest data were examined for reliability. Retest sample sizes reflect the number of twins with data for one or both of the traits examined. A brief description of each data collection (or study) phase is given in Table 2. Bold boxes are used to identify the primary sources for the Phase 3 sample.

Table 2

Overview of Studies and Data Collection Phases Shown in Figure 1

Study	Type of data collection	Summary	Description
Alcohol Challenge (1979–1981)	♠♠ In-person testing	Blood and urine samples were collected; a questionnaire on drinking and smoking habits was completed; alcohol was ingested (0.75 g ethanol/kg body weight diluted to 10% (v/v) at a steady rate over 20 minutes in a mixture of sugarless lemon and soda water); psychomotor tests were given; ratings of subjective intoxication taken; and 6 blood alcohol levels were taken during and between testing sessions	Two hundred and six twin pairs were recruited from Sydney and Canberra metropolitan areas through the Australian Twin Registry (ATR) for a clinical study of alcohol metabolism and susceptibility to intoxication (Martin et al., 1985; Whitfield & Martin, 1983). To assess repeatability, 87 participants attended a second testing occasion. The sample was not selected on personal or family history of alcoholism, but pairs were excluded if either was a lifetime abstainer (Heath & Martin, 1992; Madden, Heath, Starmer, Whitfield, & Martin, 1995). (See also Birley et al., 2007; Birley et al., 2005; Boomsma, Martin, & Molenaar, 1989; Dickson et al., 2006; Grant et al., 1999; Heath et al., 1999; Heath, Madden, Heath, & Martin, 2002; Heath & Martin, 1991a, 1991b; Heath, Whitfield et al., 2001; Madden, Heath, & Martin, 1997; Martin & Boomsma, 1989; Neale et al., 1999; Neale & Martin, 1989; Slutske et al., 1995; Whitfield & Martin, 1985, 1993, 1994, 1996; Whitfield, Nightingale, Bucholz et al., 1998; Whitfield, Starmer, & Martin, 1990; Whitfield et al., 2001)
Canberra Study (1980–1982)	⊠ Mailed questionnaire	Health and lifestyle questionnaire containing questions about personality, health, and lifestyle, including self-reported drinking habits (a history of alcohol problems was not included)	All ATR twins born 1892 to 1963 (i.e., aged 18 or over) were targeted. The questionnaire was mailed to all 5967 twin pairs and completed replies were obtained from 3808 complete pairs (64% pairwise response rate) and 576 single twins (Heath & Martin, 1994). The sample was over-represented for females and MZ pairs and under-represented for uneducated pairs, but representative of the Australian population for measures of personality, anxiety, depression, and weekly alcohol consumption (Heath & Martin, 1988). (See also Bucholz et al., 1998; Dickson et al., 2006; Heath, Jardine, & Martin, 1989; Heath, Meyer, Eaves, & Martin, 1991a, 1991b; Heath, Meyer, & Martin, 1991; Jardine & Martin, 1984; Martin, 1987, 1988, 1991; Slutske et al., 1995)
Alcohol Cohort 1 (1988–1989)	⊠ Mailed questionnaire	Health and lifestyle questionnaire that included a history of alcohol problems (based on criteria proposed by Feighner (1972))	Complete pairs who had responded in the Canberra Study (i.e. 3808 pairs) were targeted. Twins who did not respond to the mailed questionnaire were offered an abbreviated telephone interview that included alcohol consumption questions, but alcohol problems were not assessed. In total, 2997 complete pairs (79% pairwise response rate) and 335 single twins (83% individual response rate) completed questionnaires or were interviewed (Heath & Martin, 1994; Heath, Slutske et al., 1994). (See also Bucholz et al., 1998; Dickson et al., 2006; Heath et al., 1999; Heath, Todorov et al., 2002; Heath, Whitfield et al., 2001; Madden, Heath, & Martin, 1997; Slutske et al., 2002; Slutske et al., 1996; Slutske et al., 1995; Whitfield et al., 2004; Whitfield & Martin, 1996; Whitfield, Nightingale, Bucholz et al., 1998; Whitfield et al., 2000) For a spin-off study of relatives, participants were asked to provide names of parents, siblings, spouses, and children who would be prepared to answer a similar questionnaire and the names of 14,421 relatives were given. A modified version of the questionnaire was prepared for parents and another version for siblings, spouses, and children of twins. In total, 8601 relatives returned questionnaires (Lake, Eaves, Maes, Heath, & Martin, 2000)

continued over

Table 2 (CONTINUED)

Overview of Studies and Data Collection Phases Shown in Figure 1

Study	Type of data collection	Summary	Description
Alcohol Cohort 2 (1989–1992)	☒ Mailed questionnaire	Health and lifestyle questionnaire that included a history of alcohol problems (based on criteria proposed by Feighner (1972))	The target sample was twins born 1964 to 1971 who were registered with the ATR in 1989, a total of 4268 pairs. These twins were registered as children over the period 1980 to 1982 in response to media appeals and systematic appeals through the Australian school systems. Completed questionnaires were returned by 3740 respondents and a further 1319 individuals completed an abbreviated questionnaire by telephone interview (59.3% individual response rate) (Heath, Howells et al., 2001). For a spin-off study of relatives, participants were asked to provide names of parents, siblings, spouses, and children who would be prepared to answer a similar questionnaire and the names of 4999 relatives were given. A modified version of the questionnaire was prepared for parents and another version for siblings, spouses, and children of twins. In total, 2799 relatives returned questionnaires (Lake, Eaves, Maes, Heath, & Martin, 2000)
(Phase 1) SSAGA-OZ: Semi-Structured Assessment for the Genetics of Alcoholism (1992–1993)	☎ Telephone interview	Adapted from the SSAGA (Bucholz et al., 1994), a diagnostic instrument developed for genetic research on alcoholism that provides diagnostic assessments of DSM-IIIR/DSM-IV alcohol dependence and other major axis I disorders)	Twin pairs were targeted if (a) at least one twin had responded to the Alcohol Cohort 1 survey, or (b) they had participated in the Alcohol Challenge study. Data were obtained from 2685 complete pairs and 519 single twins (Heath et al., 1997). (See also Agrawal et al., 2006; Dickson et al., 2006; Grant et al., 2007; Hansell, Pang, Heath, Martin, & Whitfield, 2005; Heath, Bucholz et al., 1994; Heath et al., 1999; Heath, Slutske, Bucholz, Madden, & Martin, 1997; Heath, Slutske et al., 1994; Heath, Todorov et al., 2002; Heath, Whitfield et al., 2001; Knopik et al., 2006; Madden, Bucholz, Martin, & Heath, 2000; Slutske, Heath et al., 1998; Slutske et al., 2002; Slutske et al., 1996; Slutske et al., 1995; Whitfield et al., 2004; Whitfield, Nightingale, Bucholz et al., 1998; Whitfield, Nightingale, O'Brien et al., 1998; Whitfield et al., 2003; Whitfield et al., 2000.) Spin-off studies focused on (a) blood collection and a drinking diary (SSAGA Blood – targeted DZ pairs where one or both cotwins had not previously been bled) (see Hansell, Pang, Heath, Martin, & Whitfield, 2005; Luciano et al., 2004; Whitfield, Fletcher et al., 1998; Whitfield et al., 2003; Whitfield, Zhu, Heath, Powell, & Martin, 2001), (b) collection of data from spouses/partners (SSAGA Spouse) (see Agrawal et al., 2006; Grant et al., 2007), and longer interviews for (c) pairs where at least one twin was diagnosed with alcohol dependence based on their SSAGA interview (SSAGA-II) (see Madden et al., 1997), and (d) for female pairs where neither twin was diagnosed with alcohol dependence based on their SSAGA interview (Women's Survey) (see Madden et al., 1997).
(Phase 2) Twin 89 (1996–2000)	☎ Telephone interview	Adapted from the SSAGA (Bucholz et al., 1994), a diagnostic instrument developed for genetic research on alcoholism that provides diagnostic assessments of DSM-IIIR/DSM-IV alcohol dependence and other major axis I disorders)	As with Alcohol Cohort 2, the target sample was twins born 1964 to 1971 who were registered with the ATR (2 pairs and 1 single twin born 1972-1975 were also interviewed). Due to the poor response rate for the Alcohol Cohort 2 survey, pairs were approached whether or not they had participated in that study (Heath, Howells et al., 2001). Of 4010 pairs that could be traced, interviews were conducted for 2765 complete pairs and 735 single twins (Nelson et al., 2004; see also Knopik et al., 2004; Lynskey et al., 2005).

continued over

Table 2 (CONTINUED)

