Validity of the Lifetime Drinking History: A Comparison of Retrospective and Prospective Quantity-Frequency Measures*

LAURA B. KOENIG, PH.D.,[†] THEODORE JACOB, PH.D., AND JON RANDOLPH HABER, PH.D.

Family Research Center, Veterans Affairs Palo Alto Health Care System, 795 Willow Road (MC 151J), Menlo Park, California 94025

ABSTRACT. Objective: The Lifetime Drinking History (LDH) has been used to examine alcohol use throughout the life span. Given its retrospective nature, it is important to examine the validity of the assessment. **Method:** Building on previous work establishing the reliability and validity of the LDH, the current study examined a sample of 1,295 men in the Vietnam Era Twin Registry. The men were assessed retrospectively with the LDH in 2000, at an average age of 51. The drinking patterns of these same men were also assessed prospectively in four prior studies, taking place in 1987, 1990, 1992, and 1995. **Results:** Validity of

the LDH was examined by comparing the correspondence between the prospective and retrospective quantity-frequency measures and reported age at first regular drinking. Correlations between the retrospective and prospective assessments were high for age at first regular drinking (.42-.58) and quantity-frequency measures (.47-.69), although some mean differences in the amount of consumption existed. **Conclusions:** Results support the use of the LDH in reporting phases of drinking across the life span. (*J. Stud. Alcohol Drugs* 70: 296-303, 2009)

THE LIFETIME DRINKING HISTORY (LDH; Jacob, unpublished; Skinner and Sheu, 1982), a retrospective, interview-based procedure, is used to identify patterns of alcohol use, abuse, and dependence beginning with the onset of regular drinking and ending with the individual's current drinking pattern. Patterns of drinking over time are constructed in terms of different drinking phases, where a phase change is defined in terms of a change in quantity or frequency of drinking. The use of this pattern of questioning (asking individuals to describe different drinking phases) allows for a retrospective description of alcohol use across the entire life course, which is hard to obtain using longitudinal, prospective methods. Because there are often memory limitations in an individual's ability to recall behavior that happened many years ago, it is important to examine the reliability and validity of the LDH to determine the accuracy of these retrospective reports.

A number of studies have examined the psychometric properties of the LDH and (for the most part) have reported moderately high to highly significant test-retest correlations for a range of drinking variables in samples of problem drinkers (Sobell et al., 1988) and nonproblem drinkers (Chaikelson et al., 1994; Lemmens et al., 1997). Also, concurrent validity studies have reported moderate correlations

between the LDH and the Michigan Alcohol Screening Test and the Alcohol Use Inventory (Skinner and Schuller, 1982; Skinner and Sheu, 1982), and good agreement between subject-collateral reports on the LDH has been obtained using untreated problem drinkers (Gladysjo et al., 1992). Data based on LDH reports have also been shown to correlate with biological markers related to alcohol use (Skinner et al., 1981). Finally, recent analysis of 5-year test-retest data on a sample of nontreated alcoholics yielded moderate to high reliability coefficients on a variety of LDH variables, such as the quantity-frequency index (QFI), frequency, and drinking at home (Jacob et al., 2006).

None of these studies, however, compared the LDH retrospective reports with reports of the same variables obtained prospectively. To address this gap, the current study contrasted reports from four prospective assessments of drinking (in years 1987, 1990, 1992, and 1995) with the retrospective information given by the same individuals for the same years using the LDH in 2000. These comparisons focused on two major drinking domains: (1) age at first regular drinking and (2) the quantity and frequency of alcohol consumption.

Method

Participants

The participants were drawn from the Vietnam Era Twin Registry (Eisen et al., 1987; Goldberg et al., 1987; Henderson et al., 1990), which contains 7,375 male-male twin pairs born between 1939 and 1957, where both twins served on

Received: April 22, 2008. Revision: August 13, 2008.

*This research was supported by National Institute on Alcohol Abuse and Alcoholism grant R01 AA016402 and a Veterans Affairs Merit Award.

[†]Correspondence may be sent to Laura B. Koenig at the above address or via email at: laura.koenig@va.gov.

active military duty during the Vietnam era (1965-1975). The men used in the current study were participants in the Family Twin Study, a study of twins and their children using a children of alcoholics design (see Jacob et al., 2003), which involved the administration of the LDH to 1,295 twins, 420 of whom had a lifetime diagnosis of alcohol dependence (American Psychiatric Association, 1994), although the clustering criterion was not considered. All 1,295 men in the Family Twin Study were included in the 1987 and 1992 assessments; 1,207 (93.2%) were included in a 1990 assessment; and 860 (66.4%) were included in a 1995 assessment. A total of 814 men (62.9%) were included in all five assessments. In the sample of 420 men with a lifetime diagnosis of alcohol dependence, 323 (76.9%) were included in all five assessments. The average (SD) age of the men was 38.0 (2.6) in 1987, 40.3 (2.6) in 1990, 41.6 (2.5) in 1992, 45.8 (2.5) in 1995, and 50.9 (2.8) in 2000.

