

Comparing Numbers of Drinks: College Students' Reports From Retrospective Summary, Followback, and Prospective Daily Diary Measures*

MEGAN E. PATRICK, PH.D.,[†] AND CHRISTINE M. LEE, PH.D.[†]

Institute for Social Research, University of Michigan, Ann Arbor, Michigan 48106-1248

ABSTRACT. Objective: Retrospective summary, followback (retrospective diaries), and prospective daily diary measures of alcohol use among college students were compared across 29 days. **Method:** Participants were college students ($n = 176$; 60.2% female). Similarities in the three web-based reporting methods and both between-persons (i.e., gender, past drinking behavior, fraternity/sorority affiliation, average drinking behavior during the study period) and within-person (i.e., daily number of drinks, weekend days, Halloween, and week of study) predictors of concordance between reports of followback and prospective diaries were analyzed. **Results:** On prospective diaries, students reported a greater number of maximum drinks (compared with followback only) and a greater number of heavy drinking days in the past 2 weeks (compared

with both followback and retrospective summary measures). In followback compared with prospective diaries, students tended to provide *inflated* accounts of their drinking behavior when reporting about occasions with greater typical drinking (i.e., weekends, Halloween) and *deflated* accounts of their drinking on their own heavier drinking days, especially if they were affiliated with a fraternity/sorority. Women and students who drank more on average across study days tended to provide *deflated* estimates of their day-to-day drinking in followback compared with prospective diary. **Conclusions:** Understanding the concordance and discordance in self-reported alcohol use is an important area for continued research efforts. (*J. Stud. Alcohol Drugs*, 71, 554-561, 2010)

SELF-REPORTED ALCOHOL USE is the most common form of assessment of drinking behavior. Still, there is a need to better understand the processes involved when respondents report on their alcohol use (Del Boca and Darkes, 2003). Participants can be influenced by a myriad of factors, including their own characteristics and those of the drinking occasion or timing of the assessment. The current study compares three different measures of alcohol use—retrospective summary, followback (retrospective diaries), and prospective daily diaries—with a particular focus on for whom and on what occasions individuals are less likely to have corresponding reports of behavior.

Retrospective summary measures

Asking individuals to summarize their alcohol use and answer a few short quantity-frequency summary measures

is an easy and efficient way to assess behavior. In situations in which alcohol use is not the main focus of the study and respondent burden is a particular concern, these measures likely give very good estimates of behavior (Del Boca and Darkes, 2003; Leigh, 2000). Reliability of retrospective measures of quantity and frequency of drinking was found to be acceptable, for example, when compared with prospective written diaries over 10 weeks (Hilton, 1989). Summary measures do not provide information about variability in drinking behavior that is central to some research questions about alcohol use, however (e.g., on which occasions individuals drink more and what proximally predicts these days of particularly heavy alcohol use).

Retrospective followback reports

Another method involves asking individuals to retrospectively report their alcohol use on each day over a specified timeframe, up to 12 months, using cues such as holidays and other important personal events. Timeline Followback, in particular, is a well-validated retrospective measure of daily drinking (Sobell and Sobell, 1992). Timeline Followback can be reliably administered in person, by telephone, and by computer (Sobell et al., 1996). Followback techniques (also known as retrospective diaries) may tend to yield higher estimates of alcohol consumption, including heavy and peak drinking, compared with retrospective quantity-frequency reports (O'Hare, 1991; Shakeshaft et al., 1999). Some conflicting evidence has been found, however, with followback

Received: September 23, 2009. Revision: January 15, 2010.

*Data collection and manuscript preparation were supported by a grant from the Alcohol and Drug Abuse Institute awarded to Christine M. Lee, University of Washington, Seattle, WA. Manuscript preparation was supported by National Institute on Alcohol Abuse and Alcoholism grant F32 AA017806 awarded to Megan E. Patrick. The content here is solely the responsibility of the authors and does not necessarily represent the official views of the sponsors.

[†]Correspondence may be sent to Megan E. Patrick at the above address or via email at: meganpat@isr.umich.edu. Christine M. Lee is with the Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, WA.

measures estimating lower average drinking compared with retrospective summary measures in at least one study (Midanik et al., 1998).

