Published in final edited form as:

Addict Behav. 2010 December ; 35(12): 1138–1143. doi:10.1016/j.addbeh.2010.08.013.

How good is fine-grained Timeline Follow-back data? Comparing 30-day TLFB and repeated 7-day TLFB alcohol consumption reports on the person and daily level

Bettina B. Hoepfner, Ph.D.^a, Robert L. Stout, Ph.D.^b, Kristina M. Jackson, Ph.D.^c, and Nancy P. Barnett, Ph.D.^d

Robert L. Stout: Stout@PIRE.org; Kristina M. Jackson: Kristina_Jackson@brown.edu; Nancy P. Barnett: Nancy_Barnett@brown.edu

^a Center for Alcohol and Addiction Studies, Box G-S121-4, Brown University, Providence, RI 02912, USA

^b Decision Sciences Institute, 1005 Main Street, Suite 8120, Pawtucket, RI 02860

^c Center for Alcohol and Addiction Studies, Box G-S121-4, Brown University, Providence, RI 02912, USA

^d Center for Alcohol and Addiction Studies, Box G-S121-5, Brown University, Providence, RI 02912, USA

Abstract

Objective—This study examined the correspondence of two types of Timeline Follow-back (TLFB) methods, a web-based self-administered, repeated 7-day TLFB and an interviewer-administered 30-day TLFB of alcohol consumption.

Method—Participants were first- and second-year college students ($n=323$, 58.5% female). Day-to-day correspondence of drinking reports and correspondence of person-level indicators of drinking were assessed.

Results—Results indicated that correspondence between the TLFB-30 and TLFB-7 reports was generally good for summary indicators of drinking, but TLFB-7 data indicated a statistically significantly higher number of total drinks consumed, a higher number of days drinking 4+/5+ drinks per day, and a lower number of abstinent days than TLFB-30. Similarly, day-to-day comparison of drinking reports showed that drinking days were more frequently reported using the TLFB-7, a trend which was more pronounced for distal weekdays than recent weekdays. Correlations between TLFB-7 and TLFB-30 reports of drinks per drinking day were also lower for distal compared to recent weekdays ($r=0.61$ vs. $r=0.76$). Using a Poisson regression model, a linearly increasing trend in the absolute value of the difference between TLFB-7 and TLFB-30 drinking reports per day as length of recall increases was found ($b=0.013$, $z=4.43$, with $p < 0.001$).

Conclusions—Our results indicate that participants reported more drinking on the repeated TLFB-7 than on the standard TLFB-30. Furthermore, the result of daily-level analyses showed that

Corresponding Author: Bettina B. Hoepfner, Ph.D., Center for Alcohol and Addiction Studies, Box G-S121-4, Brown University, Providence, RI 02912, Phone: 001-401-863-6630, Fax: 001-401-863-6647, Bettina_Hoepfner@brown.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

discrepancies between methods increased as the length of recall increased. These findings suggest that TLFB assessments covering longer intervals may have reduced accuracy on a fine-grained scale.

Keywords

Timeline Follow-back; web-based surveys; retrospective recall of alcohol; consumption

1. Introduction

The Timeline Follow-back (TLFB) interview is one of the most widely used measures of alcohol consumption. Originally developed as an interviewer-administered paper-and-pencil measure, the TLFB collects retrospective estimates of daily drinking over a specified period of up to 12 months. Memory aids are provided to prompt recall, such as a calendar format and the provision of important dates. The TLFB has been used for over 30 years in clinical and nonclinical research studies (Sobell and Sobell, 2008), and its cross-cultural validity has been demonstrated in several countries (i.e., Australia, Canada, Mexico and Spain; Sobell et al., 2001). Recent research further documents and strengthens the evidence supporting this measurement's construct validity compared with other retrospective drinking assessments, such as the Quick Drinking Screen in samples of outpatient alcohol abusers (Roy et al., 2008) and non-clinical individuals with alcohol problems (Sobell et al., 2003).

The cost of using a skilled interviewer for the TLFB is considerable (Maisto et al., 2008), which has prompted the development of alternative administration methods. To date, several variations of the original interviewer-administered paper-and-pencil assessment exist, where assessment costs are reduced through the use of technology to either eliminate the need for in-person interviews through the use of phone interviews or by replacing the interviewer altogether with automated instructions via a computer program or voice-interactive system. Few studies, however, have documented the impact of different administration methods on the reporting of alcohol consumption.