Procedure

Men enrolled in the Vietnam Era Twin Registry were contacted and asked to participate in various studies, between 1985 and the present, aimed at examining a variety of psychological and health domains. A mailed survey was conducted by the Vietnam Era Twin Registry in 1987 (Survey of Health; Eisen et al., 1987; Henderson et al., 1990) and by the National Heart, Lung, and Blood Institute in 1990 (McCarren et al., 1994). Telephone interviews were conducted in 1992 (the Harvard Drug Study; Tsuang et al., 1996) and in 1995 (Health Services Utilization and Alcohol Related Problems survey; Eisen et al., 1999; Romeis et al., 1999). The Family Twin Study, which included the LDH, began in 2000 and was conducted by telephone (Jacob et al., 2003).

Measures

Each of the assessments—except the 1990 National Heart, Lung, and Blood Institute survey—asked participants to indicate the age when they began to drink regularly. In 1987, the term *regular drinking* was used without a specific definition. In 1992 and 1995, regular drinking was defined as drinking at least once a month for 6 months. In 2000, on the LDH, regular drinking was defined as drinking either at least once a month for 6 months or at least once a week for 8 weeks. A number of men reported that they had never drunk regularly: 245 in 1987, 116 in 1992, 19 in 1995, and 174 in 2000 (only 13, 4, 5, and 4 men, respectively, in the 420 lifetime alcohol-dependent sample). Only regular drinkers were included for the age-at-first-regular-drinking analyses when using the entire Family Twin Study sample.

Questions regarding quantity and frequency of drinking were asked at each of the assessments. The frequency of drinking was established by asking how many days an individual usually drank in a given timeframe (e.g., weekly,

monthly). Quantity of drinking was then assessed by inquiring about the typical number of drinks consumed on the days when the individual drank. After standardizing the quantity and frequency scores, QFIs were calculated by taking the frequency (number of days drank per month) multiplied by the quantity (number of drinks usually drank on each occasion) for a total number of drinks per month. A maximum QFI score of 360 was used to characterize the upper end of the index (which corresponds to 12 drinks a day, if an individual drank every day), and this included 4 men in 1987, 6 men in 1990, 7 men in 1992, 6 men in 1995, and 74 men for at least one drinking phase of the LDH in 2000. When only data for quantity or frequency were used, no maximum score was set. Finally, because of the skewed nature of the data, the quantity, frequency, and QFI scores were log-transformed [$\ln(\text{score}+1)$] before being correlated with one another.

In the 1987 Survey of Health study, frequency was reported as current days per week. This was changed to days per month by multiplying scores by 4.3. Men who had never had more than 20 drinks in their life, had never started drinking regularly, or had indicated they did not currently drink skipped out of the quantity-frequency questions and were, therefore, given a zero for these scores ($n = 427$).

In the 1990 National Heart, Lung, and Blood Institute study, mailed surveys asked about the frequency and quantity of drinking beer, wine, and distilled spirits in separate beverage-specific questions. Frequency was reported as the number of days in the past 2 weeks. Typical quantity consumed of each beverage was multiplied by the corresponding frequency (which was multiplied by 2.15 to indicate drinks per month), and these values were summed to create an overall QFI score (number of drinks consumed each month). Because information was not gathered on whether the drinks were consumed on different occasions, it was not possible to know the specific separate quantity and frequency estimates for all beverages combined. To estimate frequency of all beverages, the average of the sum of all beverage frequencies and the largest beverage frequency was calculated. Quantity was then estimated by dividing the QFI score by the estimated frequency score. For example, an individual who averages three beers on four occasions and four shots of distilled spirits on four occasions in a month would have a total QFI score of 28 (the number of drinks per month) but a frequency score of 6 (the average of total frequency of eight drinks and largest frequency of four drinks) and a quantity score of 4.7 (28 divided by 6). Individuals were given a zero for QFI if they had never drank more than 20 drinks in their life ($n = 42$).

The 1992 Harvard Drug Study phone interview also asked about drinking frequency based on a weekly timeframe (during the past year), which was multiplied by 4.3 to obtain the number of days per month in which drinking occurred before multiplying by drinking quantity. Unlike the other surveys, participants were still asked the quantity question when they

reported drinking (on average) zero days per week. There were 54 men who reported a quantity after saying zero frequency, and these individuals were given a QFI score of zero to be consistent with the other surveys. Individuals were also given zeros if they had reported never having a single drink, never drinking regularly, or never having more than five drinks in 1 day ($n = 88$).

In 1995, the Health Services Utilization and Alcohol Related Problems telephone survey assessed frequency of drinking in the days per month format (over the last 12 months); but, unlike other surveys, 31 days was an allowed answer. To be consistent across surveys, the frequency for the 22 participants who gave this answer was recoded to 30 days. In the Health Services Utilization and Alcohol Related Problems survey, if an individual never used alcohol regularly and had never been drunk, he skipped the alcohol section ($n = 230$). These men were given a zero for QFI.