Overview of Studies and Data Collection Phases Shown in Figure 1

Study	Type of data collection	Summary	Description
(Phase 3) NAG/IRPG: Nicotine Addiction Genetics/Interactive Research Projects Grant (2001–2005)	☎ Telephone interview	Adapted from the SSAGA (Bucholz et al., 1994), a diagnostic instrument developed for genetic research on alcoholism that provides diagnostic assessments of DSM-IIIR/DSM-IV alcohol dependence and other major axis I disorders)	The NAG and IRPG are linked projects with identical interviews, protocols, and procedures. However, the projects differed in terms of their target samples, with the NAG having a smoking focus and the IRPG having an alcohol focus. The NAG was a collaborative project involving data collection in Australia and Finland, but only the Australian data are examined here. Families were targeted based on heavy smoking index cases identified in previous interview and questionnaire surveys. The IRPG projects comprised (a) a Candidate Gene Study, which was designed as an unmatched case-control study, (b) a Big Sibships Study, designed to study only families with 5 or more offspring sharing both biological parents; and (c) an Extreme Discordant and Concordant Sibships (EDAC) Study, designed to focus on families extremely discordant or concordant for heavy drinking and alcohol dependence risk. The families targeted were drawn from earlier surveys. The Candidate Gene Study specifically targeted twins with a history of alcohol dependence, their co-twins, informative siblings, and unrelated controls. Families with 5 or more siblings were targeted for the Big Sibships Study, regardless of earlier-reported alcohol behaviors. In targeted families, biological parents, spouses of twins, and the siblings of spouses were also approached. In addition, a small number of large sibship families were recruited from studies not related to alcohol use, or through the media. For the EDAC Study, the selection of informative sibships was based upon either a history of alcohol dependence or scores on a quantitative index of alcohol consumption (below the 30th percentile and above the 85th percentile). Targeted sibships were either discordant or high-concordant. Also included in the target sample were additional siblings and biological parents. The Australian NAG/IRPG sample comprised 9688 individuals (1205 complete pairs, 706 single twins, 3586 siblings of twins, 889 parents, 666 spouses of twins, and 1431 siblings of spouses). Ancestry was primarily British and European, as shown in Table 3. Funding was provided by NIDA/NCI for the NAG project and NIAAA for the IRPG projects.

Table 3

Ancestry of the Phase 3 (NAG/IRPG) Sample Showing Number of Participants by Proportions of Ancestors Belonging to Each Ancestral Group

Ancestry	Number of Participants		
	100%	50% to 87.5%	12.5% to 37.5%
British	2456	4030	1441
Northern European	311	1459	2265
Western European	346	451	1574
Southern European	196	166	442
Eastern European	46	98	124
Middle Eastern	13	19`	39
Australian Aboriginal	7	27	129
South American	2	2	9
Asian	0	29	83
Pacific Islander	0	8	60
Native American	0	0	5
Caribbean	0	0	5
African	0	0	4
Central American	0	0	2
Unknown	200	932	1019

Note: Individuals with a proportion of 87.5% or less in an ancestral group will appear in multiple ancestral groups.

similarity and degree to which others could tell co-twins apart. If co-twins gave inconsistent answers, they were followed-up by telephone, and if inconsistency or uncertainty was still apparent, they were asked to send in photographs at various ages, from which a zygosity assignment was made by project staff. Zygosity assignment based on self-report and responses to standard informative questions has been shown to be approximately 97% accurate (Reed et al., 2005). To confirm zygosity assignment, 347 pairs were genotyped at nine independent DNA microsatellite polymorphisms (STR loci: D5S818, D13S317, D3S1358, VWA, D8S1179, D7S820, D21S11, FGA, D18S51) plus the sex marker amelogenin using the Profiler multiplex marker set (AmpFLSTR Profiler Plus, Applied Biosystems, Foster City, CA). The probability of dizygotic twins being concordant for two alleles at the multiple polymorphic loci examined when using this kit, is reported to be less than 10^{-4} (Nyholt, 2006). Analysis of genotype data confirmed the accuracy of our zygosity assignment with a 100% confirmation rate for the subset of pairs examined.

Measures

Quantity × Frequency

A quantity × frequency measure was derived from questions regarding the number of drinks consumed in a typical drinking day and drinking frequency.

Typical day. The telephone interview question at Phase 1 was 'Think of the times when you've had alcohol during the past 12 months. How many drinks do you typically drink on these days when you had an alcoholic drink?'. At Phases 2 and 3, the question was 'In the past 12 months, how many alcoholic drinks would you have on a typical day when you had any alcoholic drinks?'. At Phase 1, the actual number of drinks was recorded. However at Phases 2 and 3, participants were given response choices. Subsequently, all responses were coded into the following categories: 1 = *zero drinks*, 2 = *1–2 drinks*, 3 = *3–4 drinks*, 4 = *5–6 drinks*, 5 = *7–8 drinks*, 6 = *9–11 drinks*, 7 = *12–15 drinks*, 8 = *16–18 drinks*, 9 = *19–24 drinks*, 10 = *25–30 drinks*, 11 = *31 or more drinks*.

Frequency. The telephone interview question regarding drinking frequency was 'During the past 12 months, how often have you had alcoholic drinks?'. Response choices varied slightly between studies. They

Table 4

Demographic Characteristics of the Samples

Sample	N Individuals (Complete Pairs)	Participant age (years)			% Female	% MZ Pairs
		Range	M	SD		
Full sample						
Phase 1 (SSAGA)	5823 (2599)	28—90	43.8	12.2	65	48.8
Phase 2 (Twin 89)	6151 (2698)	23—39	29.9	2.5	55.3	42.5
Phase 3 (NAG/IRPG)	3060 (1177)	19—77	41.1	8.8	55.4	22.7[a]
Retest Samples:						
Phase 1/Phase 3						
Phase 1 (SSAGA)	1158 (446)	28—65	38.9	7.8	58.2	25.1
Phase 3 (NAG/IRPG)	—	38—77	49.9	7.8	—	—
Phase 2/Phase 3						
Phase 2 (Twin 89)	1818 (653)	24—36	30	2.4	54.2	21.4
Phase 3 (NAG/IRPG)	—	29—41	35.5	2.3	—	—

Note: The total number of participants is 12,045, that is, the sum of the full samples (5823 + 6151 + 3060) less the retest samples (1158 + 1818) less 13 individuals participating in both Phases 1 and 2.

[a] This percentage is reduced, as the primary aim for this sample was linkage analyses and MZ pairs were only targeted if they had additional siblings.

were recoded into the following six categories that are common to both the telephone interview and earlier mailed questionnaire studies (6 = *at least once daily*, 5 = *3–6 days per week*, 4 = *1–2 days per week*, 3 = *1–3 days per month*, 2 = *less often*, 1 = *never*).

The typical day categories were recoded to reflect the number of drinks per day. Similarly, the frequency categories were recoded to reflect a 'times per week' measure. Thus the quantity × frequency measure reflected the number of drinks per week.

Dependence Score

Assessment was based on lifetime prevalence. A dependence score was obtained from items based on Diagnostic and Statistical Manual of Mental Disorders, third edition, revised (DSM-III-R) criteria (American Psychiatric Association, 1987) and Diagnostic and Statistical Manual of Mental Disorders, fourth edition (DSM-IV) criteria (American Psychiatric Association, 2000) for substance dependence, which were combined as follows:

1. substance often taken in larger amounts or over a longer period than the person intended

2. persistent desire or one or more unsuccessful efforts to cut down or control substance abuse

3. a great deal of time spent in activities necessary to get the substance or recover from its effects

4. frequent intoxication or withdrawal symptoms when expected to fulfill major role obligations at work, school, or home, or when substance use is physically hazardous

5. important social, occupational, or recreational activities given up or reduced because of substance use

6. continued substance use despite knowledge of having a persistent or recurrent social, psychological, or physical problem that is caused or exacerbated by the use of the substance

7. tolerance, as defined by either of the following:

    (a) a need for markedly increased amounts of the substance to achieve intoxication or desired effect

    (b) markedly diminished effect with continued use of the same amount of the substance

8. characteristic withdrawal symptoms

9. substance often taken to relieve or avoid withdrawal symptoms

All items, with the exception of item 7, were coded as 0 for no and 1 for yes. Item 7 was coded as a three-level ordinal measure with 0 for no, 1 for a marked (but less than 50%) increase in the amount required to achieve the desired effect, and 2 for at least a 50% increase in the amount required to achieve the desired effect. To obtain a single dependence score, items were summed. Note that individuals participating in Phases 2 and 3 were not re-asked dependence-related questions at Phase 3.

Analyses

Descriptive statistics were performed using SPSS 13.0 for Windows. Quantity × frequency and dependence scores were treated as continuous variables. In addition, variables were treated as categorical data in confirmatory reliability analyses as data (particularly dependence score) were not normally distributed. Both quantity × frequency and dependence score were positively skewed. Quantity × frequency was log transformed [$\log_{10}(x + 1)$ where x is the number of drinks per week] as has been done previously (Jardine & Martin, 1984)). However, transformation did not improve the dependence score distribution and it was consequently left untransformed. Distributions for the full and retest samples are shown in Figure 2.