Two studies have examined the effect of self- and other-administration methods. The first study (Sobell et al., 1996) examined the effect of different self-administration methods (i.e., paper vs. computer) and other-administered methods (i.e., in-person vs. via phone) in two samples. The first sample ($n=63$ problem drinkers entering treatment) completed a paper-and-pencil version of the TLFB, describing their drinking during the 90 days pre-treatment. Approximately 3 weeks later, they used a computer program with the identical instructions to complete another TLFB assessment for the same days. The second sample ($n=40$ problem drinkers who had agreed to participate in a treatment study) completed a 30-day interviewer-guided TLFB via phone one month after treatment, and an interviewer-guided in-person TLFB assessment assessing drinking behavior over the 12-months after treatment, which included the 30 days assessed in the first TLFB assessment. Results from both samples indicated no differences in drinking between administration methods, where total number of drinks, days abstinent, days drinking 1-4 drinks, days drinking 5-9 drinks, days drinking > 10 drinks, and drinks per drinking day were compared. The authors concluded that reliable drinking data can be obtained using these different methods for interviewer-administration and self-administration of the TLFB.

A more recent experimental study ($n=73$ participants of the Veterans Aging Cohort Study; (VACS: Justice et al., 2006)) compared paper & pencil self-administration to interviewer-administration conducted via phone (Maisto et al., 2008). Participants completed two assessments of 30-day TLFB, completed no more than one week apart. They were randomized into four groups of assessment order (i.e., self-self, self-telephone, telephone-self, and telephone-telephone). This study as well reported no differences between administration

methods, where the indices of interest were the median number of total drinks and percentage of hazardous drinkers.

Notably, research on the impact of different administration methods on TLFB data has focused on person-level indices of drinking that summarize drinking into static variables, such as the median number of drinks across a specified timeframe. Much less is known about the day-to-day accuracy of TLFB methods, such as the number of drinks consumed on particular weekdays, for example. Accuracy on such a fine-grained scale is critical, however, if data are to be analyzed on that level. If the goal of a study lies in comparing two groups in their volume of alcohol consumption, such level of detail is not necessary. If, however, the goal is to understand and capture processes and mechanisms of change, the fine-grained details of alcohol consumption matter. A particular strength of the TLFB is its ability to capture sporadic and unpatterned drinking (Sobell, 2003), where both volume and timing of drinking are relevant. Indeed, this feature of the TLFB is particularly relevant in light of the growing evidence that shows that the pattern of drinking is an important determinant of the impact of alcohol use on health, independent of the volume of alcohol consumed (Bobak et al., 2004; Rehm et al., 2001; Room et al., 2003; Russell et al., 2004; Tolstrup et al., 2004). As such, the accuracy of different administration methods of the TLFB on a day-to-day level becomes relevant.

Only a handful of studies have examined the accuracy of the TLFB on a day-to-day basis, where TLFB methods have been compared to real-time and daily-recall methods. Carney et al. (1998) compared 28-day TLFB, ecological momentary assessment (EMA) and daily diary reports of alcohol consumption in two samples. Results indicated that compared to daily diary ($n=20$ problem drinkers) and EMA ($n=48$ moderate drinkers), 28-day TLFB data identified fewer drinking days, fewer drinks per drinking day and fewer total drinks, yet that these differences were modest. Of greater concern were the large individual differences in day-to-day correspondence, which the authors noted suggested that the 28-day TLFB may be less useful for detecting patterns of consumption. Similar disparities in correspondence between day-to-day patterns of drinking were found when TLFB methods were compared to daily interactive voice mail reports, where correlations between daily self-reports and 28-day TLFB ranged from -0.22 to 0.96, with an average of 0.51 (Perrine et al., 1995) and from 0.54 to 0.79 when compared to 7-day TLFB (Toll et al., 2006). These studies suggest that the TLFB may be limited in the accuracy with which it assesses fine-grained characteristics of alcohol consumption, particularly as the length of recall increases.

Promisingly, in addition to changing the type of administration of the TLFB, technology (e.g., internet surveying) has also made it possible to increase the frequency of TLFB administration and to shorten the length of recall intervals. Such repeated, short-interval TLFB assessments are likely subject to fewer recall biases than longer TLFB assessments. They also constitute a different type of response burden for the participant. Past research has shown that comprehensive assessments, such as the 12-month TLFB, are associated with a much lower completion rate than brief measures of alcohol consumption (22% vs. 51% for a mailed out survey) (Cunningham et al., 1999). By comparison, the short and frequent reports of EMA measures of alcohol consumption have often shown compliance rates upward of 90% (e.g., Litt et al., 2000; Todd et al., 2009), though they can also fall below 50% (e.g., Litt et al., 1998). A short-term, frequently repeated TLFB is less frequent than EMA but can be similarly brief, and thus might potentially result in better compliance than more comprehensive TLFB assessments. This study focuses on such short-term, frequently repeated TLFB and compares it to the more traditionally used 30-day TLFB.