In 2000, on the LDH, quantity and frequency of drinking were reported for each drinking phase. Phases, as stated previously, were defined by the interviewee, depending on the ages he felt that his drinking patterns had changed. The number of reported phases ranged from 0 to 11 (mean = 3.0 [1.6]), with an average length of 10.0 (8.7) years per phase (range: 1-41). Frequency was reported as days per month, which was multiplied by number of drinks per drinking occasion to derive a QFI score for each individual during each phase of his drinking history.

Analyses

To assess the correspondence between quantity-frequency measures for the retrospective LDH and the other prospective measures, each participant was given a quantity, frequency, and QFI score based on the LDH data that matched his age at each of the prospective assessments. For example, if an individual were age 38 when reporting for the Survey

of Health study in 1987, the quantity-frequency measures compared with his reported Survey of Health data would be LDH data that corresponded to age 38 (i.e., from the phase that included age 38).

The overall significance of the mean differences in age at first regular drinking was assessed by repeated-measures analysis of variance, with pairs of means tested by dependent t tests after obtaining overall significance. Mean differences in retrospective and prospective assessments of quantity-frequency were tested via PROC Mixed in SAS (Littell et al., 1996) to account for the correlated nature of the data (i.e., the twins). This correction produced almost identical significance levels as dependent t tests in SPSS Version 15 (SPSS Inc., Chicago, IL) without correction. Because of the highly skewed nature of the data, the Wilcoxon signed rank test was used to test median differences on the QFI from retrospective to prospective assessments. Because the number of comparisons completed was large, the significance level for all tests was set to .01. Correlations between the retrospective and prospective assessments assessed rank-order stability in reporting alcohol consumption.

Results

Age at first regular drinking

The average age reported at first regular drinking for each of the assessments is given in Table 1. Data for both all regular drinkers and for the 420 lifetime alcohol-dependent drinkers are reported for the entire sample as well as a subsample who reported starting regular drinking by the Survey of Health assessment (in 1987) who had data for all four assessments. This subsample does not include the "new drinkers" (i.e., those who reported starting drinking after 1987) in the later assessments that would legitimately raise the average age at first regular drinking for the later assessments.

Table 1. Descriptive statistics for age at first regular drinking, across four datasets

Year and assessment	Regular drinkers			Lifetime AD		
	Sample size	Age Mean (SD)	Range	Sample size	Age Mean (SD)	Range
Entire sample						
1987 SOH	1,030	19.3 ^a (3.0)	9-37	402	19.0 ^a (3.1)	9-37
1992 HDS	1,179	18.2 ^b (2.7)	5-38	416	17.6 ^b (2.6)	9-35
1995 HSUS	772	19.4 ^a (3.3)	9-50	329	18.9 ^a (3.0)	9-37
2000 LDH	1,094	20.1 ^c (4.7)	5-53	415	19.2 ^a (3.4)	5-38
Subsample*						
1987 SOH	748	19.2 ^a (2.8)	11-37	328	19.0 ^a (2.9)	11-37
1992 HDS	740	18.1 ^b (2.6)	6-37	327	17.7 ^b (2.5)	9-35
1995 HSUS	714	19.4 ^a (3.4)	9-50	319	18.9 ^a (3.0)	9-37
2000 LDH	720	19.9 ^c (4.1)	12-47	325	19.3 ^a (3.5)	12-38

Notes: Means with the same superscripts are not significantly different from one another at $p < .01$. The 1995 National Heart, Lung, and Blood Institute survey did not ask respondents about age at first regular drinking. Men who reported never having started regular drinking were not included. AD = alcohol dependence; SOH = Survey of Health; HDS = Harvard Drug Study; HSUS = Health Services Utilization Survey; LDH = Lifetime Drinking History administered as part of the Family Twin Study.

*Subsample assessed at all four timepoints, with age onset of drinking before SOH assessment.

Table 2. Correlations between age at first regular drinking across four datasets in the entire sample (below diagonal) and 420 lifetime alcohol-dependent cases (above diagonal)

	1987 SOH	1992 HDS	1995 HSUS	2000 LDH
1987 SOH		.49 (<i>n</i> = 401)	.52 (<i>n</i> = 319)	.49 (<i>n</i> = 399)
1992 HDS	.45 (<i>n</i> = 1,009)		.58 (<i>n</i> = 328)	.54 (<i>n</i> = 411)
1995 HSUS	.42 (<i>n</i> = 714)	.51 (<i>n</i> = 765)		.55 (<i>n</i> = 326)
2000 LDH	.42 (<i>n</i> = 970)	.45 (<i>n</i> = 1,065)	.52 (<i>n</i> = 738)	

Notes: Correlations for the entire sample are below the diagonal, and correlations for the 420 lifetime alcohol-dependent cases are above the diagonal. Men reporting never having started regular drinking were not included. SOH = Survey of Health; HDS = Harvard Drug Study; HSUS = Health Services Utilization Survey; LDH = Lifetime Drinking History administered as part of the Family Twin Study. All correlations are $p < .01$.