Both types of retrospective measures—summary and followback—are criticized for being subject to recall bias, which results in reports of less alcohol use as a result of unintentional forgetting, minimizing, or intentional under-reporting (Greenfield and Kerr, 2008). This bias is hypothesized to increase over longer reporting intervals. Gmel and Daepfen (2007) found, for example, that the amount of alcohol reported decreased with the length of recall across a 7-day retrospective diary. Adjusting for the day of the week, participant recall of the number of drinks consumed 7 days prior was nearly a full drink less than recall of the number of drinks consumed only 1 day prior.

Prospective diaries

According to a review by Gmel and Rehm (2004), prospective diary studies yield better estimates of behavior than retrospective summary measures. Prospective diaries, in which participants are asked to report their level of alcohol use each day, are particularly subject to less recall bias. Prospective written diaries are associated with higher estimates of drinking when compared with retrospective quantity-frequency alcohol questionnaires, with heavier drinkers being more likely to underestimate their drinking behavior and lighter drinkers tending to overestimate their behaviors in retrospect (Townshend and Duka, 2002). Furthermore, Lemmens et al. (1992) concluded that prospective diary measures more closely corresponded to the available alcohol sales data when compared with retrospective diary and other summary measures.

In her comprehensive review of methods for daily reporting of alcohol-use behavior, Leigh (2000) found that daily reports tended to result in a larger number of drinking occasions but similar overall quantity consumed, with smaller differences between prospective diaries and followback than between prospective diaries and retrospective summary measures. In one study comparing followback with prospective daily reports of drinking over 28 days, participants reported modestly lower levels of alcohol use (i.e., drinks per day and ounces per day) on followback measures than on daily assessments (Carney et al., 1998). Repeated-measures designs (either followback or prospective diaries) are particularly necessary for understanding patterns of within-person variations in drinking behaviors and for assessing co-occurring behaviors and processes on daily or event levels. The extent to which these two reports produce different estimates is largely unexplored.

Interindividual differences

Some research has been conducted to examine between-person variability in the correspondence of different types of

alcohol-use measures. Carney et al. (1998) found that day-to-day differences in followback and prospective diary reports were not significantly predicted by either demographic (i.e., gender, income, education) or alcohol-use (i.e., average amount consumed, variation) characteristics. Other authors have noted that participants tended to report greater alcohol use on diary measures compared with retrospective summary measures, unless they were among the heaviest drinkers, for whom the reports were more similar (Webb et al., 1990). Although prospective diaries are largely regarded as the most valid overall, retrospective measures may offer a better estimation for individuals with heavier drinking patterns than for lighter drinkers (Lemmens et al., 1992) or sporadic drinkers (Gmel and Daepfen, 2007). It may be that individuals with more variable drinking patterns have more difficulty remembering and recalling their alcohol use, whereas individuals with consistent alcohol-use patterns have less difficulty retrospectively reporting on their drinking. Among college students, it is unknown whether individuals in heavier drinking situations (e.g., those involved in fraternities and sororities) are better or worse at estimating their drinking behavior retrospectively. Fraternity/sorority affiliation has been found to provide a context for heavy drinking during college (McCabe et al., 2005; Park et al., 2009), including an association with such alcohol-related problems as alcohol dependence (Grekin and Sher, 2006), although the short-term measurement implications have yet to be examined.

Characteristics of drinking day

The characteristics of particular days may affect the discrepancies between types of alcohol measures. Regular events (e.g., drinking a glass of wine every evening with dinner) and similar episodes are easier for individuals to remember and report (Del Boca and Darkes, 2003), which suggests that retrospective measures may accurately capture these occasions. Retrospective recall bias on a 7-day followback measure was most pronounced for Fridays and Saturdays (Gmel and Daepfen, 2007). Indeed, weekend days tend to have more variability in alcohol use and represent peak drinking occasions among college students (Del Boca et al., 2004). These weekend-day effects are assessed in the current study. In addition, on holidays (e.g., Halloween in the current study) college students tend to drink more than on other days of the year (Del Boca et al., 2004; Neighbors et al., 2007). The extent to which students retrospectively report more or less alcohol use on known heavier drinking days (including weekends and holidays), and individuals for whom the discrepancies may be greater, are investigated in the current study.