The purpose of this study was to evaluate the correspondence of a web-based self-administered, repeated 7-day TLFB and an interviewer-administered 30-day TLFB of alcohol consumption

in a sample of first- and second-year college students. Of particular interest was the day-to-day correspondence of drinking reports in addition to providing information about the correspondence of person-level indicators of drinking. A focus on day-to-day correspondence addresses the accuracy of TLFB data not just in regards to the volume but also the pattern of drinking. In first- and second-year college students, who are underage drinkers and thus less likely to have established drinking patterns in that they intersperse days of abstinence and excessive drinking (Collins et al., 2008), the timing of drinking events is particularly important, rendering a comparison of methods in this population especially salient. We hypothesized that the repeated, short-term TLFB method would result in reports detailing a higher frequency and volume of alcohol consumption, particularly as length of recall increases.

2. Method

2.1. Participants

Data were from a multi-year study designed to evaluate the alcohol-related experiences of college students. For the original study, students and parents of minors received letters inviting them to enroll in the study. A total of $n=1,054$ incoming students at three colleges and universities in southern New England were enrolled (43% enrollment rate). Enrolled participants completed an online consent procedure, followed by a baseline assessment battery. For this paper, only participants who completed two types of TLFB reports (i.e., self-administered online 7-day TLFB and interviewer-administered in-person or phone 30-day TLFB) during an overlapping time period were included. Participants ($n=323$) were on average 18.4 ($SD=0.4$) years of age, and 58.5% female. The sample was 77.7% White, 9.3% Asian, 5.9% multiracial, 4.0% Black, and 3.1% other. 6.8% reported being Hispanic.

2.2. Procedure

Throughout the first two academic years of college, participants completed web-based surveys every other week about their alcohol consumption and alcohol-related consequences during the immediately preceding seven days using a 7-day TLFB format. Each week a random subsample of students who reported a negative alcohol-related experience during that week was contacted, and three months later completed an interview (93% completion rate)¹. At the interview (83.6% in-person, 16.4% via phone; $n=323$), participants completed a 30-day TLFB. Web-based biweekly surveys were not administered during summer breaks, so interviews occurring after the end of the academic year did not have corresponding online TLFB data and were not included.

2.3. Measures

Both measures of TLFB consisted of participants' retrospective recall of the number of "standard" drinks (i.e., defined as the alcohol equivalent of 12 oz of domestic (U.S.) beer, 5 oz of 12%-alcohol wine, or 1.5 oz 80-proof distilled spirits) consumed on each of the previous days. For the biweekly 7-day online TLFB (TLFB-7), students completed a grid with dates for the seven days prior to the online survey completion date. The survey instructed participants to answer for the prior seven days "starting with yesterday" and presented the days of the week in the week grid. Students were given one week to complete the survey, so times between biweekly reports could vary. Thus, for each respondent there was a flexibly-timed 7-day recall period during every *a priori* determined 14-day interval. For the 30-day interviewer-

¹The purpose of the original study was to evaluate assessment reactivity. For this purpose, 2/3 of the participants were asked to complete two interviews, where one interview was scheduled as soon as possible and the other was conducted 3 months later. The remaining 1/3 were asked to complete the 3-month interview only. At the immediate interview, participants were provided with their biweekly online drinking reports while completing their 30-day TLFB. Since the two types of data were not collected independently they are not compared in this paper.

administered TLFB (TLFB-30), participants were guided by a research assistant through a paper and pencil calendar containing dates for the preceding 30 days, starting with the day before the interview. If the interview was conducted via phone (16.4%), calendar forms were mailed prior to the interview to be used by the participant as a reference.

Days on which participants provided data on both the 7-day and the 30-day TLFB were included in the analyses. Given the flexible scheduling of the TLFB-7, the number of overlapping days could vary, and did in fact range from 4 to 21 days (mean = 14.8, SD = 2.8). Person-level indicators of drinking were calculated by summarizing drinking for the overlapping days. Means across days were first calculated per person, and were then averaged across all $n=323$ persons to compare the TLFB-7 and TLFB-30 reports. The following indicators were calculated: total number of drinks, number of days abstinent, number of days consuming 4+/5+ drinks for females and males respectively, and average number of drinks consumed per drinking day, where “drinking day” was defined in two ways: (1) if either TLFB-7 or TLFB-30 data indicated any alcohol was consumed on that day, and (2) only if both TLFB-7 and TLFB-30 data indicated that any alcohol was consumed on that day.