The average age at first regular drinking was approximately 19 for the four surveys, with slight variation. Superscripts in the table show which means were significantly different from one another. The LDH report had the latest age at onset of regular drinking, even in the subsample that excluded the later onset drinkers. Correlations between the four different reports are given in Table 2, with data for the entire sample below the diagonal and data for the 420 alcohol-dependent men above the diagonal. The correlations were all highly significant (p 's $< .01$) and ranged from .42 to .58.

Quantity-frequency measures

Tables 3, 4, and 5 present the correlations between the log-transformed prospective quantity, frequency, and QFI measures and the corresponding LDH retrospective measure (see the last column of each table). Correlations were all significant (p 's $< .01$). The correlations for quantity of drinking ranged from .47 to .62, for frequency of drinking from .58 to .69, and QFI from .57 to .69. In general, the weakest correlations were for the data reaching back to 1987, a recall period of 13 years from when the LDH was administered. The strongest correlations were for the 1995 data, in which the recall period was only 5 years.

The tables also show the descriptive statistics for the retrospective and prospective QFI measures and indicate that there were significant mean differences between retrospective and prospective assessments (see superscripts in the tables). The largest mean differences were for the QFI for the sample of lifetime alcohol-dependent drinkers (at the bottom of Table 3). Medians differed only for the 1987 and 1995 data for the entire sample, whereas all four medians differed in the alcohol-dependent sample. As seen, when the men were reporting their alcohol consumption retrospectively, they reported, on average, higher levels of consumption. Separate analysis of the quantity and frequency variables for the lifetime alcohol-dependent men (Table 5) indicated significant mean level differences for the quantity variable and fewer, weaker differences for the frequency variable.

Table 3. Descriptive statistics for prospective and retrospective assessments of the quantity-frequency index measure for the entire sample and the 420 lifetime alcohol-dependent men

Prospective reports				Retrospective reports				$r^{\dagger}$
Year and assessment	n	Median	Mean (SD)	Year in LDH	n	Median	Mean (SD)	
Entire sample								
1987 SOH ^{b,§}	1,288	7.0	30.0 (50.5)	1987	1,290	8.0	37.8 (69.5)	.60
1990 NHLBI	1,183	6.5	30.2 (54.0)	1990	1,201	8.0	33.5 (64.6)	.63
1992 HDS ^{b,§}	1,295	4.3	22.3 (46.3)	1992	1,291	6.0	34.0 (67.3)	.63
1995 HSUS ^b	849	5.0	26.6 (50.2)	1995	858	6.0	34.0 (66.1)	.69
Non-alcohol-dependent sample								
1987 SOH*	873	0.0	23.6 (42.3)	1987	870	4.5	20.9 (41.0)	.59
1990 NHLBI [§]	798	6.5	23.1 (41.0)	1990	810	4.0	19.5 (39.1)	.62
1992 HDS ^{a,§}	875	0.0	15.9 (33.0)	1992	871	4.0	20.1 (41.3)	.60
1995 HSUS	511	4.0	22.3 (42.1)	1995	517	6.0	22.0 (39.2)	.69
Lifetime alcohol-dependent sample								
1987 SOH ^{b,§}	415	17.2	43.7 (62.5)	1987	420	30.0	72.9 (97.7)	.57
1990 NHLBI ^{b,§}	385	12.9	44.9 (71.8)	1990	391	24.0	62.7 (91.8)	.63
1992 HDS ^{b,§}	420	8.6	35.5 (63.8)	1992	420	16.0	62.9 (95.7)	.66
1995 HSUS ^{b,§}	338	6.5	33.1 (59.9)	1995	341	8.0	52.2 (90.1)	.69

Notes: LDH = Lifetime Drinking History administered as part of the Family Twin Study. r = correlation (between retrospective and prospective reports of consumption); SOH = Survey of Health; NHLBI = National Heart, Lung, and Blood Institute; HDS = Harvard Drug Study; HSUS = Health Services Utilization Survey.

^a $p < .01$, ^b $p < .001$, when comparing retrospective to prospective mean via t test; * $p < .05$, [§] $p < .001$, when comparing retrospective to prospective median via Wilcoxon Signed Rank Test.

[†]The correlations between retrospective LDH and prospective assessment scores were calculated using log-transformed scores. All correlations were significant at $p < .01$.

Table 4. Descriptive statistics for prospective and retrospective assessments of quantity and frequency for the entire sample

Year and assessment	Prospective ratings			Retrospective ratings				$r^{\dagger}$
	Sample size	Mean (SD)	Range	Year from LDH	Sample size	Mean (SD)	Range	
Quantity								
1987 SOH*	1,292	1.8 (2.6)	0-20	1987	1,289	3.2 (3.9)	0-48	.52
1990 NHLBI*	1,188	2.2 (2.6)	0-21	1990	1,202	3.0 (3.8)	0-48	.56
1992 HDS*	1,207	1.8 (2.5)	0-24	1992	1,292	3.0 (4.0)	0-50	.55
1995 HSUS*	855	2.3 (3.2)	0-58	1995	858	3.0 (3.9)	0-48	.61
Frequency ^a								
1987 SOH*	1,291	8.1 (10.6)	0-30	1987	1,290	6.7 (8.8)	0-30	.61
1990 NHLBI ^b *	1,204	7.9 (10.4)	0-51	1990	1,203	6.3 (8.6)	0-30	.61
1992 HDS	1,207	6.5 (8.4)	0-30	1992	1,213	6.7 (8.9)	0-30	.59
1995 HSUS	851	7.0 (9.2)	0-30	1995	859	6.5 (9.0)	0-30	.69