Research questions

The current study was designed to address two related research questions. First, how similar are students' reports

of drinking from three measurement strategies: retrospective summary, retrospective followback, and prospective daily diaries? Second, using multilevel modeling, what predicts inconsistencies in reports from retrospective followback and prospective diary measures? Prospective diary reports are ideal for evaluating followback techniques, because both alcohol-use measures produce daily data on alcohol use (Del Boca and Darkes, 2003). In this study, daily reports from prospective Web-based diaries were compared with retrospective followback reports of alcohol use on the same 29 days to estimate for whom and on what occasions these reports are most similar and dissimilar. Both between-persons (i.e., gender, past drinking behavior, fraternity/sorority affiliation, and average drinking behavior during the study period) and within-person (i.e., daily number of drinks, weekend days, Halloween, and week of study) predictors of concordance between reports of followback and prospective diaries were included. The difference in reporting methods was conceptualized as the dependent variable, in an attempt to ascertain whether certain individuals on certain occasions were more or less likely to underreport their level of alcohol consumption in retrospect.

Method

Participants and procedure

Participants for the present study were recruited from introductory psychology classes at a large Northwest university during the fall quarter. As part of the introductory psychology program, students could participate in research projects in exchange for extra credit. During a designated class hour, students were given the opportunity to complete a large screening survey ($N = 804$ completed) with questions about alcohol use in the last month and daily access to the internet; a brief description of the study was also provided. Students were asked to provide contact information if they wished to be considered for the present study. Just over half (412; 51%) met criteria for the study (i.e., indicated drinking at least one drink in the last month, had daily internet access, and provided contact information). Of these, 300 students were randomly selected to be eligible for telephone recruitment. Recruitment continued until 200 students agreed to participate.

On the first day of the study students were emailed a link to a 45-minute baseline survey, which included demographic questions, past alcohol use and consequences measures, and other psychosocial assessments. Each day for the next 29 days, students were emailed a link to a 5-10 minute daily online questionnaire with measures of mood, alcohol use and related consequences, and daily activities and events. On the final day of the project (Day 31), students were emailed a final link to the study, with a retrospective 30-day followback measure and retrospective summary questions of alcohol use during the past

month. Participants were compensated based on a graduated schedule, receiving up to 3 extra credit hours and \$57 monetary incentive. In addition, participants were entered into a drawing for cash prizes if they completed all days.

Of the 200 students recruited to participate in the study, three did not complete the baseline survey. Among the enrolled 197 participants, 92% of target days were completed (5,620 of 6,107 possible); 69% of students completed all 31 days, and 81% completed 29 or more days. Of the 197 students, 19 reported no drinking during the 29 days of the daily reporting assessment and were excluded from the present analyses. Two additional students were excluded because of missing data on the baseline drinking measure. Thus, the final sample for the multilevel analyses comprised 176 students with a mean age of 18.97 years ($SD = 1.97$); 60.2% were female; 67.4% self-identified as white non-Hispanic, 22.1% as Asian, and 10.5% as other ethnicities. Participants included in the multilevel model did not differ ($p > .05$ for all t tests) from the full sample on age, $t(195) = -0.44$; gender, $t(195) = 0.57$; race, $t(22) = -1.93$; or fraternity or sorority membership, $t(33) = -1.91$. For the comparisons across all three types of measurement, 13 additional students were excluded for completing fewer than 14 days of diary (i.e., insufficient data to calculate diary summary measures), and 2 students were excluded for not completing the retrospective measures on Day 31. Participants ($n = 161$) included in the descriptive comparisons of types of measurement did not differ ($p > .05$ for all t tests) from the full sample on age, $t(195) = -0.37$; gender, $t(195) = -1.87$; race, $t(190) = -0.19$; or fraternity or sorority membership, $t(198) = -0.42$.

Measures

Baseline (Day 1) measures included two demographic variables: gender (0 = *male*, 1 = *female*) and fraternity/sorority affiliation (0 = *no membership*, 1 = *member or pledging*). Typical quantity of past weekend drinking was assessed by asking, "On a given weekend evening during the past month, how much alcohol did you typically drink?" with responses ranging from 0 to 25 or more drinks. These three variables were used as between-person predictors (Level 2 predictors) in the multilevel analyses.

Daily diary measures included prospective diary reports of drinking. Each day for 29 days, students were asked, "How many total standard drinks did you have yesterday?" The daily number of drinks was used in the analyses as a Level 1 variable. Average drinking across days (i.e., the mean of the 29 days of prospective diary reports) for each participant was used as a Level 2 predictor. In addition, two measures were created to describe the sample, including diary maximum number of drinks (i.e., the number of drinks reported on the highest drinking day) and diary heavy drinking in the last 2 weeks (i.e., number of days consuming five or more drinks in the last 2 weeks).