Day-to-day correspondence of drinking reports was evaluated in two ways: (1) day-to-day agreement between TLFB-7 and TLFB-30 (i.e., whether TLFB-7 and TLFB-30 agreed on which days were drinking days), and (2) when both measures agreed that a particular day was a drinking day, the number of drinks reported for that day. Given the importance of weekday patterns in college student drinking (e.g., Del Boca et al., 2004), the day-to-day correspondence of measures was summarized by weekday.

2.4. Analytic strategy

For both types of correspondence (i.e., person-level and day-to-day correspondence), correlations were calculated to describe the relationship between the TLFB-7 and TLFB-30 reports, and paired t -tests were conducted to test for differences in means. Log transformations were used if data were highly skewed. To compare the identification of drinking days on a day-to-day level, the Kappa coefficient was calculated. Since most participants completed two TLFB-7 reports within the 30 days covered by the TLFB-30, it was possible to compare weekday correspondence of day-to-day drinking reports twice per participant. Given that higher discrepancies from daily assessments have been reported for TLFB assessments with greater length of recall (i.e., Perrine et al. (1995) compared to Toll et al. (2006)), we identified weekdays as “most recent” versus “most distal”, and compared them separately. Here, “most recent” and “most distal” refer to the length of recall of the TLFB-30, which could range from 1-30 days. The length of recall for the TLFB-7 could not exceed 7 days. To assess the relationship of length of recall with alcohol reporting more explicitly, we tested for the presence of a linearly increasing trend in the absolute value of the difference between TLFB-7 and TLFB-30 drinking reports per day as length of recall increases using a Poisson regression model, where observations were modeled as nested within persons.

3. Results

3.1. Person-level Indicators

Table 1 summarizes the correspondence of TLFB-7 and TLFB-30 reports of drinking on summary indicators of drinking. In general, students reported more drinking using TLFB-7 than TLFB-30. Three statistically significant mean differences were found: TLFB-7 data indicated a higher number of total drinks consumed, a higher number of days drinking 4+/5+ drinks per day, and a lower number of abstinent days than TLFB-30. Correlations between the two types of reports generally ranged from $r=0.78$ to $r=0.89$. The correlation between TLFB-7 and TLFB-30 for the average number of drinks per drinking day was substantially lower

($r=0.38$), when “drinking day” was defined as having been identified as a drinking day by either TLFB-7 or TLFB-30 (not both).

Interviews conducted “in-person” and “via phone/mail” showed similar results as the full-sample analyses (see Table 1). Fewer differences were statistically significant in the much smaller sample of phone interviews, and the average number of drinks per drinking day (using “both” types of report to identify “drinking days”) was lower in the TLFB-7 than in the TLFB-30 reports, though this difference was not statistically significant.

3.2. Day-to-Day Correspondence

Day-to-day correspondence of drinking reports is presented in Tables 2 and 3. Table 2 summarizes the identification of drinking days for most recent and most distal days of the week. In some instances, there was overlap between “most recent” and “most distal” (e.g., when a participant only had data for one Monday), but such overlap was relatively limited (i.e., ranging from 4% of the sample reporting on only one Sunday, and 23.8% of the sample reporting on only one Monday). To describe “recent” and “distal” more concretely, the length of TLFB-30 recall is provided for each comparison. For almost all of the recent weekdays, TLFB-7 reports indicated more drinking days than TLFB-30 reports. For distal weekdays, TLFB-7 reports always indicated more drinking days than TLFB-30 reports. Furthermore, for each weekday the inter-measure reliability was higher for recent weekdays (ranging from $\kappa=0.59$ to $\kappa=0.76$) than distal weekdays (ranging from $\kappa=0.38$ to $\kappa=0.58$), the latter of which did not exceed “moderate” inter-measure reliability according to Altman (1991). Table 3 presents the number of drinks consumed per drinking days, where “drinking day” was identified as such by both TLFB-7 and TLFB-30 reports. As such, these comparisons were only possible for a small subsample. Overlap between “recent” and “distal” was limited. While the number of drinks reported for drinking days using the TLFB-7 tended to be higher than for the TLFB-30, this difference was only statistically significant for the most recent Friday. Correlations tended to be higher for recent (ranging from $r=0.48$ to $r=0.91$; average $r=0.76$) rather than distal weekdays (ranging from $r=0.45$ to $r=0.90$; average $r=0.61$).