Notes: LDH = Lifetime Drinking History administered as part of the Family Twin Study; r = correlation (between retrospective and prospective reports of consumption); SOH = Survey of Health; NHLBI = National Heart, Lung, and Blood Institute; HDS = Harvard Drug Study; HSUS = Health Services Utilization Survey. ^aThe upper limit of the frequency scores has been rounded; for weekly reports of drinking frequency, multiplying by 4.3 yields an upper limit of 30.1 days per month; ^bthe quantity and frequency scores for the NHLBI survey were calculated based on beverage-specific questions.

* $p < .01$, when comparing retrospective to prospective mean.

[†]The correlations between retrospective LDH and prospective assessment scores were calculated using log-transformed scores; all correlations were significant at $p < .01$.

To better understand the mean-level differences in the alcohol-dependent sample, analyses compared retrospective and prospective QFI means after splitting the sample by other indicators of drinking. Results indicated that mean differences were not related to number of phases reported by alcohol-dependent men or to alcohol-dependence status during the 2000 administration of the LDH. On the other hand, some of the effect was carried by those with more extreme QFI scores. For example, deleting outliers with QFIs greater than or equal to 240 drinks per month based on the LDH report (less than 10% of people at each assessment) decreased

the mean differences to nonsignificance for comparison between the LDH and the 1990 and 1995 data but not for the 1987 or 1992 data. Chronicity of alcohol dependence also affected retrospective-prospective mean differences. After trichotomizing the 420 men into three equal groups, representing the number of years an alcohol dependence diagnosis was given throughout their life (1-9 diagnoses, 10-17 diagnosis, and 18-41 diagnoses), the retrospective-prospective QFI mean differences were significant for the most chronic group; however, the mean differences for the least chronic group were not significant. Those with the most discrepant

Table 5. Descriptive statistics for prospective and retrospective assessments of quantity and frequency for the 420 lifetime alcohol-dependent men

Year and assessment	Prospective ratings			Retrospective ratings				$r^{\dagger}$
	Sample size	Mean (SD)	Range	Year from LDH	Sample size	Mean (SD)	Range	
Quantity								
1987 SOH [§]	418	2.5 (3.0)	0-20	1987	420	4.8 (5.1)	0-48	.47
1990 NHLBI [§]	386	2.7 (3.2)	0-21	1990	391	4.3 (5.2)	0-48	.55
1992 HDS [§]	418	2.3 (3.1)	0-24	1992	420	4.3 (5.5)	0-50	.59
1995 HSUS [§]	339	2.4 (2.8)	0-18	1995	341	3.7 (4.9)	0-48	.62
Frequency ^a								
1987 SOH	417	10.5 (11.4)	0-30	1987	420	10.6 (10.8)	0-30	.58
1990 NHLBI ^b	392	9.7 (11.8)	0-51	1990	391	9.7 (10.7)	0-30	.63
1992 HDS*	418	8.0 (9.7)	0-30	1992	420	9.4 (11.0)	0-30	.65
1995 HSUS	339	7.8 (10.0)	0-30	1995	341	8.0 (10.5)	0-30	.69

Notes: LDH = Lifetime Drinking History administered as part of the Family Twin Study; r = correlation (between retrospective and prospective reports of consumption); SOH = Survey of Health; NHLBI = National Heart, Lung, and Blood Institute; HDS = Harvard Drug Study; HSUS = Health Services Utilization Survey. ^aThe upper limit of the frequency scores has been rounded; for weekly reports of drinking frequency, multiplying by 4.3 yields an upper limit of 30.1 days per month; ^bnote that the quantity and frequency scores for the NHLBI survey do not actually represent true days drank or consumption per occasion because the scores were summed based on the beverage-specific questions. The quantity measure assumed all three beverages were consumed on the same day, while the frequency measure assumed all three beverages were drunk on different days.

* $p < .01$, [§] $p < .001$, when comparing retrospective to prospective mean.

[†]The correlations between retrospective LDH and prospective assessment scores were calculated using log-transformed scores. All correlations were significant at $p < .01$.

retrospective-prospective QFI scores were also more likely to have a lifetime diagnosis of antisocial personality disorder, lifetime drug disorder, higher scores on aggression and lower scores on agreeableness, and to have spent less time living with a partner. The retrospective-prospective differences were substantially larger among alcohol-dependent cases with antisocial personality disorder (around 30 individuals at each assessment had a lifetime antisocial personality disorder diagnosis) than among alcohol-dependent cases without antisocial personality disorder. Tables with these means and correlations are available on request.