Raw data on quantity × frequency and dependence score were used to regress out the effects of important covariates, including tests for sex, age at interview, quadratic effects of age (age^2), sex*age and sex*age^2. PROC REGRESS in SAS was used to regress out the effects of covariates that were significant at $p < .05$ in Phases 1, 2, and 3 separately and residuals stemming from this regression were used for both the stability and heritability calculations.

All modeling was performed using raw data in the statistical software package Mx (Neale, Boker, Xie, & Maes, 2003) using a full information maximum likelihood (FIML) estimator. The FIML method used in Mx is especially useful in handling missing data. Estimates are unbiased when data are missing completely at random and/or missing at random is predicted by other variables used in the analysis (Little & Rubin, 2000). To assess model fit, the fit of constrained models was compared to the full model by examining the difference in the –2 log likelihood, which is distributed as a chi-square for given degrees of freedom.

Stability

The following analyses were conducted for quantity × frequency and dependence score for the Phase 1/Phase 3 retest sample and for quantity × frequency for the Phase 2/Phase 3 retest sample. Residuals, as computed from the full samples, were examined. Submodels were compared to the full model, with the submodels for continuous data being:

• whether mean scores could be equated across phases for each retest sample

• whether variances could be equated across phases for each retest sample

• test–retest correlations with 95% confidence intervals

and for categorical data:

• whether thresholds could be equated across phases for each retest sample (note that each variable was divided into three categories of approximately equal size based on the first interview occasion responses)
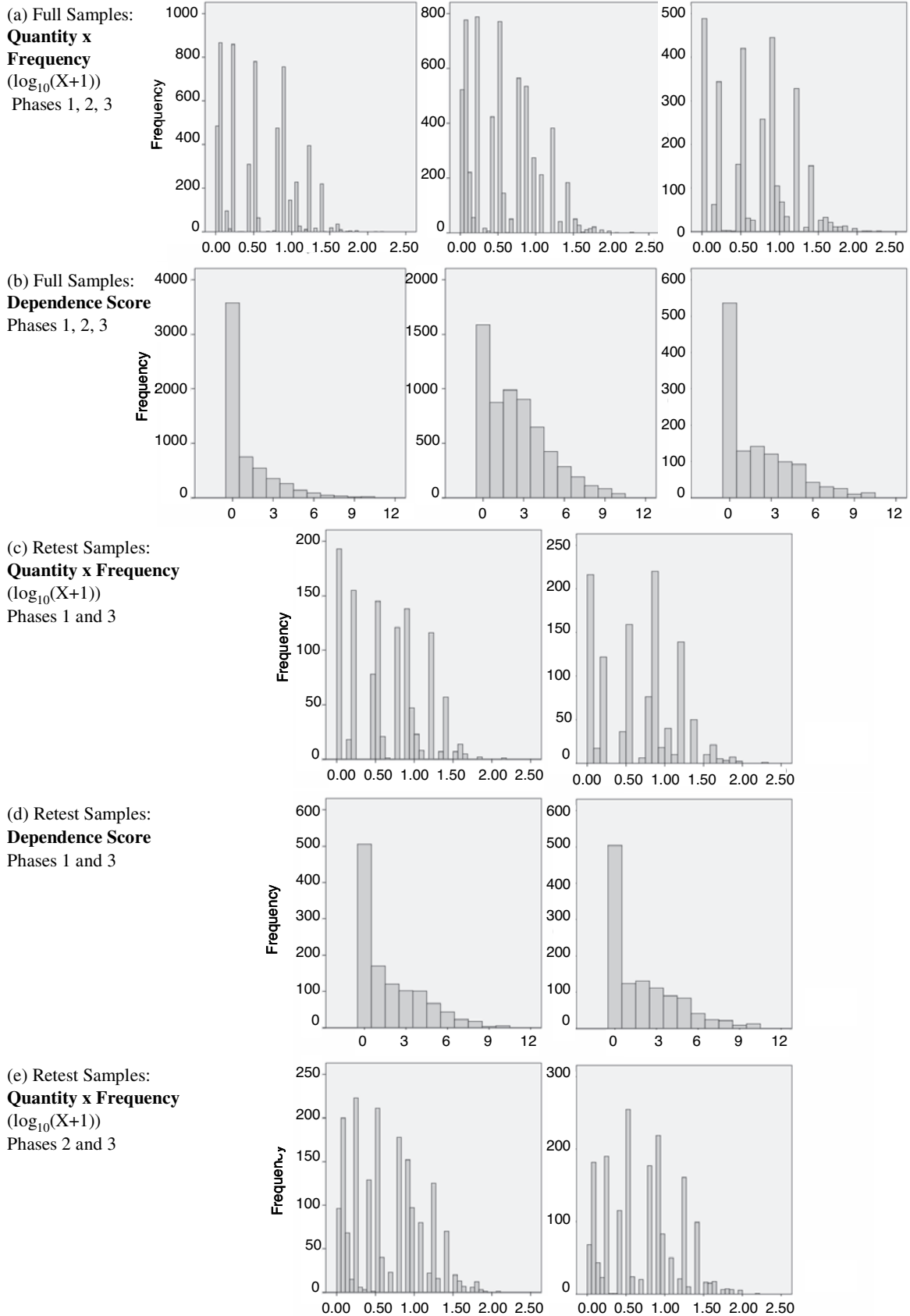
(a) Full Samples:
Quantity x Frequency
$(\log_{10}(X+1))$
Phases 1, 2, 3

(b) Full Samples:
Dependence Score
Phases 1, 2, 3

(c) Retest Samples:
Quantity x Frequency
$(\log_{10}(X+1))$
Phases 1 and 3

(d) Retest Samples:
Dependence Score
Phases 1 and 3

(e) Retest Samples:
Quantity x Frequency
$(\log_{10}(X+1))$
Phases 2 and 3

Figure 2

Distributions for quantity × frequency (transformed) and dependence score for the full samples (a) and (b), and the retest samples (c), (d), and (e).

- test–retest correlations with 95% confidence intervals.

Heritability

The extent to which additive genetic (A), shared environmental (C) and nonshared environmental (E) factors influenced population differences in quantity × frequency and the dependence score was examined. Univariate sex-equated models were fit to regression residuals for Phase 1, Phase 2, and Phase 3 data separately. In addition, models comparing quantitative differences in estimates of heritable and environmental influences on the full Phase 1 or Phase 2 samples and the subset of twins targeted for Phase 3, the latter of which was oversampled for smoking/drinking, were examined. The following submodels were tested against the full model (note that dependence score was not available at Phase 3 if collected at Phase 2, and thus a Phase 2/Phase 3 comparison could not be made):

- whether mean scores could be equated across the full Phase 1/Phase 2 sample and their respective Phase 3 components

- whether total variance could be equated across the full Phase 1/Phase 2 sample and their respective Phase 3 components.

- if the total variance could be equated, whether or not additive genetic, shared environmental and nonshared environmental influences could be constrained to be equal across Phase 1/Phase 2 and their respective Phase 3 components was examined. However, if there was statistical evidence in favor of differences in total variance, we used a scalar twin model where the total variance in Phase 3 was modeled as a scalar multiple of the variance in Phase 1/Phase 2. Note that as heritability reflects a proportion of total variance (A/A + C + E), its interpretation is determined by the magnitude of the denominator. Therefore, for instance, if variance was estimated as X in Phase 2 and differently, as Y, in Phase 3, then A/X could not be statistically equated to A/Y as a test for equality of heritable factors. Hence, we used a scalar product (or Y = X*n, where n ranges from −∞ to +∞; if n = 1 then variances are equal) for instances in which variances could not be constrained.

- tests for the statistical significance of A and C were also conducted.

Further, bivariate Cholesky triangular decomposition models of variance (Neale, Boker, Xie, & Maes, 2003) were fit to the full Phase 1/Phase 3 and full Phase 2/Phase 3 datasets to obtain genetic and environmental correlations between phases. Although only a subset of retest data was available, including the full datasets results in more stable estimates of sample characteristics.

Results

Sample Characteristics

Raw data means are shown in Table 5. Overall, the means for quantity × frequency indicate that weekly consumption is similar between the older (Phase 1) and younger (Phase 2) cohorts. In addition, there is a trend for increased consumption from the earlier collection Phases (1 and 2) to the later collection Phase (3). Means for dependence score show that, at each collection stage, males satisfied a larger number of criteria than females. In addition, the younger cohort (Phase 2) satisfied a larger number of criteria than the older cohort (Phase 1) or the combined cohort (Phase 3). Significance of covariates is shown in Table 6. Their effects were regressed from the variables of interest and residuals were examined for all remaining analyses.