Figure 1 illustrates the degree of discrepancy between TLFB-7 and TLFB-30 reports of drinks consumed per day with respect to length of recall (in days) of the TLFB-30. While the length of recall for TLFB-7 reports could not exceed seven days, recall for specific days on the TLFB-30 reports could be as long as 30 days. For Figure 1, the absolute value of the difference between TLFB-7 and TLFB-30 reports was calculated per person per day, and then averaged per day of length of TLFB-30 recall across persons. Given the intermittent and flexible timing of the TLFB-7, each daily average is based on a different sample size (ranging from $n=24$ on Day 22 to $n=282$ on Day 1), depending on who completed reports for that day. Using Poisson regression, a statistically significant positive linear trend was found ($b=0.013$, $z=4.43$, with $p<0.001$), with greater discrepancies for longer recall lengths.

4. Discussion

This study compared retrospectively reported drinking using two types of TLFB methods, self-administered, repeated 7-day TLFB and interviewer-administered 30-day TLFB. As expected, the correspondence between the TLFB-30 and TLFB-7 reports was generally good for summary indicators of drinking in our sample of first- and second-year college students. This finding is consistent with results from studies with different populations (i.e., problem drinkers) that reported no differences in drinking reports between different administration methods of the TLFB (Maisto et al., 2008; Sobell et al., 1996). Also as expected, participants reported more drinking, both in terms of volume and frequency, on the repeated TLFB-7 than on the standard TLFB-30. This finding is consistent with reports comparing 28-day TLFB to daily and real-time drinking reports (Carney, 1998).

We extended previous studies comparing TLFB administration methods by investigating the day-to-day correspondence of TLFB-7 and TLFB-30 reports of drinking. Similarly to person-level summary indicators, day-to-day comparison of drinking reports showed that drinking days were more frequently reported using the TLFB-7, a trend which was more pronounced for distal weekdays than recent weekdays. Indeed, inter-measure reliability of the identification of drinking days for distal weekdays left much to be desired.

Finally, previous research has indicated high within-person variability in correspondence between prospectively collected daily alcohol data and TLFB data, contrasted with good correspondence on overall levels of drinking (Carney et al., 1998; Perrine et al., 1995). High levels of within-person discrepancies coupled with good correspondence on overall levels of consumption could be indicative of inaccuracies in reporting the timing of drinking events. For example, if a person's alcohol reports are off by only one day on one measure but not the other, within-person correlations across all days would be quite low. Comparisons of within-person correlations can indicate who made errors in the reporting of the timing of events, but do not give information about when these errors were made. In our study, we were able to compare correspondence between TLFB-7 and TLFB-30 reports day by day between persons, thereby enabling us to compare accuracy on specific weekdays and with different lengths of recall. Similar to previous findings, we found that correlations between TLFB-7 and TLFB-30 data were generally lower for day-to-day correspondence of reports of drinks per drinking day than person-level correlations of drinks per drinking day across all recorded days (i.e., $r=0.79$). By being able to compare correlations day by day, we also found that correlations for recent weekdays (averaging $r=0.76$) were much higher than for distal weekdays (averaging $r=0.61$). Taken together with good correspondence in overall levels of drinking, this finding suggests that discrepancies in reporting the timing of drinking events may be increasingly likely for events with longer length of recall. Additionally, in testing discrepancies between TLFB-7 and TLFB-30 reports by day of recall, we found a statistically significant positive linearly increasing trend linking discrepancies with length of TLFB-30 recall. Taken together, our results suggest that the discrepancies between TLFB-7 and TLFB-30 data may be linked to the length of recall.

4.1. Strengths and Limitations

This study is the first to explore the reliability and accuracy of different types of TLFB reports of alcohol consumption on a day-to-day basis. Given the TLFB's strength in assessing sporadic and unpatterned drinking (Sobell, 2003), the accuracy with which it can estimate both volume and timing of drinking is important to investigate, particularly given the importance of fine-grained characteristics and dynamics of drinking (e.g., Del Boca et al., 2004; Greenbaum et al., 2005). The population used in this study is of particular relevance, both because TLFB methods are particularly useful for populations with sporadic drinking patterns, and because college student drinkers comprise a high-risk group that is a target for clinical research. Another strength of this study is its relatively large sample size ($n=323$). Previous studies comparing different administration methods of the TLFB or other fine-grained alcohol consumption data collection methods have used much smaller samples (Maisto et al., 2008).

