

Capturing the Moment: Innovative Approaches to Daily Alcohol Assessment

Dan J. Neal, Kim Fromme, Frances K. Del Boca, Kathleen A. Parks, Linda P. King, Ann M. Pardi, R. Lorraine Collins, Mark Muraven, Charlene Vetter, and William R. Corbin

This article is a summary of a symposium presented at the 2005 Research Society on Alcoholism annual conference organized by Dan J. Neal and chaired by William R. Corbin. Event-level data, wherein each "event" (e.g., day) is captured as its own data point, capture the complex patterns of drinking and other high-risk behaviors in ways that the typical aggregate approach cannot. Because of their richness, methodologies that incorporate event-level data are becoming more common in alcohol research. At least 3 distinct forms of event-level data can be gathered: retrospective data (those collected on a single occasion, using memory aids to help each participant reconstruct all drinking events over a specific period of time), daily monitoring data (reporting on all events for that day), and momentary assessment (those recorded immediately following a drinking event or in response to a prompt from researchers). The goal of this symposium was to address many issues associated with event-level methodology, as well as demonstrate projects that are currently implementing such innovative data collection. The 4 presentations included in this symposium were "Realizing the Promise and Avoiding the Pitfalls of Retrospective Daily Estimation Assessments of Alcohol Use" by Frances K. Del Boca; "Using Interactive Voice Response Technology to Assess the Alcohol-Victimization Link" by Kathleen Parks, Linda King, and Ann Pardi; "Methodological Issues in Using Personal Data Assistants to Self-monitor Alcohol Consumption" by R. Lorraine Collins, Mark Muraven, and Charlene Vetter; and "Collecting Event-level Data Using the World Wide Web" by Dan J. Neal and Kim Fromme.

Key Words: Event-Level Timeline Followback, EMA, Self-monitoring.

ACCURATE CHARACTERIZATION OF the relation between alcohol use, its antecedents, and its consequences is a challenge given the dynamic interplay among these variables across time. Previous research has typically assessed aggregate drinking data (e.g., number of drinking days, number of drinks consumed) utilizing heuristically based drinking indices such as quantity/frequency or the Daily Drinking Questionnaire (Collins et al., 1985). These measures assess consumption over long periods of time, ranging from 1 month to 1 year. Although these indices have considerable utility in a wide variety of contexts, they are limited in the conclusions that can be drawn from them. For example, although one may be able

to conclude that "individuals who typically drink heavily also typically engage in more high-risk sexual behaviors" one cannot conclude that "on days that individuals drink heavily they engage in high-risk sexual behaviors." Thus, in the aggregate, statements may only be made about behavior tendencies of individuals, as opposed to the specific behaviors of interest.

Recognition of the limitations of existing approaches yielded a search for methodologies that would provide greater understanding of the antecedents and consequences of alcohol use. One such approach is event-level data collection. With event-level data, each day or event is captured as an independent data point. Thus, event-level data capture the complex patterns of drinking and other high-risk behaviors in ways that aggregate data cannot. They provide a rich context for more diverse substantive research regarding frequency, quantity, and variability in alcohol use patterns and the covariation between alcohol use and associated behaviors.

Until recently, collection of event-level data was conceptually appealing but practically difficult. Recent advances in technology have enhanced the ability of researchers to collect valid event-level data and analyze the data in ways that capitalize on the richness of the data.

Event-level data may be collected in a variety of ways each of which represents a substantial advance beyond the collection of aggregate data. Early efforts to collect more ecologically valid data utilized the timeline follow-back

From the University of Texas at Austin, Austin, Texas (DJN, KF); the University of South Florida, Tampa, Florida (FKDB); the Research Institute on Addictions, Buffalo, New York (KAP, LPK, AMP, RLC, MM, CV); the University of Albany, Albany, New York (RLC, MM, CV); and Yale University, New Haven, Connecticut.

Received for publication September 30, 2005; accepted October 10, 2005.

This research was supported by grants NIAAA R01-AA013967 (KF), NIAAA 5T32-AA07471 (RG), NIAAA/NIH AA07595 (RLC), NIAAA R01-AA008333 (MSG), R01-AA011925 (MSG), NIAAA R01-AA013986 (KAP), and R01-AA014499 (KAP).

Reprint requests: Dan J. Neal, PhD, Department of Psychology, The University of Texas at Austin, 1 University Station A8000, Austin, TX 78712-0187; Fax: 512-471-5935; E-mail: neal@psy.utexas.edu

Copyright © 2006 by the Research Society on Alcoholism.

DOI: 10.1111/j.1530-0277.2006.00025.x

procedure (Sobell and Sobell, 2000) or related measures such as the Form 90 (Miller, 1996). These methods allowed for estimation of quantity of alcohol consumption, as well as risk behaviors on both drinking and nondrinking days. Prospective methods have also been used to assess daily alcohol use via paper-and-pencil daily drinking diaries and, more recently, through the use of advanced technology including palm-pilot or personal data assistant (PDA) assessments, interactive voice response (IVR), cell phones, and the Internet.

With respect to analysis of event-level data, approaches such as latent growth curve analysis, growth mixture modeling, hierarchical linear modeling (HLM), and generalized estimating equations now allow for analysis of the particular event and unlike more traditional analyses (e.g., multiple regression or analysis of variance) do not require aggregating the variables and thus losing the richness of the data. These methods also allow for analysis of temporal relations between events, both within days and between days. In short, they allow for the dynamic relations among the variables to emerge in a more realistic manner.

Although event-level data afford considerable benefits to researchers, collection of this type of data also requires that attention be given to the potential obstacles. With retrospective methods, a primary concern is the imperfect nature of human memory. Although this problem is less likely with prospective methods, other potential difficulties exist. Prospective methods may place significant burden on research participants leading to concerns about compliance with the monitoring protocol and repeated measurement of alcohol use and other behaviors also leads to concern about the potential for reactivity. Finally, although more sophisticated data analytic approaches are now available in more user-friendly interfaces, management, and analysis of event-level data remains a complex and time consuming endeavor.

The following 4 presentations discuss event-level research in further detail, with specific details regarding research studies that have implemented event-level data collection.