Table 5

Raw Data Means for Alcohol Variables

Sample	Quantity × frequency		Dependence score	
	Males	Females	Males	Females
Full Sample				
Phase 1 (SSAGA)	7.7	3.6	1.8	0.6
Phase 2 (Twin 89)	7.9	3.7	3.2	2
Phase 3 (NAG/IRPG)	10.6	5.4	2.9	1.4
Retest Sample				
Phase 1/Phase 3				
Phase 1 (SSAGA)	9.9	4.2	2.8	1.1
Phase 3 (NAG/IRPG)	11.0	5.5	2.9	1.3
Phase 2/Phase 3				
Phase 2 (Twin 89)	9.2	4.4	4	2.9
Phase 3 (NAG/IRPG)[a]	10.4	5.3	—	—

Note: [a]Dependence-related questions were not asked at Phase 3 if already asked at Phase 2.

Table 6

Significant Covariates by Alcohol Variable and Collection Phase

Significant covariates	β	SE	p
Quantity × frequency			
Phase 1: Sex	0.2551	0.0115	< .0001
Age	0.008	0.0032	.0131
Age^2	−0.0001	0	.0017
Phase 2: Sex	−1.0243	0.3896	.0086
Sex*Age	−0.8521	0.2613	.0011
Sex*Age^2	0.0139	0.0044	.0014
Phase 3: Sex	0.2522	0.0164	< .0001
Age	0.0184	0.0079	.0208
Age^2	−0.0002	0.0001	.0155
Dependence score			
Phase 1: Sex	0.908	0.1916	< .0001
Age^2	−0.0006	0.0001	< .0001
Sex*Age	−0.2516	0.0718	.0005
Sex*Age^2	0.0042	0.0008	< .0001
Phase 2: Sex	1.1982	0.0574	< .0001
Phase 3: Sex	−7.4612	1.4288	< .0001
Age	1.0499	0.1061	< .0001
Age^2	−0.0095	0.0011	< .0001
Sex*Age	−3.1639	0.6322	< .0001
Sex*Age^2	0.0276	0.0067	< .0001

Analyses of Means, Variances, and Thresholds

When examined as continuous data, quantity × frequency means (of residuals) could be equated across phases for both the Phase 1/Phase 3 and Phase 2/Phase 3 retest samples ($\Delta\chi^2_1$ ranged 0.0–2.0, $p > .1$). Variances could be equated for the Phase 2/Phase 3 retest sample ($\Delta\chi^2_1$ 0.8, $p > .3$), but not for the Phase 1/Phase 3 retest sample ($\Delta\chi^2_1$ 13.7, $p < .001$). In addition, means for dependence score could be set equal across phase for the Phase 1/Phase 3 retest sample ($\Delta\chi^2_1$ 0.3, $p > .5$). However, variances could not be equated for dependence score for the Phase 1/Phase 3 retest sample ($\Delta\chi^2_1$ 29.7, $p < .001$).

When examined as categorical data, thresholds could not be set equal across interview occasions for any variable ($\Delta\chi^2_2$ ranged 7.8–63.6, $p < .03$). For retest samples on first interview occasions, thresholds were set to divide data into three equally-sized categories (low, intermediate, high). However, for quantity × frequency, at Phase 3 for the Phase1/Phase 3 retest sample the proportion of respondents in the upper and lower tails increased to 37% in each case. Thus, at Phase 3 compared to Phase 1, there were fewer individuals drinking at an intermediate level and a larger number of individuals drinking at both low and high levels. At Phase 3 for the Phase 2/Phase 3 retest sample, 36% of respondents were in the lower tail and 31% in the upper tail. Therefore, on the second interview occasion there was the same proportion of intermediate level drinkers, but more individuals were

drinking at a low level and fewer at a high level. For dependence score, at Phase 3 for the Phase 1/Phase 3 retest sample, 48% of respondents were in the lower tail and 31% in the upper tail. This indicates a decrease over time in the number of individuals reporting either an intermediate or a high number of dependence symptoms and a corresponding increase in the number of individuals reporting a low number of symptoms. Indeed, 304 individuals (26.3% of the retest sample) reported fewer symptoms at Phase 3 compared to Phase 1. Of the remaining participants, 141 had at least one symptom reported at Phase 1 that was not reported again at Phase 3. Therefore, 38.5% of the retest sample was inconsistent in the reporting of at least one 'lifetime' symptom of alcohol dependence

Prior to the heritability analyses: Means for quantity × frequency could be constrained to be equal across the full Phase 1 and corresponding Phase 3 retest samples, as well as between full Phase 2 and Phase 3 retest samples ($\Delta\chi^2_1$ ranged 0.0–0.8, $p > .37$). Variances could be equated across Phase 2 and Phase 3 ($\Delta\chi^2_1$ 0.2, $p = .65$) but not across Phase 1 and Phase 3 ($\Delta\chi^2_1$ 11.8, $p < .001$). For dependence score, there were statistically significant differences in both the means ($\Delta\chi^2_1$ 25.5, $p < .001$) and the variances ($\Delta\chi^2_1$ 287.3, $p < .001$) when comparing Phase 1 with Phase 3. This contrasts with the results reported as preliminary to the stability analyses, where means could be set equal. It is surmised that there is a mean difference between dependence score at Phases 1 and 3, but that this difference is not statistically significant in the smaller retest samples (i.e., when the Phase 1 sample is restricted to individuals retested).

Stability

Test–retest estimates and 95% confidence intervals are shown in Table 7 for the data examined as continuous and as categorical. Estimates are similar, ranging from .55 to .61 when examined as continuous and .58 to .64 when examined as categorical.

Heritability

Quantity × frequency. In the full Phase 1, 2 and 3 samples, heritable factors explained between 31% and

Table 7

Test–Retest Correlations Showing the Stability of Quantity × Frequency and Dependence Score Over Mean Intervals of 5.5 Years (Phase 2/Phase 3) and/or 11.0 Years (Phase 1/Phase 3)

	r (95% confidence interval)	
Sample	Continuous data	Categorical data
Phase 1/Phase 3		
Quantity × frequency	.61 (.57, .65)	.64 (.59, .69)
Dependence score	.61 (.57, .64)	.58 (.52, .63)
Phase 2/phase 3[a]		
Quantity × frequency	.55 (.51, .58)	.59 (.55, .59)

Note: [a]Dependence-related questions were not asked at Phase 3 if previously asked at Phase 2.

Table 8

Means and Total Variances for the Residuals plus Additive Genetic (A) and Nonshared Environmental (E) Influences with 95% Confidence Intervals for Quantity × Frequency and Dependence Score

Sample	Mean (residuals)	Total variance	A (95% CI)	E (95% CI)
Full sample				
Phase 1 (SSAGA)				
Quantity × frequency	0	0.18	0.47 (0.43, 0.51)	0.53 (0.49, 0.57)
Dependence Score	0.003	2.64	0.46 (0.42, 0.51)	0.54 (0.49, 0.58)
Phase 2 (Twin 89)				
Quantity × frequency	0.002	0.18	0.31 (0.23, 0.36)	0.69 (0.64, 0.77)
Dependence score	0.017	5.32	0.45 (0.37, 0.49)	0.55 (0.51, 0.63)
Phase 3 (NAG/IRPG)				
Quantity × frequency	0.001	0.20	0.40 (0.24, 0.48)	0.60 (0.52, 0.76)
Dependence score	0.006	3.03	0.42 (0.26, 0.50)	0.58 (0.50, 0.74)
Combined full/retest (Phase 1/Phase 3)				
Phase 3 retest				
Quantity × frequency	0.002	0.21	0.51 (0.47, 0.54)	0.49 (0.46, 0.53)
Dependence Score	0.377	5.21	0.51 (0.48, 0.55)	0.49 (0.45, 0.52)
(Phase 2/Phase 3ª)				
Phase 3 Retest				
Quantity × frequency	0.001	0.19	0.43 (0.39, 0.53)	0.57 (0.54, 0.61)

Note: ªDependence-related questions were not asked at Phase 3 if previously asked at Phase 2, therefore there is no retest sample for dependence score.

47% of the total variance (Table 8) with the magnitude of heritable influences being somewhat higher in the older versus younger cohorts of twins. In Phase 1 and its component Phase 3, population variance could be attributed to additive genetic (51%) and nonshared environmental (49%) influences. Similarly, in Phase 2 and its component Phase 3, an additive genetic factor accounted for 43% of the total variance, with the remainder due to nonshared environmental influences.

Dependence score. Additive genetic (42–46%) and nonshared environmental factors influenced population variation in dependence score across the three full samples. However, in contrast to quantity × frequency, a statistically significant difference between the variances for Phase 1 and its component Phase 3 was found. Consequently, scalar models were pursued in this comparison. After accounting for these differences in sample variance, heritable factors explained about an equal proportion of this variance (51%) in each sample.

In each of the aforementioned models, additive genetic and nonshared environmental influences were sufficient to explain the variance (Table 8). The influence of shared environmental factors could be dropped in each case without loss of fit ($\Delta\chi^2_1$ ranged 0.0–2.4, $p > .1$). In the fully saturated models, shared environment accounted for 0.06%, 0%, and 0% of the variance of quantity × frequency at Phases 1, 2, and 3 respectively (95% confidence intervals were .0–.13, .0–.08, .0–.10). For dependence score, shared

environment accounted for 0% of the variance at all phases (confidence intervals were 0–.02, .0–.06, and 0–.10 for Phases 1, 2, and 3 respectively).

