Reactivity and ecological momentary assessment using a sample of undergraduate problem drinkers

Alan L. Shields

The University of Montana

Let us know how access to this document benefits you.
Follow this and additional works at: https://scholarworks.umt.edu/etd

Recommended Citation
Graduate Student Theses, Dissertations, & Professional Papers. 8737.
https://scholarworks.umt.edu/etd/8737

This Thesis is brought to you for free and open access by the Graduate School at ScholarWorks at University of Montana. It has been accepted for inclusion in Graduate Student Theses, Dissertations, & Professional Papers by an authorized administrator of ScholarWorks at University of Montana. For more information, please contact scholarworks@mso.umt.edu.
Permission is granted by the author to reproduce this material in its entirety, provided that this material is used for scholarly purposes and is properly cited in published works and reports.

** Please check "Yes" or "No" and provide signature **

- Yes, I grant permission  
- No, I do not grant permission

Author's Signature

Date 5/4/00

Any copying for commercial purposes or financial gain may be undertaken only with the author's explicit consent.
REACTIVITY AND ECOLOGICAL MOMENTARY ASSESSMENT USING A SAMPLE OF UNDERGRADUATE PROBLEM DRINKERS

By:

Alan L. Shields

A Thesis Presented at the University of Montana

Department of Psychology

In partial fulfillment of the requirements for the degree of

Master of Arts

Summer 2000

Approved by:

Chairperson

Dean, Graduate School

5-4-2000
Stone and Shiffman (1994) define ecological momentary assessment (EMA) as monitoring or sampling strategies that assess phenomena at the moment they occur in natural settings, thus maximizing ecological validity while avoiding retrospective recall. EMA has been used to better understand addictive behaviors among smokers (Shiffman, Hufford, et al., 1997) and problem drinkers (Collins et al., 1998; Litt et al., 1998). Research has only begun to examine the extent to which real-time monitoring impacts the behaviors and cognitions under observation. This study examined behavioral, motivational, self-efficacy, alcohol urge, and self-reported reactivity to EMA among a sample of 33 undergraduate problem drinkers. Participants completed a 2-week real-time monitoring protocol using palmtop computers that assessed mood, activities, urge to drink, and drinking behavior. Findings suggest that overall magnitude of reactivity to EMA is small. An analysis of participant characteristics that may predict reactivity revealed several interesting findings. Suggestions for future research are presented.
Reactively and Ecological Momentary Assessment Using a Sample of Undergraduate Problem Drinkers

Recently, researchers have been using new methodologies to overcome the weaknesses inherent in retrospective self-reports. One such method, Ecological Momentary Assessment (EMA), stresses the continuous investigation of individual differences in behavior, cognition, and affect in the natural environment (Shiffman & Stone, 1998). Stone & Shiffman (1994) outline four qualities that define EMA. First, EMA assesses phenomena at the moment they occur in order to reduce recall biases. Second, EMA is dependent upon careful timing of assessments in order to ensure that an unbiased sample of the target behavior is assessed. Third, EMA methods typically obtain a significant number of repeated observations in order to produce more reliable samples and allow researchers to delineate covarying relationships over time. Finally, EMA measurements are made in the environment that participants typically inhabit. These characteristics of EMA are essential for researchers interested in understanding dynamic processes as they evolve over time (e.g., Marlatt, 1985; Prochaska, DiClemente, & Norcross, 1992; Vallacher & Nowak, 1994). Currently, EMA is being used to understand smoking (e.g., Shiffman, Paty, Gnys, Kassel, & Hickcox, 1996; Shiffman, Hufford et al., 1997), drinking (e.g., Collins et al., 1998; Hufford et al., under review), the relationship between coping and chronic pain (e.g., Affleck et al., 1998), stress, coping, and mood (Stone et al., 1998), and stress and cardiovascular reactivity (e.g., Kamarck et al., 1998).

EMA has utilized electronic technology such as beepers (Johnson & Larson, 1982) and programmed wristwatches (Litt, Coony, & Morse, 1998) to prompt participants to complete self-report questionnaires. Ecological validity has been further...
improved through the use of palm-top computers (Shiffman et al., 1994; Shiffman, Hufford et al., 1997) which collect data in near real-time. In this way, all responses from participants are date and time stamped, and participants can be monitored for compliance thereby increasing the validity of their self-reports (Stone & Shiffman, 1994).