REALIZING THE PROMISE AND AVOIDING THE PITFALLS OF RETROSPECTIVE DAILY ESTIMATION ASSESSMENTS OF ALCOHOL USE

Frances K. Del Boca

Roughly 2 decades have passed since the introduction of the Timeline Followback procedure for estimating daily alcohol consumption (TLFB: Sobell et al., 1986; Sobell and Sobell, 2000). Since then, adaptations of the basic TLFB have been developed (e.g., Form 90: Miller, 1996), and retrospective daily estimation (RDE) techniques have become one of the primary means of measuring drinking behavior. Respondents are presented with calendars depicting an assessment window of a specified duration (e.g., previous month, past 90 days) and, using a variety of memory aids, indicate the amount of alcohol (most often,

the number of standard drinks) consumed on each day during the period. Despite the richness of the information produced, investigators have most often used the daily drinking data generated by RDE methods to compute global summary measures of the quantity and frequency of alcohol use. As argued here, however, the daily records themselves are potentially useful for addressing a wide range of clinical and research questions regarding the temporal patterning of drinking behavior.

The Promise: Potential Applications and Illustrative Examples. Although prospectively gathered data are regarded as more valid than drinking estimates obtained retrospectively, prospective data collection is often simply not feasible (e.g., baseline or longer term follow-up assessments in treatment outcome studies). Further, as indicated in other contributions to this article, numerous technical issues, as well as considerable expense, may be encountered in implementing a prospective protocol. Relative to such approaches, the typical administration mode for RDE methods provides considerable control over the assessment process (e.g., interviewers can assure respondent sobriety, answer questions, and clarify responses). Evidence also suggests that prospective daily estimation procedures can be reactive (e.g., Helzer et al., 2002). Finally, both prospective and retrospective assessments often share some of the same potential weaknesses (e.g., misuse of the standard drink metric; memory lapses caused by blackouts) (see also Leigh, 2000). Thus, as with all assessment techniques, both prospective and retrospective methods have strengths and limitations, and each may be more or less appropriate depending on the specific research question.

Retrospective daily estimation methods have numerous advantages when compared with assessments composed of questions regarding typical (or usual) quantities and frequencies of drinking. Because respondents are asked to recall specific drinking episodes, rather than to formulate a summary description of their behavior, RDE procedures capture atypical drinking experiences and may be less subject to the cognitive heuristics involved in forming summary judgments (e.g., overreliance on salient or recent events). The data produced are extremely versatile, permitting the creation of a wide range of summary measures for differing time frames as well as time-to-event and event duration variables. They allow an analysis of *patterns* of drinking and changes in consumption as a function of temporal events; idiographic (within-person), as well as aggregate, data analyses are possible. Similarly, RDE data invite the application of a variety of statistical procedures, including standard regression and ANOVA, time-series analysis, latent growth curve modeling, HLM, and event history analysis. If data regarding the duration of drinking episodes are also obtained, blood alcohol concentrations can be estimated. Additionally, it may be possible to collect information regarding the circumstances associated with episodes of alcohol use (e.g., events, activities, others present, other substance use).

Retrospective daily estimation procedures require considerable time and effort from both the respondent and the investigator, and the resulting detailed drinking records are not necessary to address many research questions. Such data are, nevertheless, uniquely well suited for many research purposes. In treatment-outcome studies, the versatility of RDE data can serve many aims. Such data permit the construction of measures that can be specifically tailored to interventions that are thought to achieve their results through different mechanisms of action. For example, the adoption of the Form 90 in the COMBINE trial enabled investigators to compute different criterion variables for the 2 pharmacotherapies under study (naltrexone and acamprosate; Anton and Randall, 2005), an approach which increases the likelihood of detecting treatment effects (Stout, 2003). Similarly, the flexibility of RDE data permits construction of outcome variables that can facilitate comparisons of results across studies with differing follow-up durations or primary outcome measures. Further, the use of RDE procedures in treatment trials is consistent with a conceptualization of outcome in terms of process, rather than as a static endpoint, and the resulting data base allows examination of the alternating periods of abstinence and relapse that tend to characterize postintervention drinking (see also Sobell et al., 2003).

The insights afforded by daily estimation data can be further illustrated by an investigation of factors affecting college student drinking. Participants were 301 freshmen recruited over 3 years (48% male; M age = 18.4 years) with a 95% retention rate. Timeline Followback interviews were completed monthly, and assessment periods were linked to produce a continuous, synchronized daily drinking record for the academic year (see Del Boca et al., 2004, for a complete description of procedures). These records revealed that most alcohol use occurred on Thursdays, Fridays, and Saturdays; consequently, weekly quantity estimates were computed to examine fluctuations over time. As shown in Fig. 1 (which includes the data from those who abstained during the entire period), alcohol use varied from week to week in a pattern that replicated across the study's 3 years: The Fall semester began with relatively high levels of drinking; elevations were observed for national holidays (e.g., Christmas), local celebrations (e.g., Guavaween), and Spring Break; and drinking tended to decline during midterm and final exam weeks. A latent growth curve fit to the weekly data included intercept, slope, and "holiday" parameters, and gender, residence, and alcohol expectancies measured at baseline significantly predicted these parameters in a model that controlled for other risk factors (Del Boca et al., 2004). Although drinking was influenced by temporal events, considerable individual variation was also observed. Latent growth curve mixture modeling provided evidence of 5 distinct classes of drinkers, which were differentiated by gender and alcohol expectancies (Greenbaum et al., 2005). Future

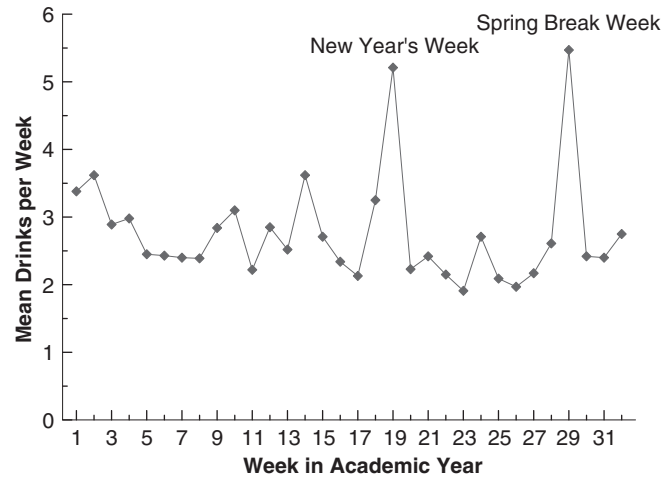


Fig. 1. Weekly quantity estimates over the course of the academic year.