Discussion

The major conclusion to be drawn from current analyses is that the retrospective LDH is a valid measure of drinking history. Onset of regular drinking, reported retrospectively at each assessment, was highly similar over time. This finding speaks to the general reliability of retrospective reporting. For QFI variables, there was high rank-order correspondence between the prospective reports and the retrospective LDH reports from the same year that the prospective reports were collected. Mean differences, however, were significant, especially for problem drinkers, who reported drinking more when being assessed retrospectively. As expected, correspondence between retrospective and prospective reports was highest when comparing the most recent prospective assessment with the retrospective data. Several aspects of these findings deserve further comment.

First, finding mean differences but high rank-order stability indicates that an individual's actual amount of alcohol consumed may be biased in the retrospective reports, although there is still high consistency in reporting low, medium, or high levels of consumption. That is, individuals reporting high levels of consumption on the LDH were consuming large amounts of alcohol based on the prospective measures, but possibly not as much as they report retrospectively. The purpose of the LDH is to identify individuals with alcohol problems and to follow increases and decreases in alcohol use over time, and the current study supports this use. On the other hand, the use of the LDH may be problematic when accurate reporting of the volume of alcohol consumed is necessary, such as examining the effects of alcohol intake on chronic disease or health. If individuals are overreporting the volume of alcohol intake on the LDH, health recommendations based on these reports for what is considered problematic consumption would be too generous.

Second, retrospective QFI reports were found to yield higher values than the prospective reports, an unexpected finding that contrasts with most findings in this area, in which retrospective reporting has most often lead to underestimates of behavior (Searles et al., 2002; Wynn, 2000). In the current study, the presence of significant mean differences was most pronounced for the quantity (vs frequency)

measure and for the alcohol-dependent versus the non-alcohol-dependent group. Furthermore, the effect observed within the alcohol-dependent group was, to a significant extent, accounted for by a smaller subset of alcohol-dependent cases who represented the heaviest drinkers, those with the most chronic course of alcohol dependence, those with the highest prevalence of comorbid disorders (including antisocial personality disorder and drug-use disorder), those with higher levels of aggression, and those living without a partner. From the larger alcoholism typology literature, such a profile is most often associated with *antisocial alcoholism*, characterized by a more severe and chronic form of the disorder and by more limited, close personal relationships (Zucker, 1994). But why would this subset of alcoholics have an artificially high score on the retrospective QFI assessments and/or an artificially low QFI score on the prospective assessments? One possibility is that these alcohol-dependent cases could have been underreporting during the prospective assessments when they were more keenly aware of their problem drinking. But when these individuals moved beyond these periods of high consumption, they may have felt more comfortable reporting their actual drinking levels at those earlier times, knowing that those times are in the past. Certainly, no interpretation of current findings lends itself to a definitive explanation, and more research needs to be done to disentangle reporting bias from memory effects on the reporting of alcohol consumption. The important conclusion for this study, however, is that the rank-order stability of consumption was high from prospective to retrospective ratings; therefore, there would still be predictive validity to the retrospective ratings.

Third, the sample included individuals who had never regularly drank, which created skewness in the data. Setting 360 as a maximum QFI score and correcting scores with log-transformation reduced this skewness, although there were still many people with zero scores at the low end of the distribution. These men reported never drinking on the prospective measures and would be likely to correctly remember the absence of drinking on the retrospective LDH. In this way, nondrinkers might raise the correlations between the prospective and retrospective quantity-frequency measures. But, when data for nondrinkers were excluded, the correlations remained significant (ranging from .52 to .62 for QFI, with sample sizes between 611 and 1,161). Correlations for the raw scores (without nondrinkers, without a maximum cutoff of 360, and without log-transformation) were also significant and ranged from .40 to .54. Thus, in the subsample of regular alcohol users, the retrospective LDH exhibits notable validity, even among individuals who might have had a more difficult time remembering the quantity and frequency of alcohol use for a given year because of their more frequent, and likely changing, patterns of use.

Fourth, the reports from the 1990 National Heart, Lung, and Blood Institute survey were hardest to compare with the

other prospective measures and with the retrospective LDH, given the different beverage-specific format. Previous studies have shown that this format of asking beverage-specific questions leads to higher reports of consumption (Dawson, 1998; Russell et al., 1991). Although the quantities and frequencies reported were higher, the correlations with the retrospective LDH were comparable to those from the other prospective surveys. Additionally, the different reference periods (e.g., "past 2 weeks," "current days per week," and "days per month over past year") used in the different assessments may have affected reports of quantity and frequency. Difficulties would clearly exist for minimal drinkers who drink only once or twice a month and for individuals who drink heavily in a given week but then have a period of abstinence when they are asked about drinking "per week." Using different reference periods for the different assessments allows for more measurement error in comparing QFIs across assessments.

Fifth, several limitations to the current study exist. Only quantity-frequency variables were included in all assessments, and little information exists to validate the LDH for alcohol abuse/dependence symptoms, diagnoses, or other drinking-related variables. The generalizability of the current findings to these variables is unknown; therefore, subsequent efforts addressing these variables would be strongly recommended. Also, although the sample included only men, the results likely apply also to women, and there is no reason to suspect that women would remember their drinking patterns either better or worse than men.