For quantity × frequency, genetic influence was found to be highly stable across the 5.5 and 11.0 year intervals examined (r_g = .96 for both intervals). In addition, the stability of genetic influence over 11.0 years was moderately strong for dependence score (r_g = .62). However, nonshared environmental influences were less stable (r_e = .33 for quantity × frequency over 5.5 years and .34 over 11.0 years, r_e = .40 for dependence score over 11.0 years).

Discussion

The aims of the present article were first, to examine the long-term stability of alcohol dependence symptoms and a quantity × frequency measure of consumption, and second, to examine the heritability of these measures for different collection phases. The results, which are outlined more fully in following paragraphs, indicate that (a) both measures have moderate long-term stability, (b) our dependence measure may not reflect a true 'lifetime' measure, (c) both measures are moderately heritable, (d) quantity × frequency appears less heritable in a younger versus older cohort, (e) dependence appears to be less heritable when examined as symptom count than as a yes/no dichotomous measure, (f) selection for drinkers and smokers did not significantly affect the heritability of the measures, and (g) genetic influence

was highly stable over the intervals examined, particularly for consumption.

In the current paper, the stability of alcohol dependence symptoms and a quantity × frequency measure were examined in an older cohort (ages ranging 28–65 years at baseline) over a mean interval of 11.0 years, and for quantity × frequency only, in a younger cohort (ages ranging 24–36 years at baseline) over a mean interval of 5.5 years. For both measures and intervals, moderate stability is found with correlations ranging from .55 to .64. Typically, stability decreases slowly over time (e.g., Pape & Hammer, 1996; Whitfield et al., 2004). However, in the present case stability is higher, although not significantly, over the longer compared to the shorter interval (.61–.64 vs. .55-.59 for quantity × frequency). This may be due to the younger age of participants examined for the shorter interval (41% vs. 2% < 30 years). It has previously been found that the stability of alcohol consumption is lower in younger samples (Kerr, Fillmore, & Bostrom, 2002) and that level of consumption and drinking problems in adolescents and young adults is not highly predictive of later drinking habits and problems (Fillmore et al., 1991).

As expected, the long-term stability estimates are considerably lower than reports of test–retest reliability (.58–.62 vs. .84–.90, (Bucholz et al., 1994), for alcohol dependence symptoms, and .55–.64 vs. .73–.91, (Friesema et al., 2004; Webb, Redman, Gibberd, & Sanson-Fisher, 1991) for quantity × frequency) indicating changes in drinking habits over time. The stability of dependence symptoms over 11 years is similar to that reported by Slutske, True et al., (1998) for an interval of only 15 months (.58–.61 vs. .61). Further, it is consistent with the general finding of degradation of stability over time, as stability was found to be .77 for dependence symptomatology in a subset of females from the same sample retested after an interval of approximately 2 years (Heath, Bucholz et al., 1997).

Interestingly, close examination of the dependence symptomatology retest data shows that over a third of participants reported at least one symptom at the first interview, which was not subsequently reported at the second interview. However, the interview questions are based on a lifetime diagnosis, that is, they asked if a symptom had ever been experienced rather than if a symptom is currently being experienced or was recently experienced. One explanation for this is that many participants may have responded positively only for symptoms that had been *recently* experienced, rather than *ever* experienced. Consequently, the measure of dependence symptomatology examined does not appear to be a true measure of lifetime dependence. Such a finding is not unique to the present study. In a sample of Vietnam era veterans, Slutske, True et al. (1998) found that 13% of their sample had a lifetime alcohol dependence diagnosis at time one, but not 15 months later at time two.

Stability estimates for the longer interval for quantity × frequency are similar to those based on the Phase 1 telephone interview and earlier questionnaire surveys from the same sample and reported by Whitfield et al. (2004); .61–.64 over ~11 years vs .57–.60 over ~9 years and .54-.56 over ~13 years). These estimates are higher than those reported by Kerr et al. (2002) for quantity × frequency measures collected by questionnaire over a 10 year interval (.26–.36). However, current estimates for the shorter interval are consistent with estimates reported for a 5-year interval by Kerr et al. (2002); .55–.59 vs .50). Pape and Hammer (1996) reported lower estimates (.35–.54 for 2–6 year intervals), but they examined a younger sample (aged 19–22 at baseline).

As mentioned previously, in addition to having moderate long-term stability, alcohol dependence symptomatology and quantity × frequency measures both appear to have high test–retest reliability in the short term(Bucholz et al., 1994; Gerstel, Harford, & Pautler, 1980) (Friesema et al., 2004; Webb, Redman, Gibberd, & Sanson-Fisher, 1991). This is an important factor for genetic studies because measurement reliability has a direct impact on the heritability values estimated. The greater the error variance in the measures, the less the measures reflect the true level of individual difference, and the smaller the upper bound of the heritability estimate. Although reliability was not directly investigated in this sample, reports of high reliability for the measures examined suggest a high upper bound for the heritability estimate.

Moderate heritabilities are found for both measures (.42–.51 for dependence symptomatology, .31–.51 for quantity × frequency). All remaining variance is due to environmental influences unique to the individual. The heritabilities found for dependence symptoms are at the lower end of those generally reported. For example, estimates in the range of .50 to .60 were reported for dependence in a review by Dick and Bierut (2006). Similarly, heritabilities ranging .45 to .65 were reported in a review by Heath, Slutske et al. (1994) for alcoholism in community samples. However, it is notable that a heritability of .64 was found by Heath, Bucholz et al. (1997) in an earlier study of the Phase 1 sample, compared to .46 found in the current study. The major difference between the studies is in how dependence was quantified, which subsequently influenced how it was analyzed. In the Heath, Bucholz et al. (1997) study, dependence was quantified as a dichotomous variable (based on the DSM-IIIR criteria, whereby individuals with three or more symptoms were diagnosed as being dependent) and analyzed using a threshold model. In the present study, symptoms based on the DSM-IIIR and DSM-IV criteria are summed and analyzed as a continuous

variable. These two approaches differ in power, with analysis of continuous data having considerably greater power than analysis of threshold data (Neale, Eaves, & Kendler, 1994). For the Phase 2 sample, similar measurement and methodological differences between the current and an earlier study did not differentiate heritability estimates for dependence symtomatology, with estimates of .45 found in the current study and .47 in the earlier study (Knopik et al., 2004).

Heritability estimates for the quantity × frequency measure fall within the general range of heritabilities reported for alcohol consumption patterns (i.e., .30–.60, Heath & Martin, 1994). As expected, it is similar to that reported previously for data collected from the same sample at Phase 1 and in earlier questionnaire surveys (.40-.54). Of note is the finding of significantly lower heritability in the younger compared to the older cohort (.31 vs. .47). Other studies have hinted at increasing heritability with age for alcohol consumption in adolescents and young adults, but a recent meta-analysis by Bergen et al. (2007) did not find differences to be significant. The difference found in the current analyses may reflect a decreasing environmental influence from young adulthood to middle age. Heritabilities for alcohol consumption appear to be consistent through middle to older age. For example, heritabilities ranging .44 to .47 over a 14- to 18-year period have been reported for older adult males (Reed et al., 1994).

The source of genetic influence on the measures of interest appears to be highly stable over the two intervals examined. For quantity × frequency, almost all of the genes influential at baseline were still influential after intervals of 5.5 or 11.0 years. In addition, most of the genes influencing the dependence symptoms score at baseline were still influential after an 11-year interval. In contrast, the influence of nonshared environmental factors was considerably less stable over time. However, this may be partly due to measurement error, which is included in the estimate of nonshared environmental influence, and which would have a negative impact on stability.

A number of limitations are evident for this study. First, the alcohol phenotypes are based on self-report, which may be associated with problems related to recall, willingness to impart information, or correctly interpreting the question. Nevertheless, self-report measures of alcohol consumption are reported to demonstrate reasonable levels of reliability and validity and have the advantage of being relatively inexpensive, noninvasive, and acceptable to respondents (Del Boca & Darkes, 2003). Second, the data were collected for reasons other than to examine stability, and consequently, in terms of a stability study, the data for interval length are confounded by cohort age. Thus, a clear comparison of

stability for interval length cannot be made. Third, the data, particularly for alcohol dependence symptoms, are not normally distributed, but are analyzed as continuous data. This may bias results. However, to address this limitation threshold models were also used to examine data stability and no significant differences were found. Lastly, genetic modeling shows that the variance for both measures is strongly influenced by nonshared environmental factors, but the nature of these influences is not assessed in this study.

The results of the present study will inform decisions made regarding the structure of a combined sample for future linkage and association studies. These decisions include choice of data to analyze when data were collected on two occasions. Due to significantly lower heritability found for alcohol consumption in the younger cohort compared to the older cohort, choosing data collected at Phase 3 rather than at Phases 1 or 2 may maximize power. Further, the heritability estimates derived in the current analyses will be informative regarding the specification of heritability in future quantitative linkage analyses.