Weekend day was calculated based on the day of week of the report with Thursdays, Fridays, and Saturdays coded as 1 (all other days coded as 0). Halloween occurred for all participants during the fall assessment period and was represented by whether it was (1 = *Sunday, October 31; Day 13 of the study*) or was not (0 = *all other days*) the holiday. Week of study was also computed, such that 1 = *Days 1-7*, 2 = *Days 8-14*, 3 = *Days 15-21*, and 4 = *Days 22-29*.

Postdaily (Day 31) assessment measures included a modified retrospective Timeline Followback (Sobell and Sobell, 1992). Students were presented with a calendar of the previous 30 days and asked to fill in the number of standard drinks they consumed on each day. For the purposes of the present study, only the 29 days matching the prospective daily diary assessments were used. The main outcome variable is the difference between the number of drinks reported for each day on the daily diary and the number of drinks on the corresponding followback day. Two additional variables were created for the purposes of describing the sample's drinking, including followback maximum number of drinks (i.e., the number of drinks reported on the highest drinking day) and followback heavy drinking in the last 2 weeks (i.e., number of days consuming five or more drinks in the last 2 weeks). Discrepancy between prospective diary and followback reports was the outcome variable of the multilevel model and computed to reflect the extent to which students underestimated their drinking on the followback measure (i.e., prospective daily diary report minus followback report).

Retrospective alcohol measures were also assessed in the postdaily (Day 31) assessment after the followback calendar. Retrospective maximum drinks was assessed by asking participants to "Think of the occasion you drank the most this past month. How much did you drink?" with response options from 0 to 25 or more drinks. Retrospective heavy drinking was measured by asking participants to "Think back over the last two weeks. How many times have you had 5 or more drinks in a row?"

Plan of analysis

The first research question focused on the similarity of reports from retrospective summary, retrospective followback, and prospective daily diary measures of alcohol use. Means and standard deviations of reports of maximum drinks and frequency of heavy drinking in the last 2 weeks (i.e., last 14 days of diary and followback) from each of the three measures were compared with paired-samples *t* tests. To address the second research question, multilevel models were conducted to estimate the discrepancies between prospective daily diary and followback reports (i.e., the extent to which followback reports were underestimates of prospective diary reports). Between-persons (Level 2) predictors included gender (γ_{01}), average drinking for the past month at baseline (γ_{02}), affiliation with a fraternity or sorority (γ_{03}),

and the person mean number of drinks across the 29 days of diary data (γ_{04}). Within-person (Level 1) predictors were the number of drinks reported that day from the prospective diary (β_1), whether it was a weekend day (β_2), whether it was Halloween (β_3), and the week of the study (β_4). Interaction effects were also tested. Level 2 variables of gender, average past drinking, and fraternity/sorority affiliation were examined as moderators of each of the Level 1 variables (i.e., daily number of drinks, weekend days, Halloween, and week of study). All interactions for gender and average drinking were nonsignificant. Fraternity/sorority affiliation was found to moderate the effects of daily number of drinks and weekend days (effects are shown in Table 2) but not Halloween or week of study (therefore, these interactions are not presented).

Results

Similarity of reports

The first research question focused on the similarity of reports from the three measures of alcohol use. Table 1 shows the number of drinks reported on a given day by prospective diary and followback measures and the within-person difference (prospective daily diary minus followback) at Level 1 (note that there are no within-person reports from the retrospective summary measures to compare). Prospective diary reports of maximum drinks were not significantly different than retrospective reports of maximum drinks, $t(160) = 0.74$, *n.s.* Both prospective diary reports, $t(160) = 2.37$, $p < .01$, d [effect size] = 0.10, and retrospective reports, $t(160) = 3.09$, $p < .01$, $d = 0.08$, yielded higher means on maximum number of drinks compared with followback reports. In the prospective diaries, students reported a greater number of heavy drinking days compared with followback, $t(160) = 8.49$, $p < .001$, $d = 0.49$, and retrospective summary reports, $t(160) = 8.55$, $p < .001$, $d = 0.59$. There was no difference between followback and retrospective summary reports of number of heavy drinking days, $t(160) = 1.86$, *n.s.*

Interindividual differences: Gender, past drinking, fraternity/sorority affiliation, and average daily alcohol use predicting discrepancy