The reliability and validity of retrospective reports are sometimes criticized as inadequate. In the assessment of personality and attitude variables, for example, Ross (1989) found that, after individuals experienced an experimental manipulation designed to change their attitudes, they reported their past attitudes to be more congruent with their current (changed) attitudes than they really were. On the other hand, research has found retrospective reporting to be quite reliable for a variety of other, more behaviorally based measures, including childhood experiences (Cournoyer and Rohner, 1996; Hardt et al., 2006), psychopathology (including anxiety and conduct disorders [Holmshaw and Simonoff, 1996] and attention-deficit/hyperactivity disorder [Wierzbicki, 2005]), smoking (Bernaards et al., 2001; Kenkel et al., 2003), and alcohol use (Windle, 2005). Even so, more frequently occurring behaviors appear to be more difficult to recall with accuracy than infrequently occurring behaviors. For example, Bernaards et al. (2001) found that agreement between retrospective and prospective assessments of pack-years of smoking was lower for individuals who were heavier smokers. In general, however, it seems that individuals can adequately assess past behaviors and that asking about specific experiences may help in accuracy.

In summary, the prospective data described in the current report support the findings of past psychometric research involving the LDH using test-retest (Chaikelson et al.,

1994; Jacob et al., 2006; Sobell et al., 1988), subject-collateral (Gladsjo et al., 1992), and concurrent validity methods (Skinner and Schuller, 1982; Skinner and Sheu, 1982). The current study is of particular importance because it is the first to prospectively examine the validity of the retrospective LDH. Given current results, it can be concluded that LDH reports can be used with a greater degree of certainty as to their validity and interpretability, at least for ages 35 to 60. We are confident that the psychometric strengths of the LDH will yield meaningful and interpretable information regarding the lifetime drinking patterns of our sample.

References

- AMERICAN PSYCHIATRIC ASSOCIATION. Diagnostic and Statistical Manual of Mental Disorders (DSM-IV), Washington, DC, 1994.
- BERNAARDS, C.M., TWISK, J.W.R., SNEL, J., VAN MECHELEN, W., AND KEMPER, H.C.G. Is calculating pack-years retrospectively a valid method to estimate life-time tobacco smoking? A comparison between prospectively calculated pack-years and retrospectively calculated pack-years. *Addiction* **96**: 1653-1661, 2001.
- CHAIKELSON, J.S., ARBUCKLE, T.Y., LAPIDUS, S., AND GOLD, D.P. Measurement of lifetime alcohol consumption. *J. Stud. Alcohol* **55**: 133-140, 1994.
- COURNOYER, D.E. AND ROHNER, R.P. Reliability of retrospective reports of perceived maternal acceptance-rejection in childhood. *Psychol. Rep.* **78**: 147-150, 1996.
- DAWSON, D.A. Volume of ethanol consumption: Effects of different approaches to measurement. *J. Stud. Alcohol* **59**: 191-197, 1998.
- EISEN, S., TRUE, W., GOLDBERG, J., HENDERSON, W., AND ROBINETTE, C.D. The Vietnam Era Twin (VET) Registry: Method of construction. *Acta Genet. Med. Gemellol.* **36**: 61-66, 1987.
- EISEN, S.A., WATERMAN, B., SKINNER, C.S., SCHERRER, J.F., ROMEIS, J.C., BUCHOLZ, K., HEATH, A., GOLDBERG, J., LYONS, M.J., TSUANG, M.T., AND TRUE, W.R. Sociodemographic and health status characteristics with prostate cancer screening in a national cohort of middle-aged male veterans. *Urology* **53**: 516-522, 1999.
- GLADJSJO, J.A., TUCKER, J.A., HAWKINS, J.L., AND VUCHINICH, R.E. Adequacy of recall of drinking patterns and event occurrences associated with natural recovery from alcohol problems. *Addict. Behav.* **17**: 347-358, 1992.
- GOLDBERG, J., TRUE, W., EISEN, S., HENDERSON, W., AND ROBINETTE, C.D. The Vietnam Era Twin (VET) Registry: Ascertainment bias. *Acta Genet. Med. Gemellol. Twin Res.* **36**: 67-78, 1987.
- HARDT, J., SIDOR, A., BRACKO, M., AND EGLE, U.T. Reliability of retrospective assessments of childhood experiences in Germany. *J. Nerv. Ment. Dis.* **194**: 676-683, 2006.
- HENDERSON, W.G., EISEN, S., GOLDBERG, J., TRUE, W.R., BARNES, J.E., AND VITEK, M.E. The Vietnam Era Twin Registry: A resource for medical research. *Publ. Hlth Rep.* **105**: 368-373, 1990.
- HOLMSHAW, J. AND SIMONOFF, E. Retrospective recall of childhood psychopathology. *Int. J. Meth. Psychiat. Res.* **6**: 79-88, 1996.
- JACOB, T. Modified Lifetime Drinking History, 1998, unpublished measure.
- JACOB, T., SEILHAMER, R.A., BARGEIL, K., AND HOWELL, D.N. Reliability of Lifetime Drinking History among alcohol dependent men. *Psychol. Addict. Behav.* **20**: 333-337, 2006.
- JACOB, T., WATERMAN, B., HEATH, A., TRUE, W., BUCHOLZ, K.K., HABER, R., SCHERRER, J., AND FU, Q. Genetic and environmental effects on offspring alcoholism: New insights using an offspring-of-twins design. *Arch. Gen. Psychiat.* **60**: 1265-1272, 2003.
- KENKEL, D., LILLARD, D.R., AND MATHIOS, A. Smoke or fog? The usefulness of retrospectively reported information about smoking. *Addiction* **98**: 1307-1313, 2003.