An important limitation of this study was that the two compared TLFB methods differed in multiple ways: length of recall (i.e., 7 vs. 30 days), type of administration (i.e., self vs. interviewer) and presentation medium (i.e., internet, paper & pencil, phone). Because this study is based on data collected for a different purpose, these characteristics were not experimentally controlled. It should be noted, however, that previous research has not found differences in reports of drinking due to type of administration, and our study revealed few differences in findings for in-person versus telephone/mail interviews. The finding that discrepancies between TLFB-7 and TLFB-30 increased with length of recall further supports the idea that

observed differences in TLFB-7 and TLFB-30 reports are primarily due to length of recall rather than type of administration or presentation medium. Nevertheless, this study cannot disentangle these effects.

5. Conclusions

Our results indicate that participants reported more drinking on the repeated TLFB-7 than on the standard TLFB-30. Importantly, this difference was observed both in terms of volume and frequency. In investigating the day-to-day correspondence of drinking reports, it could be seen that the discrepancy in the identification of drinking days was more pronounced for distal than recent days, and indeed that the discrepancy in reports increased as the length of recall increased. These findings suggest that TLFB assessments covering longer intervals may have reduced accuracy in documenting the timing of drinking events. For research studies that are primarily or exclusively interested in the volume of alcohol consumption without reference to its timing, the reduced accuracy is largely irrelevant. For studies, however, that are interested in the processes underlying drinking behaviors, fine-grained details are important, and the accuracy with which both volume and timing of consumption are reported matters. For such studies, the TLFB-7 would be more appropriate than the TLFB-30.

References

- Altman, DG. Practical Statistics for Medical Research. London, England: Chapman and Hall; 1991.
- Bobak M, Room R, Pikhart H, Kubinova R, Malyutina S, Pajak A, et al. Contribution of drinking patterns to differences in rates of alcohol related problems between three urban populations. *Journal of Epidemiology and Community Health* 2004;58(3):238–242. [PubMed: 14966239]
- Carney MA, Tennen H, Affleck G, Del Boca FK, Kranzler HR. Levels and patterns of alcohol consumption using Timeline Follow-Back, Daily Diaries and Real-Time “Electronic Interviews”. *Journal of Studies on Alcohol* 1998;59(4):447–454. [PubMed: 9647427]
- Collins RL, Kashdan TB, Koutsky JR, Morsheimer ET, Vetter CJ. A self-administered Timeline Followback to measure variations in underage drinkers' alcohol intake and binge drinking. *Addictive Behaviors* 2008;33(1):196–200. [PubMed: 17720324]
- Cunningham JA, Ansara D, Wild TC, Toneatto T, Koski-Jannes A. What is the price of perfection? The hidden costs of using detailed assessment instruments to measure alcohol consumption. *Journal of Studies on Alcohol* 1999;60(6):756–758. [PubMed: 10606486]
- Del Boca FK, Darkes J, Greenbaum PE, Goldman MS. Up close and personal: Temporal variability in the drinking of individual college students during their first year. *Journal of Consulting and Clinical Psychology* 2004;72(2):155–164. [PubMed: 15065951]
- Greenbaum PE, Del Boca FK, Darkes J. Variation in the drinking trajectories of freshmen college students. *Journal of Consulting and Clinical Psychology* 2005;73(2):229–238. [PubMed: 15796630]
- Justice A, Dombrowski E, Conigliaro J, Fultz S, Gibson D, Madenwald T, et al. Veterans Aging Cohort Study (VACS): Overview and description. *Medical Care* 2006;44(8 Suppl. 2):S13–S24. [PubMed: 16849964]
- Litt MD, Cooney NL, Morse P. Ecological momentary assessment (EMA) with treated alcoholics: methodological problems and potential solutions. *Health Psychology* 1998;17:48–52. [PubMed: 9459069]
- Litt MD, Cooney NL, Morse P. Reactivity to alcohol related stimuli in the laboratory and in the field: Predictors of craving in treated alcoholics. *Addiction* 2000;95:889–900. [PubMed: 10946438]
- Maisto S, Conigliaro JC, Gordon AJ, McGinnis KA, Justice AC. An experimental study of the agreement of self-administration and telephone administration of the Timeline Followback interview. *Journal of Studies on Alcohol and Drugs* 2008;69(3):468–471. [PubMed: 18432391]
- Perrine MW, Mundt JC, Searles JS, Lester LS. Validation of daily self-reported alcohol consumption using interactive voice response (IVR) technology. *Journal of Studies on Alcohol* 1995;56(5):487–490. [PubMed: 7475027]