Note: This figure is adapted with permission by APA from page 160 of Del Boca FK, Darkes J, Greenbaum PE, Goldman MS (2004) Up close and personal: temporal variability in drinking during the first year of college. *J Consult Clin Psychol* 72:155–164.

analyses will examine the dynamic interplay between drinking and other variables over time.

This illustrative study demonstrates the feasibility of collecting daily drinking data from college students using RDE procedures. Although limited to a single university, the results have methodological and policy implications. For example, the data suggest that college drinking prevalence estimates may be influenced by the timing of surveys and that intervention studies should be sensitive to natural variations in alcohol use. Regarding prevention, the contingency-driven nature of college drinking (consumption is heaviest during periods of low academic demand) suggests that a restructuring of course requirements might result in drinking declines.

Potential Pitfalls. Several caveats regarding the use of RDE methods should be noted. First, there are validity concerns that apply to all self-report measures of alcohol use (Del Boca and Darkes, 2003; Del Boca and Noll, 2000). The available literature suggests, however, that self-report accuracy can be enhanced by attention to data collection circumstances (e.g., confidentiality of responses, minimization of perceived negative contingencies).

Second, to ensure comparability across participants and studies, most RDE methods instruct respondents to convert the amounts of alcohol consumed into standard drink units, a practice associated with several potential pitfalls (e.g., drinks vary in size and source; they may be partially consumed or shared; beverage brands and proofs are frequently unknown to respondents) (e.g., Leigh, 2000). To reduce estimation errors, investigators should consider the use of props (e.g., sample beverage containers) and the recording of specific information (e.g., beverage brand) that can later be used in the computation of standard drink equivalents (see Miller, 1996).

Third, although the reliability and validity of *summary* measures based on RDE data have been demonstrated

(e.g., Tonigan et al., 1997), the potential utility of *daily* drinking records depends on the accuracy of these estimates. Comparisons between retrospective and prospective results (see Leigh, 2000, for a review) suggest that, although summary indices tend to be highly correlated, prospective methods often yield larger numbers of drinking episodes. Moreover, retrospective reports tend to vary by drinker type: frequent imbibers tend to underestimate, whereas infrequent drinkers tend to overestimate, consumption. Discrepancies may also occur with respect to the precise synchrony of the daily estimates.

The nature of the observed inconsistencies between prospective and retrospective methods suggests that they may be reduced by greater attention to the cognitive demands of RDE assessments. Although more studies are needed that examine directly the cognitive and motivational processes that operate during RDE interviews, the existing research on human memory suggests techniques that can be used to facilitate accurate recall. For example, because autobiographical memories are not easily cued by calendar dating, interviewers should, in addition to eliciting personal milestones, encourage respondents to employ multiple retrieval cues when reconstructing memories of specific drinking experiences. Additionally, ample time should be provided for completion of drinking histories, a factor that should be considered in determining assessment frequency and time frames (see Del Boca and Noll, 2000, for a fuller discussion).

Conclusion. Despite their history of use in alcohol research, the data produced by RDE methods have not been fully exploited. Retrospective daily estimation data records hold considerable promise for enhancing our understanding of the processes of recovery and relapse in treatment research and for increasing our knowledge regarding the natural patterning of alcohol use over time. Concerns regarding data accuracy can be addressed, at least in part, by administration procedures that are informed by relevant research and implemented by well-trained interviewers.

USING INTERACTIVE VOICE RESPONSE TECHNOLOGY TO ASSESS THE ALCOHOL-VICTIMIZATION LINK

Kathleen A. Parks, Linda P. King, and Ann M. Pardi

Interactive voice response (IVR) represents state-of-the-science technology for measuring daily alcohol consumption. Interactive voice response has been shown to be an efficient and accurate means of collecting data on daily alcohol consumption (Mundt et al., 1995; Perrine et al., 1995; Searles et al., 1995, 2000, 2002). Indeed, IVR measurement of alcohol use has been shown to be more accurate than retrospective calendar measurement (e.g., TLFB) of daily alcohol use (Perrine et al., 1995). Interactive voice response is easy and quick (< 5 min/d) for participants to use, is minimally sensitive to impairment because of alcohol, and does not appear to be sensitive to impairment as a result of fatigue (Mundt et al., 1997). Interactive voice re-

sponse has been used successfully to measure other behaviors, such as binge eating (Bardone et al., 2000).

Most people are familiar with IVR through everyday use of the touch-tone keypad of their telephone to navigate automated systems for banks and pharmacies. In 2 ongoing studies reported here, IVR is being used to assess the temporal relationship between alcohol use and victimization among community and college women. Women provide information about their mood, alcohol use, and victimization experiences through daily telephone calls to an automated IVR system. To our knowledge, these are the first applications of IVR data collection for assessing victimization experiences.

Study 1

Methods. In Study 1 we are assessing the relationship between alcohol consumption and sexual victimization and risk taking associated with the drinking context of bars. Young (18–30 years), single women who report drinking in bars at least weekly are recruited through advertisements in the major local news and entertainment papers. At the present time, we have screened 225 women for the study. Among those women, 40% ($n = 91$) have been eligible to participate and 50 (55%) have entered the study. Women who complete 4 or more weeks of the protocol will be retained for later data analyses.