Multilevel modeling results are shown in Table 2. Positive coefficients indicate that there was a greater discrepancy, such that participants underestimated their drinks consumed in the followback compared with the prospective diary. Between persons (Level 2), the intercept (β_0) indicates that on average, across days, participant prospective diary and followback reports did not significantly differ after taking into account the other variables in the model. Women were more likely to underestimate their drinking in the followback than men (γ_{01}). Individuals who reported more drinking in

TABLE 1. Descriptive statistics comparing types of measurement

| Variable | <i>M</i> | <i>SD</i> | Range |
|---|----------|-----------|----------|
| Drinks per day (Level 1) ^a | | | |
| Diary drinks per day | 0.97 | 2.67 | 0 – 25 |
| Followback drinks per day | 1.08 | 2.59 | 0 – 25 |
| Difference of diary and Followback | -0.13 | 2.40 | -19 – 25 |
| Maximum drinks (Level 2) ^b | | | |
| Diary maximum drinks | 7.45 | 5.00 | 1 – 25 |
| Followback maximum drinks | 6.97 | 4.26 | 1 – 25 |
| Retrospective maximum drinks | 7.30 | 4.37 | 0 – 25 |
| Frequency of heavy drinking (≥5) in last 2 weeks (Level 2) ^b | | | |
| Diary heavy drinking days | 2.54 | 2.79 | 0 – 13 |
| Followback heavy drinking days | 1.41 | 1.73 | 0 – 7 |
| Retrospective heavy drinking occasions | 1.22 | 1.52 | 0 – 10 |

Notes: ^a*n* = 176 students on *N* = 4,579 person-days; ^b*n* = 161.

the past month were less likely to underestimate their drinking in the followback (γ_{02}). Students who drank more during the prospective daily assessment period were more likely to underestimate the number of drinks they consumed in the followback, as compared with the prospective daily diary reports (γ_{04}). Fraternity/sorority affiliation was not a significant between-person predictor of discrepancy in reports.

Characteristics of the drinking day: Daily drinking, weekend days, and Halloween predicting discrepancy

Within-person (Level 1), on days individuals reported consuming more drinks than their own mean in the prospective diary, they were also more likely to underestimate their number of drinks in the followback (γ_{10}). This underestimation was especially pronounced for individuals who were

affiliated with a fraternity or sorority (γ_{11}). On weekends, students tended to overestimate their drinking in the followback, compared with the diary (γ_{20}), especially if they were affiliated with a fraternity or sorority (γ_{21}). In addition, on Halloween, students tended to overestimate their drinking in the followback compared with the diary reports they made for that day (γ_{30}). There were no significant systematic differences in discrepancy by week of the study across the 4 weeks (γ_{40}).

Discussion

On prospective diary reports, students reported a greater number of maximum drinks (compared with followback only) and a greater number of heavy drinking days in the past 2 weeks (compared with followback and retrospective

TABLE 2. Multilevel model predicting the difference of diary and Followback reported number of drinks

| Variable | <i>B</i> (<i>SE</i>) |
|---|------------------------|
| Average over days intercept, β_0 | -0.15 (0.11) |
| Female gender, γ_{01} | 0.22 (0.10)* |
| Average past drinking quantity, γ_{02} | -0.04 (0.01)** |
| Fraternity affiliation, γ_{03} | 0.12 (0.10) |
| Person mean drinks across days, γ_{04} | 0.15 (0.05)** |
| Average fluctuations in daily total drinks, β_1 | |
| Intercept, γ_{10} | 0.46 (0.04)*** |
| Fraternity affiliation, γ_{11} | 0.17 (0.07)* |
| Average fluctuations with weekend, β_2 | |
| Intercept, γ_{20} | -0.19 (0.06)** |
| Fraternity affiliation, γ_{21} | -0.82 (0.19)*** |
| Average fluctuations with Halloween, β_3 | |
| Intercept, γ_{30} | -1.64 (0.26)*** |
| Average fluctuations by week of study, β_4 | |
| Intercept, γ_{40} | 0.03 (0.03) |

Notes: *n* = 176 students on *N* = 4,579 person days.