- LEMMENS, P.H., VOLOVICS, L., AND DE HAAN, Y. Measurement of lifetime exposure to alcohol: Data quality of self-administered questionnaire and impact on risk assessment. *Contemp. Drug Probl.* **24**: 581-600, 1997.
- LITTELL, R.C., MILLIKEN, G.A., STROUP, W.W., AND WOLFINGER, R.D. *SAS System for Mixed Models*, Cary, NC: SAS Institute, 1996.
- MCCARREN, M., GOLDBERG, J., RAMAKRISHNAN, V., AND FABSITZ, R. Insomnia in Vietnam Era veteran twins: Influence of genes and combat experience. *Sleep* **17**: 456-461, 1994.
- ROMEIS, J.C., WATERMAN, B., SCHERRER, J.F., GOLDBERG, J., EISEN, S.A., HEATH, A.C., BUCHOLZ, K.K., SLUTSKE, W.S., LYONS, M.J., TSUANG, M.T., AND TRUE, W.R. The impact of sociodemographics, comorbidity and symptom recency on health-related quality of life in alcoholics. *J. Stud. Alcohol* **60**: 653-662, 1999.
- ROSS, M. Relation of implicit theories to the construction of personal histories. *Psychol. Rev.* **96**: 341-357, 1989.
- RUSSELL, M., WELTE, J.W., AND BARNES, G.M. Quantity-frequency measures of alcohol consumption: Beverage-specific vs global questions. *Brit. J. Addict.* **86**: 409-417, 1991.
- SEARLES, J.S., HELZER, J.E., ROSE, G.L., AND BADGER, G.J. Concurrent and retrospective reports of alcohol consumption across 30, 90, and 366 days: Interactive voice response compared with the Timeline Follow Back. *J. Stud. Alcohol* **63**: 352-362, 2002.
- SKINNER, H.A., HOLT, S., AND ISRAEL, Y. Early identification of alcohol abuse: I. Critical issues and psychosocial indicators for a composite index. *Can. Med. Assoc. J.* **124**: 1141-1152, 1981.
- SKINNER, H. AND SCHULLER, R. Validation of the Lifetime Drinking History: Biochemical, Medical History and Psychosocial Evidence, Toronto, Canada: Addiction Research Foundation, 1982.
- SKINNER, H.A. AND SHEU, W.-J. Reliability of alcohol use indices: The Lifetime Drinking History and MAST. *J. Stud. Alcohol* **43**: 1157-1170, 1982.
- SOBELL, L.C., SOBELL, M.B., RILEY, D.M., SCHULLER, R., PAVAN, D.S., CANCELLA, A., KLAJNER, F., AND LEO, G.I. The reliability of alcohol abusers' self-reports of drinking and life events that occurred in the distant past. *J. Stud. Alcohol* **49**: 225-232, 1988.
- TSUANG, M.T., LYONG, M.J., EISEN, S.A., GOLDBERG, J., TRUE, W., LIN, N., MEYER, J.M., TOOMEY, R., FARAONE, S.V., AND EAVES, L. Genetic influences on DSM-III-R drug abuse and dependence: A study of 3,372 twin pairs. *Amer. J. Med. Genet.* **67**: 473-477, 1996.
- WIERZBICKI, M. Reliability and validity of the Wender Utah Rating Scale for college students. *Psychol. Rep.* **96**: 833-839, 2005.
- WINDLE, M. Retrospective use of alcohol and other substances by college students: Psychometric properties of a new measure. *Addict. Behav.* **30**: 337-342, 2005.
- WYNN, P.A. The reliability of personal alcohol consumption estimates in a working population. *Occupat. Med.* **50**: 322-325, 2000.
- ZUCKER, R.A. Pathways to alcohol problems and alcoholism: A developmental account of the evidence for multiple alcoholisms and for contextual contributions to risk. In: ZUCKER, R., BOYD, G., AND HOWARD, J. (Eds.) *The Development of Alcohol Problems: Exploring the Biopsychosocial Matrix of Risk*. NIAAA Research Monograph No. 26, NIH Publication No. 94-3495, Bethesda, MD: National Institute on Alcohol Abuse and Alcoholism, 1994, pp. 255-289.