The protocol for this study involves an initial in-depth, face-to-face interview, followed by 12 weeks of daily calls to the IVR system and 3 monthly interviews about any sexual aggression or risk taking reported in the daily calls. We will only be discussing the daily reporting through IVR in the current summary. To encourage compliance with daily reporting we employ a compensation strategy that is similar to that used by Searles et al. (2000), whereby participants receive compensation for daily reports (i.e., \$1.00/day completed) and a maximum bonus for 7 consecutive days (i.e., \$10) but are not penalized for missing a day. When a participant misses an IVR report, she is able to provide the missed report for the preceding day via the automated system. If she has missed more than 1 day of reporting, she is called (or calls the project) and provides an oral report for the missing days. Providing missing reports reinstates the majority (i.e., \$7) of the bonus for the 7-day period.

Women are asked to call the automated system once during each 24-hour period (midnight–11:59 PM). Women are asked to report on their current mood through 6 questions measured on a 5-point scale ranging from not at all to extremely (e.g., *How sad are you feeling today?*). They are then asked about the amount and location of any alcohol or other drug use as *this time yesterday*. They also are asked about any sexual activity (consensual and non-consensual) and risky sexual behavior (i.e., whether a condom or dental dam was used) they have engaged in since *this time yesterday*.

Results. Fifty women have entered the study since April 2005. Currently, 21 (42.0%) women are active in the protocol, 16 (76.2%) have completed 4 or more weeks. Fifteen (30.0%) women completed the full 12-week protocol, 6 (12.0%) women dropped out of the study after completing 4 or more weeks, and 8 (16.0%) women dropped out of the study before completing 4 weeks of the protocol. The average age of participants was 22.7 years (SD 2.3). The majority of women were European American (64.8%), with 16.2% African American, 10.8% Asian American, and the remaining 8.2% of another or mixed ethnicity.

The following are preliminary data for the 37 women who have completed a minimum of 4 weeks of the protocol. Daily calling compliance has been excellent (approximately 95%). The majority (74.3%) of daily reports have been provided on time. Of the 25.7% of reports provided late, 18.2% were reported 1 day late and 7.5% were reported 2 to 4 days late.

Women reported drinking on 40.8% of the report days, and 50.5% of these drinking days involved drinking in a bar. Women reported sexual activity on 16.9% of days. Sexual activity has occurred more often on drinking (22.4%) versus nondrinking days (13.2%; $\chi^2 = 32.4$, $df = 1$, $p < 0.001$). However, women were equally as likely to engage in sexual activity after drinking in a bar (22.1%) as they were after drinking in some other location (22.3%).

Study 2

Methods. In Study 2 we are assessing the relationship between alcohol consumption and victimization (verbal, sexual, and nonsexual physical) among college women across 4 years. All women entering a large public university in Western New York state were invited to participate in a brief web survey during the Fall semester 2004. A total of 996 women participated in the web survey. A subsample of 204 of the women who participated in the Web survey participated in the IVR portion of the study during the Spring 2005 semester.

The IVR portion of the study involved an initial interview and IVR training session, followed by 8 weeks of daily calls to the IVR system, and 2 brief Web surveys (at 4 and 8 weeks). Women who reported 1 or more aggressive incidents during daily reporting were recruited for in-person interviews to gather detailed information on the incidents. We will only be presenting the data from the daily IVR reporting. Women were compensated for their participation in daily calls in the same way as Study 1, with the exception that women received \$0.50 for each daily report rather than \$1.

Women were asked to call the automated system once during each 24-hour period (midnight–11:59 PM). As in Study 1, they were asked to report on their current mood through 6 questions measured on a 5-point scale ranging from not at all to extremely (e.g., *How sad are you feeling*

today?). They were then asked about the quantity, location, and context of any alcohol or other drug use since calling on the previous day. They also were asked about any aggression they had experienced. They were asked to report on the type of aggression (verbal, physical nonsexual, or physical sexual), as well as the gender, substance use, and relationship to the other person(s) involved in the incident.

Results. Of the women who began the protocol, 95.6% ($n = 195$) completed the full 8 weeks. The average age was 18.7 years (SD 1.9). The majority of the women were of European descent (61.2%), with 14.9% of African descent, 13.4% of Asian descent, and the remaining 10.5% of another or mixed ethnic origin. Daily calling compliance was excellent (93.2%). The majority (82.4%) of daily reports were provided on time. Of the 17.6% of reports provided late, 76.5% were reported 1 day late and 23.5% were reported 2 to 4 days late.

Women drank on 6.6% of the reported days and experienced aggression on 1.0% of the reported days. Aggression occurred on 2.7% of drinking days and on 0.9% of nondrinking days. Preliminary HLM analyses indicate that women were significantly more likely to experience aggression on drinking compared with nondrinking days (odds ratio (OR) 2.5, confidence interval (CI) 1.7, 3.7).

Discussion. It has been well established that IVR can be used effectively to collect daily data on alcohol use (e.g., Mundt et al., 1995; Perrine et al., 1995; Searles et al., 2000, 2002). Although IVR has been used to measure other outcome variables (e.g., binge eating; Bardone et al., 2000), these are the first studies that have used IVR to assess daily substance use, victimization, and risky sexual behavior.

Our preliminary findings serve to highlight some of the advantages in using IVR technology for daily data collection. The IVR system is quick and easy for participants to use, allows flexibility in the time of participation, eliminates interviewer bias, limits retrospective recall bias, and provides privacy for participants when answering sensitive questions.

Compliance for our studies is very good, suggesting that participant burden is not excessive and can be balanced with appropriate incentives. In addition, these early findings suggest that daily data collection through the use of IVR will provide valuable insight into the role of alcohol use in related risky sexual behavior and experiences with victimization.

METHODOLOGICAL ISSUES IN USING PERSONAL DATA ASSISTANTS TO SELF-MONITOR ALCOHOL CONSUMPTION

R. Lorraine Collins, Mark Muraven, and Charlene J. Vetter

Use of PDAs for Ecological Momentary Assessment Ecological Momentary Assessment. Ecological Momentary Assessment (EMA) involves intensive, real-time, data collection focused on momentary changes in internal states and/or behavior. It combines experience sampling in

which individuals are prompted to report on their momentary states with event-related monitoring in which individuals self-initiate reports contingent on an event such as drinking or smoking. We began to use EMA methods because of our interest in the unfolding processes encompassed in the limit violation effect (LVE), which we will describe later. Given the limitations of traditional methods of self-monitoring (e.g., unknown compliance, poor quality of data), we also were interested in the enhanced validity of the data collected using PDAs. In particular, because the PDA software time stamps each entry and failures to respond to prompts are tagged and stored, it is nearly impossible to fake compliance with self-monitoring (Shiffman, 2000). Reactivity to EMA procedures was similar to that found with other approaches to self-monitoring, even among alcoholics (Litt et al., 1998). Of conceptual interest, the random prompts provided base-rate data as a context for understanding the event-specific assessments of drinking episodes.