Level 1: $Y = \beta_0 + \beta_1$ (daily drinks) + β_2 (weekend) + β_3 (Halloween) + β_4 (week of study) + r_{it}

Level 2: $\beta_0 = \gamma_{00} + \gamma_{01}$ (gender) + γ_{02} (average drinking past month) + γ_{03} (fraternity affiliation) + γ_{04} (person mean drinks) + U_0

$\beta_1 = \gamma_{10} + \gamma_{11}$ (fraternity affiliation)

$\beta_2 = \gamma_{20} + \gamma_{21}$ (fraternity affiliation)

$\beta_3 = \gamma_{30}$

$\beta_4 = \gamma_{40}$

p* < .05; *p* < .01; ****p* < .001.

summary measures). Discrepancies between daily prospective diaries and followback reports were found to vary based on characteristics of the person and the drinking occasion. Some students provided *deflated* accounts of their drinking behavior in followback compared with prospective diaries, including women and those who had more to remember (i.e., drank more on average across the study or on that day, in particular). Reporting on drinking occasions that occur frequently requires a large amount of recall effort (Del Boca and Darkes, 2003). Therefore, individuals who drank more during the study period, when controlling for the amount they consumed in the past, had more to recall when completing the assessments. This may also reflect the fact that it is more difficult for sporadic drinkers (i.e., who engaged in one or more very heavy drinking events) to recall their behavior (Gmel and Daepfen, 2007). Similar to Carney et al. (1998), we found significant between-person variability in the correspondence of the measures. Contrary to previous research, however, the day-to-day differences in diary and followback reports in our study were significantly predicted by gender and previous alcohol use. The contrary findings may result from differences between the two studies, including differences in sample eligibility, the younger age and college status of our participants, and our larger sample size.

Characteristics of the drinking day were also predictors of the discrepancies between reports. Specifically, students tended to provide *inflated* accounts of their drinking behavior, in followback compared with prospective diaries, when they were reporting about occasions when students typically drink more (i.e., weekends and Halloween). This may reflect students relying on typical patterns or heuristics when estimating their drinking behavior in retrospect, rather than a true memory of the drinking event. In addition, students who reported more alcohol in the past (before the study) tended to *overestimate* their behavior during the 29 days in followback compared with prospective diary. These individuals may believe themselves to be heavier drinkers than they are in reality.

Retrospective reports of drinking and followback reports showed no difference on number of heavy drinking days, but students reported significantly greater numbers of maximum drinks on retrospective reports compared with followback reports. Although some studies show that followback techniques tend to yield higher estimates of alcohol consumption (including heavy and peak drinking) compared with retrospective quantity-frequency reports (O'Hare, 1991; Shakeshaft et al., 1999), there are conflicting findings that followback measures yield lower drinking estimates compared with retrospective summary measures (Midanik et al., 1998). Further research regarding when and for whom different measurement strategies yield discrepant findings is needed.

Although daily reports are sometimes assumed to be more accurate, comparing data from retrospective followback and

from prospective diaries is simply a means of establishing the level of *intermethod reliability* (Leigh, 2000). We do not make assumptions about the infallibility of diary reports. In fact, diaries suffer from the same limitations as other methods of self-reported alcohol use, including variation in what is called "a drink," losing count of quantity consumed as drinks are refilled, or becoming too intoxicated to remember (Leigh, 2000). The definition of a drink is sometimes unclear to participants: Across countries and cultures, the definition of a drink varies from 8 to 20 g of pure ethanol (Gmel and Rehm, 2004), for example. In addition, fatigue or disengagement from the research protocol as a result of repeatedly answering the same questions each day are potential limitations for diary assessments. Biological methods (e.g., breath alcohol testing) and observational studies of the number and type of drinks consumed are more appropriately designed to measure the validity of all types of self-reports.

Implications

Given that some individuals either underestimate or overestimate their behavior in retrospect, some college students may benefit from behavioral monitoring as a sort of reality check. Fraternity and sorority members, for example, tend to provide inflated accounts of their heavy drinking occasions in retrospect, which may be the (erroneous) basis for believing that they can or should drink that same amount in the future. These overestimates may be the result of telling and retelling stories from these drinking episodes within a context that supports heavy alcohol use (Park et al., 2009). These events may then be remembered as more positive or extreme than they were in actuality and may lead to more extreme future plans (e.g., see Morewedge et al., 2005). Many clinical interventions for alcohol problems incorporate a monitoring component (e.g., Brief Alcohol Screening and Intervention for College Students, or BASICS; Marlatt et al., 1998), and highlighting this type of discrepancy between daily reports and what is remembered up to 1 month later may provide an additional teaching point. Research has documented inflated social norms for college students, such that they believe other students drink more than they do in reality (e.g., Borsari and Carey, 2003). This study suggests that some college students may also have inflated *personal* norms, or inaccurate beliefs about their own behavior.