More generally, the results suggest that many individuals may not accurately self-report lifetime symptoms of alcohol dependence. Rather, their responses may be disproportionately influenced by current or recent symptoms. Second, oversampling for smokers and drinkers did not significantly influence heritability for the alcohol traits examined. However, heritability did vary significantly for cohort, suggesting that the influence of an individual's unique life experiences on patterns of alcohol consumption may decrease as they move from young adulthood to middle age. Last, the influence of genetic factors was highly stable over the intervals examined.

References

Agrawal, A., Heath, A. C., Grant, J. D., Pergadia, M. L., Statham, D. J., Bucholz, K. K., Martin, N. G., &

Madden, P. A. F. (2006). Assortative mating for cigarette smoking and alcohol consumption in female Australian twins and their spouses. *Behavior Genetics, 36*, 553–566.

American Psychiatric Association. (1987). *Diagnostic and Statistical Manual of Mental Disorders* (Third Edition, Revised). Washington, DC: American Psychiatric Association.

American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of Mental Disorders* (Fourth Edition, Text Revision). Washington, DC: American Psychiatric Association.

Bergen, S. E., Gardner, C. O., & Kendler, K. S. (2007). Age-related changes in heritability of behavioral phenotypes over adolescence and young adulthood: A meta-analysis. *Twin Research and Human Genetics, 10*(3), 423–433.

Birley, A. J., James, M. R., Dickson, P. A., Montgomery, G. W., Heath, A. C., Whitfield, J. B., & Martin, N. G. (2007). Association of the grastric alcohol dehydrogenase gene ADH7 with variation in alcohol metabolism. *Human Molecular Genetics, 17*(2), 179–189.

Birley, A. J., Whitfield, J. B., Neale, M. C., Duffy, D. L., Heath, A. C., Boomsma, D. I., & Martin, N. G. (2005). Genetic time-series analysis identifies a major QTL for in vivo alcohol metabolism in the ADH region not predicted by in vitro studies of structural protein polymorphism at the ADH1B or ADH1C loci. *Behavior Genetics, 35*, 509–524.

Boomsma, D. I., Martin, N. G., & Molenaar, P. C. M. (1989). Factor and simplex models for repeated measures: Application to two psychomotor measures of alcohol sensitivity in twins. *Behavior Genetics, 19*, 79–96.

Bucholz, K. K., Cadoret, R., Cloninger, C. R., Dinwiddie, S. H., Hesselbrock, V. M., Nurnberger Jr., J. I., Reich, T., Schmidt, I., & Schuckit, M. A. (1994). A new semi-structured psychiatric interview for use in genetic linkage studies: A report on the reliability of the SSAGA. *Journal of Studies on Alcohol, 55*, 149–158.

Bucholz, K. K., Heath, A. C., Madden, P. A. F., Slutske, W. S., Statham, D. J., Dunne, M. P., & Martin, N. G. (1998). Drinking in an older population: cross-sectional and longitudinal data from the Australian Twin Registry. In E. S. L. Gomberg, A. M. Hegedus & R. A. Zucker (Eds.), *Alcohol problems and aging* (Vol. NIAAA Research Monograph 33, pp. 41–62): USDHHS Publ.

Bucholz, K. K., Heath, A. C., Reich, T., Hesselbrock, V. M., Kramer, J. R., Nurnberger Jr, J. I., & Schuckit, M. A. (1996). Can we subtype alcoholism? A latent class analysis of data from relatives of alcoholics in a multicenter family study of alcoholism. *Alcoholism, Clinical and Experimental Research, 20*(8), 1462–1471.

Bucholz, K. K., Hesselbrock, V. M., Shayka, J. J., Nurnberger Jr., J. I., Schuckit, M. A., Schmidt, I., &

Reich, T. (1995). Reliability of individual diagnostic criterion items for psychoactive substance dependence and the impact on diagnosis. *Journal of Studies on Alcohol, 56*, 500–505.

Del Boca, F. K., & Darkes, J. (2003). The validity of self-reports of alcohol consumption: state of the science and challenges for research. *Addiction, 98*(Suppl. 2), 1–12.

Dick, D. M., & Bierut, L. J. (2006). The genetics of alcohol dependence. *Current Psychiatry Reports, 8*(2), 151–157.

Dickson, P. A., James, M. R., Heath, A. C., Montgomery, G. W., Martin, N. G., Whitfield, J. B., & Birley, A. J. (2006). Effects of variation at the ALDH2 locus on alcohol metabolism, sensitivity, consumption and dependence in Europeans. *Alcoholism, Clinical and Experimental Research, 30*, 1093–1100.

Feighner, J. P., Robins, E., Guze, S. B., Woodruff, R. A., Winokur, G., & Munoz, R. (1972). Diagnostic criteria for use in psychiatric research. *Archives of General Psychiatry, 26*, 57–63.

Fillmore, K. M., Hartka, E., Johnstone, B. M., Leino, E. V., Motoyoshi, M., & Temple, M. T. (1991). A meta-analysis of life course variation in drinking. *British Journal of Addiction, 86*(10), 1221–1268.

Friesema, I. H. M., Veenstra, M. Y., Zwietering, P. J., Knottnerus, J. A., Garretsen, H. F. L., & Lemmens, P. H. H. M. (2004). Measurement of lifetime alcohol intake: Utility of a self-administered questionnaire. *American Journal of Epidemiology, 159*(8), 809–817.

Gerstel, E. K., Harford, T. C., & Pautler, C. (1980). The reliability of drinking estimates obtained with two data collection methods. *Journal of Studies on Alcohol, 41*, 89–94.

Grant, B. F., Harford, T. C., Dawson, D. A., Chou, P. S., & Pickering, R. P. (1995). The alcohol use disorder and associated disabilities interview schedule (AUDADIS): Reliability of alcohol and drug modules in a general population sample. *Drug and Alcohol Dependence, 39*, 37–44.

Grant, J. D., Heath, A. C., Bucholz, K. K., Madden, P. A. F., Agrawal, A., Statham, D. J., & Martin, N. G. (2007). Spousal concordance for alcohol dependence: Evidence for assortative mating or spousal interaction effects? *Alcoholism, Clinical and Experimental Research, 31*(5), 717–728.

Grant, J. D., Heath, A. C., Madden, P. A. F., Bucholz, K. K., Whitfield, J. B., & Martin, N. G. (1999). An assessment of the relationship between alcohol metabolism and alcoholism risk in Europeans. *Behavior Genetics, 29*, 463–472.

Gruenewald, P. J., & Johnson, F. W. (2006). The stability and reliability of self-reported drinking measures. *Journal of Studies on Alcohol, 67*, 738–745.

Hansell, N. K., Pang, D., Heath, A. C., Martin, N. G., & Whitfield, J. B. (2005). Erythrocyte aldehyde dehydrogenase activity: Lack of association with alcohol use

and dependence of alcohol reactions in Australian twins. *Alcohol & Alcoholism, 40*(5), 343–348.

Heath, A. C. (1995). Genetic influences on drinking behavior in humans. In H. Begleiter & B. Kissin (Eds.), *The genetics of drinking behavior in humans.* New York: Oxford University Press.

Heath, A. C., Bucholz, K. K., Madden, P. A. F., Dinwiddie, S. H., Slutske, W. S., Bierut, L. J., Stratham, D. J., Dunne, M. P., Whitfield, J. B., & Martin, N. G. (1997). Genetic and environmental contributions to alcohol dependence risk in a national twin sample: consistency of findings in women and men. *Psychological Medicine, 27,* 1381–1396.

Heath, A. C., Bucholz, K. K., Slutske, W. S., Madden, P. A., Dinwiddie, S. H., Dunne, M. P., Statham, D. B., Whitfield, J. B., Martin, N. G., & Eaves, L. J. (1994). The assessment of alcoholism in surveys of the general community: What are we measuring? Some insights from the Australian twin panel interview survey. *International Review of Psychiatry, 6,* 295–307.

Heath, A. C., Howells, W., Kirk, K. M., Madden, P. A. F., Bucholz, K. K., Nelson, E. C., Slutske, W. S., Statham, D. J., & Martin, N. G. (2001). Predictors of nonresponse to a questionnaire survey of a volunteer twin panel: Findings from the Australian 1989 Twin Cohort. *Twin Research, 4*(2), 73–80.

Heath, A. C., Jardine, R., & Martin, N. G. (1989). Interactive effects of genotype and social environment on alcohol consumption in female twins. *Journal of Studies on Alcohol, 50,* 38–48.

Heath, A. C., Madden, P. A., Bucholz, K. K., Dinwiddie, S. H., Slutske, W. S., Bierut, L. J., Rohrbaugh, J. W., Statham, D. J., Dunne, M. P., Whitfield, J. B., & Martin, J. B. (1999). Genetic differences in alcohol sensitivity and the inheritance of alcoholism risk. *Psychological Medicine, 29,* 1069–1081.