Use of PDAs in Alcohol Research. In collaboration with Saul Shiffman, who originated the use of PDAs to collect EMA data, we inaugurated the use of PDAs for collecting data on alcohol use. In more than 11 years of research, we have collected data for 4 different studies in which these electronic diaries (EDs) were used to self-monitor alcohol use and related phenomena (Collins and Muraven, in press). Depending on the study, participants carried their EDs for periods that ranged from 2 to 3 weeks (Muraven et al., 2005a) to 8 weeks (Collins et al., 1998).

Research Participants. We have conducted EMA studies that included persons with a variety of sociodemographic and drinking characteristics. Our populations have included adult men and women who range in age from 18 to 45 years and persons who self-identify as European American, African American, and other minorities. Participants have been college students or persons employed in occupations that ranged from “white collar” professions (e.g., stockbroker) to manual labor (e.g., roofer). Average weekly alcohol intake ranged from 15.8 drinks (Muraven et al., 2005a) to 22 drinks (Collins et al., 1998).

EMA Procedures. Participants received 1 hour of individualized training on how to use the ED. When they were comfortable with the ED, they took it into the field for 2 days and then returned to the training site for a feedback session. While participants were in the field, research staff were available by cellular telephone to help with ED emergencies. Participants returned to the research site for weekly individual 1-hour sessions for downloading data, receiving feedback on their compliance with the protocol, and receiving compensation. These visits also provided opportunities to build rapport, thereby increasing the participants’ engagement with the research.

Interactions with the ED. The nature and number of interactions with the ED were based on the research question(s). In our typical alcohol study, participants interacted with the ED in response to multiple (4–6)

random prompts each day, during which they provided base-rate data that were not linked to drinking. Participants also initiated ED interviews at specific times each day; typically they were asked to turn on the ED at the start of each day (*morning interview*), in the evening (*evening report*), at the start of each drinking episode (*begin drinking*), at the end of each drinking episode (*end drinking*), and just before going to sleep.

Content of the ED Interviews. The content of the ED interviews also varied based on the research question(s) for which EMA data was being collected. For example, our *random-prompt interview* included collection of base-rate information on mood (positive, negative), activities (e.g., eating, driving), social context (e.g., the presence of others), and location (e.g., bar, vehicle). The *end drinking interview* included questions about the number of drinks consumed, intoxication, reasons for drinking, and causal attributions for drinking.

The Nature of the Data Collected Using the ED

Compliance with the ED Protocol. Software for the ED was custom designed for specific research protocols. It included features to prevent missing data and out of range responses. To assess overall compliance with the ED protocol, we focused on responses to the *random prompts*. In our experience, most random prompts were responded to as soon as they were received. For example, in a 2-week study, participants received 2,113 random prompts ($M = 44.82/\text{person}$) and responded to 1,903 of them (80.6%) within 2 minutes of being signaled (Muraven et al., 2005a). Prompts that occurred at inopportune times could either be suspended (i.e., turn off prompting for up to 2 hours) or delayed (i.e., put off responding for up to 20 minutes).

Limit Violation Effect. One major reason for using EMA methods was the potential to test a dynamic theoretical model that was difficult to capture in the laboratory (cf. Collins et al., 1994) or using cross-sectional questionnaire data (cf. Collins and Lapp, 1991). Thus, our principal focus was to test hypotheses derived from our model of the LVE. The LVE describes sequential relationships among situational subjective reactions and drinking behavior, which could move some social drinkers into excessive drinking. It occurs in a cycle that begins with the individual’s concern about drinking and decision to self-impose a limit on alcohol intake. When individuals violate their drinking limits and attribute the violations to themselves, they experience negative affective reactions to the limit violation. Excessive drinking occurs as a way to repair the negative affect/mood. Personal data assistants provided an ecologically valid way to capture the subjective and transient components of the LVE, in real time. Thus, we could examine relationships within and across drinking episodes and within and across persons.

Findings Regarding the LVE. We have used EMA data to examine different aspects of the LVE. For example, we

conducted studies in which we examined: (1) changes in mood as a function of alcohol intake during a drinking episode (Muraven et al., 2005a), and (2) the role of mood and attributions in subsequent drinking episodes (Muraven et al., 2005b). In each case we found support for the LVE. To examine changes in mood at the end of a drinking episode, we focused on the real-time measurement of mood at the start and end of each drinking episode. We found that the more drinking during the episode combined with the more internal attributions (i.e., self-blame) for drinking, the greater the decline in mood by the end of the drinking episode (Muraven et al., 2005a). At the next step in the LVE cycle, we examined mood and subsequent drinking and found that mood after drinking predicted alcohol intake in next drinking episode. Social drinkers returned to drinking sooner, following drinking episodes in which they reported a more negative mood. Self-blame for excessive drinking at the end of 1 episode was related to increased alcohol intake in the next episode (Muraven et al., 2005b).

Advantages and Challenges of Using PDAs. Our experiences have highlighted and confirmed the many advantages of using PDAs to collect data on alcohol use and related phenomena. Personal data assistants provide an ecologically valid way to collect prospective data about ongoing internal and external events, in real time. Situation-specific and event-contingent data can be recorded in the context of the events of interest. There are fewer distortions or biases in participant recall and retrieval as well as tight control over data collection. Compliance cannot be faked.

The benefits of using PDAs outweigh their challenges, but these challenges need to be acknowledged. Clearly, the intensive assessments and multiple interviews can pose a monitoring burden for participants and therefore may limit the generalizability of findings, particularly because good data collection necessitates the recruitment of highly motivated and engaged participants. There can be technical glitches with the PDA's hardware and software. As found with other forms of self-monitoring, high levels of alcohol use can produce drug effects that result in declines in cognitive and motor performance, which can interfere with the mechanics of data collection. Ecological momentary assessment methods rely on self-report.