- Rehm J, Greenfield TK, Rogers JD. Average volume of alcohol consumption, patterns of drinking, and all-cause mortality: Results from the US National Alcohol Survey. *American Journal of Epidemiology* 2001;153(1):64–71. [PubMed: 11159148]
- Room R, Graham K, Rehm J, Jernigan D, Monteiro M. Drinking and its burden in a global perspective: Policy considerations and options. *European Addiction Research* 2003;9(4):165–175. [PubMed: 12970585]
- Roy M, Dum M, Sobell LC, Sobell MB, Simco ER, Manor H, et al. Comparison of the quick drinking screen and the alcohol Timeline Followback with outpatient alcohol abusers. *Substance Use & Misuse* 2008;43(14):2116–2123. [PubMed: 18825590]
- Russell M, Light JM, Gruenewald PJ. Alcohol consumption and problems: The relevance of drinking patterns. *Alcoholism: Clinical and Experimental Research* 2004;28(6):921–930.
- Sobell LC, Agrawal S, Annis H, Ayala-Velazquez H, Echeverria L, Leo GI, et al. Cross-cultural evaluation of two drinking assessment instruments: Alcohol Timeline Followback and Inventory of Drinking Situations. *Substance Use & Misuse* 2001;36(3):313–331. [PubMed: 11325169]
- Sobell LC, Agrawal S, Sobell MB, Leo GI, Young LJ, Cunningham JA, et al. Comparison of a Quick Drinking Screen with the Timeline Followback for individuals with alcohol problems. *Journal of Studies on Alcohol* 2003;64(6):858–861. [PubMed: 14743950]
- Sobell LC, Brown J, Leo GI, Sobell MB. The reliability of the Alcohol Timeline Followback when administered by telephone and by computer. *Drug and Alcohol Dependence* 1996;42(1):49–54. [PubMed: 8889403]
- Sobell, LC.; Sobell, MB. Alcohol Timeline Followback (TLFB). In: American Psychiatric Association. , editor. *Textbook of Psychiatric Measures*. Washington, DC: American Psychiatric Association; 2008. p. 477–479.
- Todd M, Armeli S, Tennen H. Interpersonal problems and negative mood as predictors of within-day time to drinking? *Psychology of Addictive Behaviors* 2009;23:205–215. [PubMed: 19586137]
- Toll BA, Cooney NL, McKee SA, O'Malley SS. Correspondence between interactive voice response (IVR) and Timeline Followback (TLFB) reports of drinking behavior. *Addictive Behaviors* 2006;31(4):726–731. [PubMed: 15975732]
- Tolstrup JS, Jensen MK, Tjenneland A, Overvad K, Gronbaek M. Drinking pattern and mortality in middle-aged men and women. *Addiction* 2004;99(3):323–330. [PubMed: 14982545]

Figure 1.

Discrepancies between TLFB-7 and TLFB-30 reports are displayed in the absolute value of the difference per day, averaged across persons. Due to the flexible timing of TLFB-7 assessments, each daily average is based on varying n . Standard errors around means are given. Time is expressed as the length of recall (i.e., number of days between interview date and day of drinking). The grey line represents the statistically significant linear time effect estimated by the Poisson regression model.

Table 1
Correspondence on summary indicators of drinking

	<i>n</i>	TLFB-7 mean (<i>SD</i>)	TLFB-30 mean (<i>SD</i>)	Pearson <i>r</i>	Paired <i>t</i> -test
All interviews					
Total # of drinks †	323	15.3 (17.2)	13.6 (15.4)	0.78	2.16 *
# of days abstinent	323	11.7 (3.5)	12.1 (3.4)	0.89	-4.70 ***
# of days drinking 4+/5+ drinks per day †	323	1.7 (2.1)	1.5 (1.8)	0.78	2.50 **
Avg. # of drinks per "drinking day" (either) †	293	3.7 (2.5)	3.5 (2.8)	0.38	1.90
Avg. # of drinks per "drinking day" (both) †	237	5.2 (3.1)	5.1 (3.2)	0.79	1.65
In-person interviews only					
Total # of drinks †	270	15.7 (17.8)	13.9 (16.0)	0.79	1.97 *
# of days abstinent	270	11.8 (3.5)	12.2 (3.3)	0.88	-4.08 ***
# of days drinking 4+/5+ drinks per day †	270	1.7 (2.1)	1.5 (1.8)	0.78	2.33 *
Avg. # of drinks per "drinking day" (either) †	243	3.8 (2.6)	3.5 (2.9)	0.38	1.72
Avg. # of drinks per "drinking day" (both) †	200	5.2 (3.2)	5.0 (3.3)	0.78	1.86
Phone/Mail interviews only					
Total # of drinks †	53	13.2 (13.7)	11.7 (11.9)	0.73	0.87
# of days abstinent	53	11.4 (3.6)	11.9 (3.6)	0.95	-2.70 ***
# of days drinking 4+/5+ drinks per day †	53	1.5 (1.7)	1.3 (1.4)	0.77	0.90
Avg. # of drinks per "drinking day" (either) †	50	3.6 (2.2)	3.3 (2.2)	0.38	0.80
Avg. # of drinks per "drinking day" (both) †	37	5.3 (2.9)	5.4 (2.8)	0.83	-0.37