Heath, A. C., Madden, P. A. F., Bucholz, K. K., Nelson, E. C., Todorov, A., Price, R. K., Whitfield, J. B., & Martin, N. G. (2002). Genetic and genotype ↔ environment interaction effects on risk of dependence on alcohol, tobacco, and other drugs: New research challenges. In R. Plomin, J. C. DeFries, I. W. Craig & P. McGuffin (Eds.), *Behavioral genetics in the postgenomic era* (pp. 309–334). Washington DC: APA Books.

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.*

Heath, A. C., & Martin, N. G. (1991a). The inheritance of alcohol sensitivity and of patterns of alcohol use. *Alcohol & Alcoholism*(Suppl. 1), 141–145.

Heath, A. C., & Martin, N. G. (1991b). Intoxication after an acute dose of alcohol: An assessment of its association with alcohol consumption patterns using twin data. *Alcoholism, Clinical and Experimental Research, 15,* 122–128.

Heath, A. C., & Martin, N. G. (1992). Genetic differences in psychomotor performance decrement after alcohol: A multivariate analysis. *Journal of Studies on Alcohol, 53,* 262–271.

Heath, A. C., & Martin, N. G. (1994). Genetic influences on alcohol consumption patterns and problem drinking: Results from the Australian NH&MRC twin panel follow-up survey. *Annals of the New York Academy of Sciences, 708,* 72–85.

Heath, A. C., Meyer, J., Eaves, L. J., & Martin, N. G. (1991a). The inheritance of alcohol consumption patterns in a general population twin sample: I. Multidimensional scaling of quantity/frequency data. *Journal of Studies on Alcohol, 52,* 345–352.

Heath, A. C., Meyer, J., Eaves, L. J., & Martin, N. G. (1991b). The inheritance of alcohol consumption patterns in a general population twin sample: II. Multidimensional scaling of quantity/frequency data. *Journal of Studies on Alcohol, 52,* 345–352.

Heath, A. C., Meyer, J. M., & Martin, N. G. (1991). Inheritance of alcohol consumption patterns in the Australian Twin Survey. In H. Begleiter & C. R. Cloninger (Eds.), *Banbury report 33: Genetics and biology of alcoholism* (pp. 3–13). New York: Cold Spring Harbour Laboratory Press.

Heath, A. C., Slutske, W. S., Bucholz, K. K., Madden, P. A. F., & Martin, N. G. (1997). Behavior genetic methods in prevention research: an overview. In K. J. Bryant, M. Windle & S. G. West (Eds.), *The science of prevention: Methodological advances from alcohol and substance abuse research* (pp. 123–163). Washington, DC: American Psychological Association.

Heath, A. C., Slutske, W. S., Madden, P. A. F., Bucholz, K. K., Dinwiddie, S. H., Whitfield, J. B., Dunne, M. P., Statham, D., & Martin, N. G. (1994). Genetic effects on alcohol consumption patterns and problems in women. *Alcohol & Alcoholism* (Suppl. 2), 53–57.

Heath, A. C., Todorov, A. A., Nelson, E. C., Madden, P. A. F., Bucholz, K. K., & Martin, N. G. (2002). Gene–environment interaction effects on behavioral variation and risk of complex disorders: The example of alcoholism and other psychiatric disorders. *Twin Research, 5,* 30–37.

Heath, A. C., Whitfield, J. B., Madden, P. A. F., Bucholz, K. K., Dinwiddie, S. H., Slutske, W. S., Bierut, L. J., & Martin, N. G. (2001). Towards a molecular epidemiology of alcohol dependence: Analysing the interplay of genetic and environmental risk factors. *British Journal of Psychiatry, 178*(Suppl 40), S33–40.

Hesselbrock, M., Easton, C., Bucholz, K. K., Schuckit, M., & Hesselbrock, V. (1999). A validity study of the SSAGA: A comparison with the SCAN. *Addiction, 94,* 1361–1370.

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

Kaprio, J., Viken, R., Koskenvuo, M., Romanov, K., & Rose, R. J. (1992). Consistency and change in patterns of social drinking: A 6-year follow-up of the Finnish Twin Cohort. *Alcoholism, Clinical and Experimental Research, 16*(2), 234–240.

Kendler, K. S., Kuo, P. H., Tai, J. P., Webb, B. T., Kalsi, G., Neale, M. C., Sullivan, P. F., Walsh, D., Patterson, D. G., Riley, B., & Prescott, C. A. (2006). A joint genomewide linkage analysis of symptoms of alcohol dependence and conduct disorder. *Alcoholism, Clinical and Experimental Research, 30*(12), 1972–1977.

Kerr, W. C., Fillmore, K. M., & Bostrom, A. (2002). Stability of alcohol consumption over time: evidence from three longitudinal surveys from the United States. *Journal of Studies on Alcohol, 63*, 325–333.

Knopik, V. S., Heath, A. C., Jacob, T., Slutske, W. S., Bucholz, K. K., Madden, P. A. F., Waldron, M., & Martin, N. G. (2006). Maternal alcohol use disorder and offspring ADHD: Disentangling genetic and environmental effects using a children-of-twins design. *Pscyhological Medicine, 36*, 1461–1472.

Knopik, V. S., Heath, A. C., Madden, P. A. F., Bucholz, K. K., Slutske, W. S., Nelson, E. C., Statham, D., Whitfield, J. B., & Martin, N. G. (2004). Genetic effects on alcohol dependence risk: Re-evaluating the importance of psychiatric and other heritable risk factors. *Psychological Medicine, 34*, 1519–1530.

Lake, R. I. E., Eaves, L. J., Maes, H. H. M., Heath, A. C., & Martin, N. G. (2000). Further evidence against the environmental transmission of individual differences in neuroticism from a collaborative study of 45,850 twins and relatives on two continents. *Behavior Genetics, 30*(3), 223–233.

Little, R. J., & Rubin, D. G. (2000). Causal effects in clinical and epidemiological studies via potential outcomes: Concepts and analytical approaches. *Annual Review of Public Health, 21*, 121–145.

Luciano, M., Zhu, G., Kirk, K. M., Whitfield, J. B., Butler, R., Heath, A. C., Madden, P. A., & Martin, N. G. (2004). Effects of dopamine receptor D4 variation on alcohol and tobacco use and on novelty seeking: Multivariate linkage and association analysis. *American Journal of Medical Genetics Part B (Neuropsychiatric Genetics), 124*(1), 113–123.

Lynskey, M. J., Nelson, E. C., Neuman, R. J., Bucholz, K. K., Madden, P. A. F., Knopik, V. S., Slutske, W., Whitfield, J. B., Martin, N. G., & Heath, A. C. (2005). Limitations of DSM-IV operationalizations of alcohol abuse and dependence in a sample of Australian twins. *Twin Research and Human Genetics, 8*, 574–584.

Madden, P. A. F., Bucholz, K. K., Dinwiddie, S. H., Slutske, W. S., Bierut, L. J., Statham, D. J., Dunne, M. P., Martin, N. G., & Heath, A. C. (1997). Nicotine withdrawal in women. *Addiction, 92*(7), 889–902.

Madden, P. A. F., Bucholz, K. K., Martin, N. G., & Heath, A. C. (2000). Smoking and the genetic contri-

bution to alcohol-dependence risk. *Alcoholism, Research & Health, 24*, 209–214.

Madden, P. A. F., Heath, A. C., & Martin, N. G. (1997). Smoking and intoxication after alcohol challenge in women and men: Genetic influences. *Alcoholism, Clinical and Experimental Research, 21*, 1732–1741.

Madden, P. A. F., Heath, A. C., Starmer, G. A., Whitfield, J. B., & Martin, N. G. (1995). Alcohol sensitivity and smoking history in men and women. *Alcoholism, Clinical and Experimental Research, 19*, 1111–1120.

Martin, N. G. (1987). Genetic differences in drinking habits, alcohol metabolism and sensitivity in unselected samples of twins. In H. W. Goedde & D. P. Agarwal (Eds.), *Genetics and alcoholism* (pp. 109–119). New York: Alan R Liss.

Martin, N. G. (1988). Twin studies of alcohol consumption, metabolism and sensitivity. *Australian Drug and Alcohol Review, 7*, 9–12.

Martin, N. G. (1991). Twin studies of alcohol consumption, metabolism and sensitivity. In H. Begleiter & C. R. Cloninger (Eds.), *Banbury Report 33: Genetics and Biology of Alcoholism* (pp. 15–29). New York: Cold Spring Harbour Laboratory Press.

Martin, N. G., & Boomsma, D. I. (1989). Willingness to drive when drunk and personality: A twin study. *Behavior Genetics, 1989*, 97–111.

Martin, N. G., Perl, J., Oakeshot, J. G., Gibson, J. B., Starmer, G. A., & Wilks, A. V. (1985). A twin study of ethanol metabolism. *Behavior Genetics, 15*, 93–109.