To address these challenges, we see the following as essential for successful use of PDAs for collecting data. It is important to plan for the inevitable technical glitches and to hire and train highly motivated staff who can establish and maintain rapport and participant engagement with the research. The thorough screening of participants helps to reduce attrition related to the monitoring burden. It is important to compensate participants; it helps to keep them engaged in the intensive monitoring procedures. In our experience, compliance and attrition are better during shorter monitoring periods (2–3 weeks). Researchers also should be aware that EMA procedures produce very large data sets, which can mean that data management is com-

plicated and time-consuming. Analyses of these data require sophisticated statistical techniques such as HLM (Schwartz and Stone, 1998). Personal data assistants represent only 1 of many technical approaches to collecting data in real time. As technologies evolve, researchers will have a variety of options, including IVR and cellular telephone technologies (e.g., Collins et al., 2003), for how best to collect self-monitoring data in real time.

COLLECTING EVENT-LEVEL DATA USING THE WORLD WIDE WEB

Dan J. Neal and Kim Fromme

The Internet, or the World Wide Web, is an innovative method for collecting event-level data. Benefits of Web-based data collection are many. First, because responses are immediately saved to the Web server and all data are date and time stamped, the resulting data set, which can be assessed for compliance to research protocol, is available for analysis in real time. Second, large samples are easily accommodated; there is essentially no limit to the number of participants who can be given access to the research Website. Third, for participants who are computer-savvy, Web-based data collection is simple to use and requires little training. Fourth, Web-based surveys are generally dynamic and flexible, allowing for skip-out questions that may not be relevant to an individual, while also allowing for specific follow-up questions if a particular behavior is reported.

There are, however, also drawbacks to Web-based surveys. First, in terms of accessibility, there is a question of whether all participants have the ability and the knowledge required to log into the website and provide data. Second, the potential for technical problems can occur in terms of server-side errors, such as website unavailability or mistakes in computer code that may prohibit completion. Likewise, participant-side errors such as Web-browser incompatibility or obsolete computer equipment can also interfere with successful submission of data. Third, if participants learn that they can save time by minimizing their reports (and thus take advantage of built-in skip-out routines) bias may be introduced into the data set.

Daily Monitoring and "The UT Experience!" "The UT Experience!" is a 5-year project, funded by the National Institute on Alcohol Abuse and Alcoholism (NIAAA) that is tracking the alcohol use, consequences, and behavioral risks of college students. As a part of this study, participants engage in 30 days of Web-based daily self-monitoring each year. Data collection is ongoing and current preliminary analyses report data provided during participants' first academic year. Of those invited ($n = 1,654$), 81% ($n = 1,337$) provided some data. Current analyses were restricted to participants who provided at least 14 days of data and reported at least 1 occasion of alcohol use; thus, the final sample size was 664. Average age of participants was 18.9 (0.4) years and the sample was 64%

female and 61% Caucasian, 15% Hispanic, 15% Asian, and 3% African American. For each day, participants reported consumption in standard drinks, time spent drinking, and subjective intoxication. Next, participants reported whether they engaged in any of 11 risk behaviors, both when they were drinking and when they were not drinking. Finally, participants reported whether they experienced any consequences as a result of the specific risk behaviors that they engaged in, both when drinking and when not drinking. Participants were asked to log into the website each day, to provide data for the previous day. However, participants had access to the 7 previous days to allow for participants to fill in missed days or make corrections.

Preliminary Results. Participants' compliance to the experimental protocol was adequate. The average number of days reported was 27.3; 47% of participants provided all 30 days of data, and another 43% provided between 21 and 29 days of data. However, only 5 individuals (1%) completed all 30 days consecutively.

Alcohol use was reported on 12.0% of days for females, and 11.4% of days for males; average blood alcohol concentration (BAC) per drinking occasion was 0.09 (0.06) for females, 0.08 (0.05) for males. Average peak BAC for females was 0.16 (0.12) for females and 0.14 (0.11) for males. Risk behaviors assessed included academic delinquency (60%), any sexual behavior (41%), aggressive behavior (29%), illicit drug use (22%), drinking and driving (21%), gambling (17%), unsafe sexual behavior (16%), being a victim of sexual coercion (15%), theft (8%), perpetration of sexual coercion (7%), and vandalism (6%).

Exploratory analyses were then conducted to assess the association between alcohol intoxication and the occurrence of 2 risk behaviors (aggressive behavior and unsafe sex) and, conditional on the occurrence of a risk behavior, the association between alcohol intoxication and the occurrence of a consequence. Generalized estimating equations (GEE; e.g., Hardin and Hilbe, 2003) based on logistic regression (e.g., Hosmer and Lemeshow, 2000) was implemented for each analysis. Blood alcohol concentration and gender were the 2 primary predictor variables. Blood alcohol concentration was person-centered and entered as *average BAC* and *daily BAC*, as well as the *average × daily BAC* interaction. For unsafe sexual behavior, a main effect for *gender* as well as *gender × average BAC* and *gender × daily BAC* interactions were included.

Analysis of risk behaviors indicated that for aggression, heavier drinkers were more likely to engage in aggressive behavior regardless of whether alcohol had been consumed, and the risk of aggressive behavior increased steadily as intoxication increased. For unsafe sex, heavier drinkers were more likely to engage in unsafe sex than lighter drinkers, and unsafe sex became more likely as intoxication increased. Furthermore, the degree to which the risk of unsafe sex increased was moderated by average

intoxication, with heavier drinkers showing less increase in risk than lighter drinkers.

Analysis of consequences indicated that for consequences of aggression, lighter drinkers were more likely to report consequences of aggressive behavior than heavier drinkers; however, the likelihood of consequences did not change as an individual became more intoxicated. Consequences of unsafe sex showed a different pattern. Daily alcohol intoxication was positively associated with increased likelihood of consequences of unsafe sex. Furthermore, significant gender by average consumption interaction emerged. In particular, heavy drinking males were much more likely than light drinking males to report experiencing consequences; conversely, light drinking females were more likely than heavy drinking females to report experiencing consequences.