Note : Means were first averaged within persons (4-21 obs per person), then across persons; for variables

Table 2
Day-to-day correspondence in determining drinking day by weekday

	n	Length of Recall (in days) M (SD)	Drinking Days		Kappa
			TLFB-7 %	TLFB-30 %	
Most recent weekday					
Monday	315	10.8 (6.1)	10.9	7.9	0.68
Tuesday	317	8.4 (5.3)	11.0	9.5	0.64
Wednesday	320	6.9 (4.5)	13.4	10.0	0.59
Thursday	320	5.9 (4.2)	21.6	15.6	0.66
Friday	320	4.8 (3.8)	36.3	37.2	0.72
Saturday	323	4.0 (4.5)	35.9	36.8	0.67
Sunday	323	2.8 (4.2)	14.6	12.1	0.76
Most distal weekday					
Monday	315	21.6 (5.8)	14.6	7.0	0.42
Tuesday	317	20.4 (5.4)	9.8	7.6	0.38
Wednesday	320	19.5 (5.3)	12.2	10.3	0.56
Thursday	320	18.5 (5.1)	21.3	15.0	0.58
Friday	320	17.3 (5.1)	36.9	33.8	0.52
Saturday	323	27.5 (5.6)	37.8	37.5	0.45
Sunday	323	26.6 (5.6)	11.5	8.4	0.41

Note : “Most recent” were the same as “most distal” weekdays for 4.0% (Sunday) to 23.8% (Monday) of the cases.

Table 3
Day-to-day correspondence in number of drinks per drinking day by weekday

	<i>n</i>	Length of Recall (<i>in days</i>) M (<i>SD</i>)	Drinks per Drinking Day			Pearson's <i>r</i>	Paired <i>t</i> -test
			TLFB-7 M (<i>SD</i>)	TLFB-30 M (<i>SD</i>)			
Most recent weekday							
Monday	21	12.1 (7.2)	5.1 (3.6)	4.8 (3.3)	0.91	0.70	
Tuesday	22	7.2 (5.0)	3.9 (2.7)	4.7 (4.2)	0.48	-0.96	
Wednesday	24	6.1 (3.5)	4.5 (3.1)	4.1 (3.2)	0.90	1.19	
Thursday	43	6.3 (4.2)	5.5 (3.7)	5.0 (3.6)	0.70	1.00	
Friday	97	5.0 (3.8)	6.3 (3.7)	5.6 (3.6)	0.77	2.79 **	
Saturday	93	4.0 (4.6)	5.9 (3.6)	5.8 (3.9)	0.64	0.32	
Sunday	34	1.9 (2.7)	5.5 (4.2)	4.9 (3.9)	0.90	1.88	
Most distal weekday							
Monday	16	23.2 (5.6)	5.3 (2.5)	5.1 (3.5)	0.48	0.16	
Tuesday	12	16.8 (6.9)	3.8 (2.6)	3.5 (2.4)	0.63	0.42	
Wednesday	22	18.9 (6.4)	3.5 (1.9)	3.5 (1.9)	0.63	-0.04	
Thursday	38	20.1 (5.4)	4.4 (2.9)	5.1 (4.3)	0.69	-1.38	
Friday	78	17.3 (5.5)	5.5 (4.0)	5.2 (3.4)	0.49	0.70	
Saturday	80	26.6 (6.3)	5.7 (3.6)	6.3 (4.0)	0.45	-1.36	
Sunday	15	25.9 (7.8)	5.5 (6.5)	4.6 (4.5)	0.90	1.05	

Note : "Most recent" were the same as "most distal" weekdays for *n* = 1 (Sunday) to *n* = 8 (Friday);

**
p<0.01