Limitations

This study is subject to several limitations. First, completing a prospective diary may increase memory for drinking occasions, and therefore inflate the concordance between prospective and followback and retrospective reports (Leigh, 2000). In addition, in this study the followback calendar was completed before the retrospective reports, possibly increasing recall accuracy on retrospective reports. Therefore, the

discrepancy estimates are conservative. Randomly assigning participants to receive either the followback calendar or the retrospective reports first would be worthwhile to examine potential ordering effects and implications for findings. Second, the definition of heavy drinking used in the diary and followback measures was that of consuming five or more drinks on a single day and the sum of these days across 2 weeks, whereas the retrospective report defined heavy drinking as the number of times five or more drinks were consumed in a row in the last 2 weeks (a definition used for comparison with other studies; e.g., Johnston et al., 2008). Although we attempted to compare across the types of measurement as accurately as possible, we recognize it as a limitation that these definitions do not exactly correspond, as well as that these definitions do not match other widely used definitions of heavy episodic use (e.g., four or more drinks for women and five or more drinks for men in about 2 hours; National Institute on Alcohol Abuse and Alcoholism, 2003). Although we are limited by knowing only how many drinks were consumed in a 24-hour period for the diary and followback reports, we believe comparisons across prospective daily diary reports and followback reports at the daily level are fairly comparable with retrospective measures using these definitions. Third, in the followback protocol, only main school and cultural events were listed (e.g., first day of school or Halloween); students were not provided the opportunity to enter personally important events into the calendar. Several other limitations include issues of generalizability. Students in the present study were recruited from introductory psychology classes and the sample was primarily White and Asian American; in addition, the majority of the participants were in their first or second year of college. Thus, these participants had relatively recent drinking histories and these results may not generalize to individuals who have more established drinking patterns, the broader college student population, and to other universities.

Future directions

There are costs and benefits to all methods of data collection. In the interest of making the most efficient use of resources, various measures of drinking behaviors should match the research questions of interest as much as possible (Del Boca and Darkes, 2003; Leigh, 2000). Retrospective summary measures are appropriate for studies attempting to estimate an average level of alcohol use. Followback studies may help inform researchers about general patterns of use (e.g., whether some participants engage in heavy use on most weekend days). For some research questions, however—including those concerned with real-time predictors, correlates, and consequences of alcohol use, involving other variables that vary within person across days—prospective diaries may be most appropriate. Furthermore, understanding

for whom different reports are most likely to be inaccurate is an important area for continued research efforts.