Thus in general, analyses indicated that aggressive behavior and unsafe sex, as well as their related consequences, had differential associations with alcohol intoxication.

Summary and Future Directions. The majority of participants provided daily event-level data and a wide range of risk- and alcohol-use behaviors were endorsed. A few technical problems occurred, primarily resulting from participants having difficulties accessing the online surveys and minor programming bugs. For those individuals who wish to use Web-based procedures in the future, they should be aware that they will constantly need to monitor their website and provide computer support to participants when needed.

One of the benefits of event-level research is the virtually unlimited number of hypotheses that can be addressed. Although temporal changes in drinking across days, weeks, months, semesters, and years are of interest, the strength of event-level data is that they provide a framework for complex patterns of drinking and risk behaviors to be identified. For example, preliminary data analyses demonstrated the complex nature of the association between average intoxication, daily intoxication, aggressive behavior, and unsafe sex. Data on several other risk behaviors have not yet been analyzed. Just as aggression and unsafe sex had differential patterns of association with alcohol intoxication, the frequency and consequences of other risk behaviors may differ in their relation with alcohol.

Analyses reported here also investigated the role of gender in the context of a single risk behavior. Not only can gender be investigated with all other risk behaviors, a wide variety of other potential moderators of the alcohol use/risk behaviors and alcohol use/consequences association can be examined. Personality factors such as impulsivity and sensation seeking, demographics such as family history and living situation, and situational factors such as weekends, holidays, vacations, and social settings may all influence daily alcohol use and risk behaviors. Reciprocal associations could also be probed, to determine whether severe consequences or very heavy intoxication influence future drinking events. Furthermore, these data provide the opportunity to compare other predictors of

risk behaviors, such as subjective and objective indices of intoxication.

Conclusion. In general, Web-based daily monitoring of "The UT Experience!" has been a tremendous success. Compliance to the protocol was good, a wide variety of risk behaviors were endorsed, and preliminary analyses are yielding interesting and valuable findings with regards to alcohol-related risk taking behaviors.

CONCLUSIONS

The 4 papers presented as a part of this symposium illustrated the use of the TLFB, IVR, PDAs, and the World Wide Web to collect and analyze event-level data. Each presentation demonstrated interesting and innovative uses of daily- and event-level data, both from a theoretical and from an empirical framework; the presentations shed light on the complex nature of alcohol use and its association with a variety of other risk behaviors. Given the richness of the data provided by these assessment approaches, it is not surprising that the data were analyzed using a wide variety of complex statistical techniques, which yielded informative patterns and conclusions.

The reader may be left wondering which of the approaches described is most appropriate for use in ongoing or future studies. As outlined in the preceding sections, all of these methods have unique benefits and costs associated with their use that must be considered when planning for event-level data collection. The TLFB provides a versatile and relatively inexpensive approach to collecting data that can be used to answer research questions at both the aggregate and the event level. For studies that seek a more detailed understanding of alcohol use patterns but for which event-level associations are not of central importance, the TLFB is an excellent option. For example, treatment outcome studies are not solely focused on the interrelations of alcohol and associated behaviors but may nonetheless benefit from data that allow evaluation of more complex patterns of alcohol use across time. With respect to limitations, the primary concern with the TLFB is the potential for difficulty with recall over long periods of time. Despite this concern, the TLFB represents a significant advance relative to aggregate quantity frequency indices that have been widely used in the past.

Although IVR and Web-based methods utilize distinctly different technologies, the 2 studies presented yielded remarkably similar data. Both approaches were used to capture data regarding drinking and drinking-related variables for each day over a prespecified period of time. In many respects these methods are quite similar to the TLFB, with the exception that the data are collected prospectively or within a shorter retrospective recall period. It is the prospective nature of the data collected via these methods that make them such powerful approaches. Although the concern for the quality of retrospective recall is largely eliminated using these methods, in many respects it is replaced

with concerns regarding compliance, which become paramount with these methods of data collection. Fortunately, as demonstrated in the 2 previous presentations, steps can be taken by researchers to ensure acceptable compliance and minimize the intrusiveness of the approach to research participants. In summary IVR and Web-based methods are recommended when the research questions are primarily focused on the relations among alcohol and other behaviors that may change dynamically over time. The confidence one can place in prospective data is crucial for drawing accurate inferences in this type of research.

Perhaps the most complex of the event-level methods presented was the PDA data collection method. As illustrated, the primary benefit of using PDAs is that they allow researchers to move beyond capturing data regarding the day (as done with the TLFB, IVR, and Web-based methods) by capturing data in the moment. Personal data assistants can be used at multiple times within a single day and participants can initiate a report as they engage in a behavior such as drinking or in response to random prompts throughout the day. The increased level of specificity relative to other event-based methods enhances the ability of researchers to engage in hypothesis testing and theory development that requires insight into thoughts, emotions, and behaviors occurring from moment to moment. Personal data assistants allow researchers to capture these moments and model them in subsequent analyses.

In summary, decisions regarding which method to use should be based both on a consideration of the benefits and costs of each method and on the types of questions the researcher hopes to answer. The TLFB is time and cost-effective and provides reliable and valid data for capturing alcohol use patterns that would not be identified using quantity-frequency indices. Because of concerns about recall bias using retrospective methods like the TLFB, they may not be ideal for research studies designed specifically to probe relations between alcohol and other risk behaviors that unfold dynamically across time. With respect to IVR and Web-based methods, the population under study is an important consideration. The greater access to touch-tone phones may make this approach better suited for use with certain populations whereas the ease of use of Web-based approaches may make this approach more useful with others. For large-scale studies of diverse populations, some combination of IVR and Web-based methods may be optimal. Personal data assistants are very time and effort intensive for participants and may therefore be unsuitable for very large-scale studies. At the same time, PDAs provide the opportunity for the richest form of event-level data and may therefore be ideal for addressing certain types of very specific and theoretically important questions.

REFERENCES

- Anton RF, Randall CL (2005) Measurement and choice of drinking outcome variables in the COMBINE Study. *J Stud Alcohol* 15:104-109.