Neale, M. C., Boker, S. M., Xie, G., & Maes, H. H. (2003). *Mx: Statistical modeling* (6th ed.). Richmond, VA: Department of Psychiatry.

Neale, M. C., Cherny, S. S., Sham, P., Whitfield, J. B., Heath, A. C., Birley, A. J., & Martin, N. G. (1999). Distinguishing population stratification from genuine allelic effects with Mx: Association of ADH2 with alcohol consumption. *Behavior Genetics, 29*, 233–243.

Neale, M. C., Eaves, L. J., & Kendler, K. S. (1994). The power of the classical twin study to resolve variation in threshold traits. *Behavior Genetics, 24*(3), 239–258.

Neale, M. C., & Martin, N. G. (1989). The effects of age, sex and genotype on self-report drunkenness following a challenge dose of alcohol. *Behavior Genetics, 19*, 63–78.

Nelson, E. C., Heath, A. C., Bucholz, K. K., Madden, P. A. F., Fu, Q., Knopik, V., Lynskey, M. J., Whitfield, J. B., Statham, D. J., & Martin, N. G. (2004). Genetic epidemiology of alcohol-induced blackouts. *Archives of General Psychiatry, 61*, 257–263.

Nyholt, D. (2006). On the probability of dizygotic twins being concordant for two alleles at multiple polymorphic loci. *Twin Research and Human Genetics, 9*(2), 194–197.

Pape, H., & Hammer, T. (1996). How does young people's alcohol consumption change during the transition to early adulthood? A longitudinal study of changes at aggregate and individual level. *Addiction, 91*(9), 1345–1357.

Reed, T., Plassman, B. L., Tanner, C. M., Dick, D. M., Rinehart, S. A., & Nichols, W. C. (2005). Verification of self-report of zygosity determined via DNA testing in a subset of the NAS-NRC twin registry 40 years later. *Twin Research and Human Genetics, 8*(4), 362–367.

Reed, T., Slemenda, C. W., Viken, R. J., Christian, J. C., Carmelli, D., & Fabsitz, R. R. (1994). Correlations of alcohol consumption with related covariates and heritability estimates in older adult males over a 14-to18-year period: The NHLBI Twin Study. *Alcoholism, Clinical and Experimental Research, 18*(3), 702–710.

Reich, T., Edenberg, H. J., Goate, A., Williams, J. T., Rice, J. P., Van Eerdewegh, P., Foroud, T., Hesselbrock, V., Schuckit, M. A., Bucholz, K., Porjesz, B., Li, T. K., Conneally, P. M., Nurnberger Jr., J. I., Tischfield, J. A., Crowe, R. R., Cloninger, C. R., Wu, W., Shears, S., Carr, K., Crose, C., Willig, C., & Begleiter, H. (1998). Genome-wide search for genes affecting the risk for alcohol dependence. *American Journal of Medical Genetics, 81*(3), 207–215.

Saha, T. D., Chou, S. P., & Grant, B. F. (2006). Toward an alcohol use disorder continuum using item response theory: Results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Psychological Medicine, 36*(7), 931–941.

Slutske, W. S., Heath, A. C., Dinwiddie, S. H., Madden, P. A. F., Bucholz, K. K., Dunne, M. P., Stratham, D. J., & Martin, N. G. (1998). Common genetic risk factors for conduct disorder and alcohol dependence. *Journal of Abnormal Psychology, 107*, 363–374.

Slutske, W. S., Heath, A. C., Madden, P. A., Bucholz, K., Stratham, D. J., & Martin, N. G. (2002). Personality and the genetic risk for alcohol dependence. *Journal of Abnormal Psychology, 111*(1), 124–133.

Slutske, W. S., Heath, A. C., Madden, P. A. F., Bucholz, K. K., Dinwiddie, S. H., Dunne, M. P., Statham, D., & Martin, N. G. (1996). Reliability and reporting biases for perceived parental history of alcohol-related problems: Agreement between twins and differences between discordant pairs. *Journal of Studies on Alcohol, 57*, 387–395.

Slutske, W. S., Heath, A. C., Madden, P. A. F., Bucholz, K. K., Dinwiddie, S. H., Dunne, M. P., Statham, D., Whitfield, J. B., & Martin, N. G. (1995). Is alcohol-related flushing a protective factor for alcoholism in Caucasians? *Alcoholism, Clinical and Experimental Research, 19*, 582–592.

Slutske, W. S., True, W. R., Scherrer, J. F., Goldberg, J., Bucholz, K. K., Heath, A. C., Henderson, W. G., Elsen, S. A., Lyons, M. J., & Tsuang, M. T. (1998). Long-term reliability and validity of alcoholism diag-

noses and symptoms in a large national telephone interview survey. *Alcoholism, Clinical and Experimental Research, 22*(3), 553–558.

Webb, G. R., Redman, S., Gibberd, R. W., & Sanson-Fisher, R. W. (1991). The reliability and stability of a quantity-frequency method and a diary method of measuring alcohol conumption. *Drug and Alcohol Dependence, 27*, 223–231.

Wennberg, P., Bohman, M., & Andersson, T. (2000). Variations and stability in drinking patterns in a cohort of Swedish males. *Scandinavian Journal of Public Health, 28*(4), 312–316.

Whitfield, J. B., Fletcher, L. M., Murphy, T. L., Powell, L. W., Halliday, J., Heath, A. C., & Martin, N. G. (1998). Smoking, obesity, and hypertension alter the dose-response curve and test sensitivity of carbohydrate-deficient transferrin as a marker of alcohol intake. *Clinical Chemistry, 44*(12), 2480–2489.

Whitfield, J. B., Gu, Z., Madden, P. A., Neale, M. C., Heath, A. C., & Martin, N. G. (2004). The genetics of alcohol intake and of alcohol dependence. *Alcoholism, Clinical and Experimental Research, 28*(8), 1153–1160.

Whitfield, J. B., & Martin, N. G. (1983). Inheritance and alcohol as factors influencing plasma uric acid levels. *Acta Geneticae Medicae et Gemellologiae Twin Research, 32*, 117–126.

Whitfield, J. B., & Martin, N. G. (1985). Individual differences in plasma ALT, AST and GGT: contributions of genetic and environmental factors, including alcohol consumption. *Enzyme, 33*, 61–69.

Whitfield, J. B., & Martin, N. G. (1993). Aversive reactions to alcohol use in Europeans. *Alcoholism, Clinical and Experimental Research, 17*, 131–134.

Whitfield, J. B., & Martin, N. G. (1994). Alcohol consumption and alcohol pharmacokinetics: interactions within the normal population. *Alcoholism, Clinical and Experimental Research, 18*, 238–243.

Whitfield, J. B., & Martin, N. G. (1996). Alcohol reactions in subjects of European descent: effects on alcohol use and on physical and psychomotor responses to alcohol. *Alcoholism, Clinical and Experimental Research, 20*, 81–86.

Whitfield, J. B., Nightingale, B. N., Bucholz, K. K., Madden, P. A. F., Heath, A. C., & Martin, N. G. (1998). ADH genotypes and alcohol use and dependence in Europeans. *Alcoholism, Clinical and Experimental Research, 22*, 1463–1469.

Whitfield, J. B., Nightingale, B. N., O'Brien, M. E., Heath, A. C., Birley, A. J., & Martin, N. G. (1998). Molecular biology of alcohol dependence, a complex polygenic disorder. *Clinical Chemistry and Laboratory Medicine, 36*, 633–636.

Whitfield, J. B., O'Brien, M. E., Nightingale, B. N., Zhu, G., Heath, A. C., & Martin, N. G. (2003). ADH genotype does not modify the effects of alcohol on

high-density lipoprotein. *Alcoholism, Clinical and Experimental Research, 27,* 509–514.

Whitfield, J. B., Pang, D., Bucholz, K. K., Madden, P. A. F., Heath, A. C., Statham, D. J., & Martin, N. G. (2000). Monoamine oxidase: Associations with smoking, alcohol dependence, and other measures of psychopathology and personality. *Psychological Medicine, 30*(443–454).

Whitfield, J. B., Starmer, G. A., & Martin, N. G. (1990). Alcohol metabolism in men and women. *Alcoholism, Clinical and Experimental Research, 14,* 785–786.

Whitfield, J. B., Zhu, G., Duffy, D. L., Birley, A. J., Madden, P. A. F., Heath, A. C., & Martin, N. G. (2001). Variation in alcohol pharmokinetics as a risk factor for alcohol dependence. *Alcoholism, Clinical and Experimental Research, 25,* 1257–1263.

Whitfield, J. B., Zhu, G., Heath, A. C., Powell, L. W., & Martin, N. G. (2001). Effects of alcohol consumption on indices of iron stores, and of iron stores on alcohol intake markers. *Alcoholism, Clinical and Experimental Research, 22,* 1037–1045.