References

- Borsari, B., & Carey, K. B. (2003). Descriptive and injunctive norms in college drinking: A meta-analytic integration. *Journal of Studies on Alcohol, 64*, 331-341.
- Carney, M. A., Tennen, H., Affleck, G., Del Boca, F. K., & Kranzler, H. R. (1998). Levels and patterns of alcohol consumption using timeline follow-back, daily diaries and real-time "electronic interviews." *Journal of Studies on Alcohol, 59*, 447-454.
- 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. No. 2), 1-12.
- Del Boca, F. K., Darkes, J., Greenbaum, P. E., & Goldman, M. S. (2004). Up close and personal: Temporal variability in the drinking of individual college students during their first year. *Journal of Consulting and Clinical Psychology, 72*, 155-164.
- Gmel, G., & Daepfen, J. (2007). Recall bias for seven-day recall measurement of alcohol consumption among emergency department patients: Implications for case-crossover designs. *Journal of Studies on Alcohol and Drugs, 68*, 303-310.
- Gmel, G., & Rehm, J. (2004). Measuring alcohol consumption. *Contemporary Drug Problems, 31*, 467-540.
- Greenfield, T. K., & Kerr, W. C. (2008). Alcohol measurement methodology in epidemiology: Recent advances and opportunities. *Addiction, 103*, 1082-1099.
- Grekin, E. R., & Sher, K. J. (2006). Alcohol dependence symptoms among college freshmen: Prevalence, stability, and person-environment interactions. *Experimental and Clinical Psychopharmacology, 14*, 329-338.
- Hilton, M. E. (1989). A comparison of a prospective diary and two summary recall techniques for recording alcohol consumption. *British Journal of Addiction, 84*, 1085-1092.
- Johnston, L. D., O'Malley, P. M., Bachman, J. G., & Schulenberg, J. E. (2008). *Monitoring the Future national survey results on drug use, 1975-2007. Volume I: Secondary School Students* (NIH Publication No. 08-6418A), Bethesda, MD: National Institute on Drug Abuse.
- Leigh, B. C. (2000). Using daily reports to measure drinking and drinking patterns. *Journal of Substance Abuse, 12*, 51-65.
- Lemmens, P., Tan, E. S., & Knibbe, R. A. (1992). Measuring quantity and frequency of drinking in a general population survey: A comparison of five indices. *Journal of Studies on Alcohol, 53*, 476-486.
- McCabe, S. E., Schulenberg, J. E., Johnston, L. D., O'Malley, P. M., Bachman, J. G., & Kloska, D. D. (2005). Selection and socialization effects of fraternities and sororities on US college student substance use: A multi-cohort national longitudinal study. *Addiction, 100*, 512-524.
- Marlatt, G. A., Baer, J. S., Kivlahan, D. R., Dimeff, L. A., Larimer, M. E., Quigley, L. A., ... Williams, E. (1998). Screening and brief intervention for high-risk college student drinkers: Results from a 2-year follow-up assessment. *Journal of Consulting and Clinical Psychology, 66*, 604-615.
- Midanik, L. T., Hines, A. M., Barrett, D. C., Paul, J. P., Crosby, G. M., & Stall, R. D. (1998). Self-reports of alcohol use, drug use and sexual behavior: Expanding the timeline follow-back technique. *Journal of Studies on Alcohol, 59*, 681-689.
- Morewedge, C. K., Gilbert, D. T., & Wilson, T. D. (2005). The least likely of times: How remembering the past biases forecasts of the future. *Psychological Science, 16*, 626-630.
- National Institute on Alcohol Abuse and Alcoholism. (2003). *Task Force on Recommended Alcohol Questions - National Council on Alcohol Abuse and Alcoholism recommended sets of alcohol consumption questions - October 15-16, 2003*. Bethesda, MD: Author. Retrieved from <http://www.niaaa.nih.gov/Resources/ResearchResources/TaskForce.htm>

- Neighbors, C., Walters, S. T., Lee, C. M., Vader, A. M., Vehige, T., Szigethy, T., & DeJong, W. (2007). Event-specific prevention: Addressing college student drinking during known windows of risk. *Addictive Behaviors, 32*, 2667-2680.
- O'Hare, T. (1991). Measuring alcohol consumption: A comparison of the retrospective diary and the quantity-frequency methods in a college drinking survey. *Journal of Studies on Alcohol, 52*, 500-502.
- Park, A., Sher, K. J., Wood, P. K., & Krull, J. L. (2009). Dual mechanisms underlying accentuation of risky drinking via fraternity/sorority affiliation: The role of personality, peer norms, and alcohol availability. *Journal of Abnormal Psychology, 118*, 241-255.
- Shakeshaft, A. P., Bowman, J. A., & Sanson-Fisher, R. W. (1999). A comparison of two retrospective measures of weekly alcohol consumption: Diary and quantity/frequency index. *Alcohol and Alcoholism, 34*, 636-645.
- Sobell, L. C., Brown, J., Leo, G. I., & Sobell, M. B. (1996). The reliability of the Alcohol Timeline Followback when administered by telephone and by computer. *Drug and Alcohol Dependence, 2*, 49-54.
- Sobell, L. C., & Sobell, M. B. (1992). Timeline follow-back: A technique for assessing self-reported alcohol consumption. In R. Z. Litten & J. P. Allen (Eds.), *Measuring alcohol consumption: Psychosocial and biochemical methods* (pp. 41-72). Totowa, NJ: Humana Press.
- Townshend, J. M., & Duka, T. (2002). Patterns of alcohol drinking in a population of young social drinkers: A comparison of questionnaire and diary measures. *Alcohol and Alcoholism, 37*, 187-192.
- Webb, G. R., Redman, S., Sanson-Fisher, R. W., & Gibberd, R. W. (1990). Comparison of a quantity-frequency method and a diary method of measuring alcohol consumption. *Journal of Studies on Alcohol, 51*, 271-277.