- Bardone AM, Krahn DD, Goodman BM, Searles JS (2000) Using interactive voice response technology and timeline follow-back methodology in studying binge eating and drinking behavior: different answers to the same question? *Addict Behav* 25:1–11.
- Collins RL, Kashdan TB, Gollnisch G (2003) The feasibility of using cellular phones to collect ecological momentary assessment data: application to alcohol consumption. *Exp Clin Psychopharm* 11:73–78.
- Collins RL, Lapp WM (1991) Restraint and attributions: evidence of the abstinence violation effect in alcohol consumption. *Cog Ther Res* 15:69–84.
- Collins RL, Lapp WM, Izzo CV (1994) Affective and behavioral reactions to the violation of limits on alcohol consumption. *J Stud Alcohol* 55:475–486.
- Collins RL, Morsheimer ET, Shiffman S, Paty JA, Gnys M, Papandonatos GD (1998) Ecological momentary assessment in a behavioral drinking moderation training program. *Exp Clin Psychopharmacol* 6:1–10.
- Collins RL, Muraven M (in press) Ecological momentary assessment for alcohol consumption. in *The Science of Real-Time Data Capture: Self Report in Health Research* (Stone AA, Shiffman S, Atienza AA, Nebeling L eds.) Oxford University Press, Oxford, UK.
- Collins RL, Parks GA, Marlatt GA (1985) Social determinants of alcohol consumption: the effects of social interaction and model status on the self-administration of alcohol. *J Consult Clin Psychol* 53:189–200.
- Del Boca FK, Darkes J (2003) The validity of self-reports of alcohol consumption: state of the science and challenges for research. *Addiction* 98S:1–12.
- Del Boca FK, Darkes J, Greenbaum PE, Goldman MS (2004) Up close and personal: temporal variability in drinking during the first year of college. *J Consult Clin Psychol* 72:155–164.
- Del Boca FK, Noll JA (2000) Truth or consequences: the validity of self-report data in health services research on addictions. *Addiction* 95S:347–360.
- Greenbaum PE, Del Boca FK, Darkes J, Wang C-P, Goldman MS (2005) Variability among freshman college students' drinking trajectories. *J Consult Clin Psychol* 73:229–238.
- Hardin JW, Hilbe JM (2003) *Generalized Estimating Equations*. Chapman & Hall/CRC, Boca Raton, FL.
- Helzer JE, Badger GJ, Rose GL, Mongeon JA, Searles JS (2002) Decline in alcohol consumption during two years of daily reporting. *J Stud Alcohol* 63:551–558.
- Hosmer DW, Lemeshow S (2000) *Applied Logistic Regression*. John Wiley & Sons Inc., New York.
- Leigh BC (2000) Using daily reports to measure drinking and drinking patterns. *J Substance Abuse* 12:51–65.
- Litt MD, Cooney NL, Morse P (1998) Ecological momentary assessment (EMA) with treated alcoholics: methodological problems and potential solutions. *Health Psychol* 17:48–52.
- Miller WR (1996) *Manual for Form 90: A Structured Assessment Interview for Drinking and Related Behaviors*. Project MATCH Monograph Series. Vol. 5. NIAAA, Rockville, MD.
- Mundt JC, Kelleher PH, Perrine MW, Searles JS (1997) Psychological performance assessment via interactive voice response systems. *Behav Res Methods Instrum Comput* 29:506–518.
- Mundt JC, Perrine MW, Searles JS, Walter D (1995) An application of interactive voice response (IVR) technology to longitudinal studies of daily behavior. *Behav Res Methods Instrum Comput* 27:351–357.
- Muraven M, Collins RL, Morsheimer ET, Shiffman S, Paty JA (2005a) One too many: a test of the LVE in predicting future alcohol consumption. *Exp Clin Psychopharm* 13:127–136.
- Muraven M, Collins RL, Morsheimer ET, Shiffman S, Paty JA (2005b) The morning after: limit violations and the self-regulation of alcohol consumption. *Psych Addict Beh* 19:253–262.
- Perrine MW, Mundt JC, Searles JS, Lester LS (1995) Validation of daily self-reported alcohol consumption using interactive voice response (IVR) technology. *J Stud Alcohol* 56:487–490.
- Schwartz JE, Stone AA (1998) Strategies for analyzing Ecological Momentary Assessment data. *Health Psychol* 17:6–16.
- Searles JS, Helzer JE, Rose GL, Gadger GJ (2002) 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.
- Searles JS, Perrine MW, Mundt JC, Helzer JE (1995) Self-report of drinking using touch-tone telephone: extending the limits of reliable daily contact. *J Stud Alcohol* 56:375–382.
- Searles JS, Walter DE, Helzer JE (2000) Comparison of drinking patterns measured by daily reports and timeline follow back. *Psychol Addict Behav* 14:277–286.
- Shiffman S (2000) Real time self-report of momentary states in the natural environment: computerized ecological momentary assessment. in *The Science of Self-Report: Implications for Research and Practice* (Stone AA, Turkkan J, Bachrach C, Jobe J, Kurtzman H, Cain V eds.), pp 49–61. Lawrence Erlbaum Associates, Mahwah, NJ.
- Sobell LC, Sobell MB (2000) *Alcohol Timeline Followback (TLFB)*. *Handbook of Psychiatric Measures*. pp 477–479. American Psychiatric Association, Washington, DC.
- Sobell LC, Sobell MB, Connors GJ, Agrawal S (2003) Assessing drinking outcomes in alcohol treatment efficacy studies: a yardstick of success. *Alcohol Clin Exp Res* 27:661–1666.
- Sobell MB, Sobell LC, Klajner F, Pavan D, Basian E (1986) The reliability of a timeline method for assessing normal drinker college students' recent drinking history: utility for alcohol research. *Addict Behav* 11:148–162.
- Stout RL (2003) Methodological and statistical considerations in measuring alcohol treatment effects. *Alcohol Clin Exp Res* 27:1686–1691.
- Tonigan JS, Miller WR, Brown JM (1997) The reliability of Form 90: an instrument for assessing alcohol treatment outcomes. *J Stud Alcohol* 58:358–364.