Work with EMA has demonstrated that traditional retrospective self-report measures are vulnerable to inaccuracies and recall bias. Shiffman, Hufford et al. (1997) showed that never-smokers asked about a “typical smokers” relapse produced extremely similar relapse profiles as compared with the retrospective report of smokers. This finding suggests that schemas about a target behavior can significantly influence retrospective recall. Both the never-smokers’ schemas and smokers’ retrospective recall failed to correlate to real-time reports of actual smoking lapses. This implies that, no matter how reasonable and believable a behavioral account may sound, it can still be inaccurate and systematically biased (Shiffman, Hufford et al., 1997). This is a significant finding given that a large number of addictive behavior researchers rely on retrospective methods to gather data. Retrospective reports may be able to address the static variables associated with addiction, but EMA appears to be the superior methodology for the accurate understanding of the day-to-day and moment-to-moment phenomena theorized to be crucial in the study of addictive behaviors (Brownell, Marlatt, Lichtenstein, & Wilson, 1986).

McKay (1999) called EMA the most important contribution to contemporary addictive behavior research. He concluded his critical review of addictive behavior research methodologies by encouraging the use of a multitrait-multimethod design for the understanding of addictive behaviors. In particular, he called for the use of new
technologies (e.g., palm-top computers) in addictive behavior assessment as methods that reduce memory inaccuracies and biases. If EMA represents the emerging methodology in addictive behavior research, then researchers must begin to understand some of its potential weaknesses and limitations.

EMA: Methodological Limitations and Reactivity

Despite many advantages, EMA may have unique limitations. One potential source of inaccuracy and bias is reactivity. A reactive effect describes the degree to which the intensity, frequency, and/or quality of a target variable will change when being observed, monitored, or assessed (Nelson, 1977). EMA researchers frequently cite reactivity as a potential confound in their studies (e.g., Shiffman, Hufford et al., 1997; Collins et al., 1998; Shiffman et al., 1998; Affleck et al., 1998; Kamarck et al., 1998; Litt et al., 1998); however, few studies (Cruise Broderick, Porter, Kaell, & Stone, 1996) exist that systematically explore the reactivity phenomenon in EMA specifically.

Ecological momentary assessment may be particularly prone to reactivity as it often involves a creative combination of two existing research orientations: (1) self-monitoring (Nelson, 1977) and (2) the experience sampling method (Csikszentmihalyi & Larson, 1987).

Self-monitoring. Self-monitoring is theorized to involve two-stages (Nelson, 1977). First, the participant must be able to discriminate aspects of his or her behavior and effectively determine if the target behavior has indeed occurred. Second, the participant must make a self-recording response. In other words, participants must then use whatever method is being employed to record the occurrence of the target event. Recent definitions of SM have been expanded to assess contextual variables associated
with the target behavior. Thus, observations regarding both one’s internal and external environment are incorporated into the SM procedure (Korotitsch & Nelson-Gray, 1999).

Researchers have demonstrated that a wide variety of behaviors (e.g., academic performance, cigarette smoking, eye-blink response, food consumption) can be reactive to self-monitoring (e.g., Lam, Cole, Shapiro, & Bambara, 1994; Abrams & Wilson, 1979; Sieck & McFall, 1976; Romanczyk, Tracy, Wilson, & Thorpe; 1973). The literature regarding self-monitoring of drinking behavior has yielded mixed results. Some studies have found no evidence of reactivity (e.g., Camey, Tennen, Affleck, Del Boca, & Kranzler, 1998; Samo, Tucker, & Vuchinich, 1989, Sobell, Bogardis, Schuller, Leo, & Sobell, 1989), while others suggest the presence of behavioral reactivity as a function of self-monitoring (Uchalik, 1979).

**Experience sampling.** In contrast to self-monitoring, the experience sampling method relies on an electronic instrument to signal a participant to report his or her immediate activities, thoughts, and moods (Csikszentmihalyi & Larson, 1987). It is a *random sampling* technique that has been successful in gathering base rate data of behaviors and their antecedents (e.g. Delespaul & DeVries, 1987). This assessment strategy may produce reactivity in some circumstances (e.g., Larson, 1989; Magneberg, 1998). For example, a study of heavy drinkers found that monitoring activities and mood resulted in decreased drinking over a three-week monitoring protocol (Magneberg, 1998).

**Summary.** Understanding the reactive effects of EMA is important for several reasons. If EMA produces reactivity, then a loss of methodological validity must be acknowledged. Also, this information can assist both the EMA researcher in minimizing reactivity and serve as a guide to the applied clinician who may want to maximize
therapeutic reactivity. Moreover, understanding reactive effects of EMA can lead to more interpretable research conclusions. For example, if EMA is introduced along with a therapeutic intervention (e.g., Collins et al., 1998), changes in behavior due to reactivity may make it hard to evaluate the effects of the intervention. In other words, reactivity may be a source of unwanted variance in the dependent variables of interest. In addition, if EMA is used to establish the frequency “baseline” of behaviors, researchers may be misled as the EMA may alter the frequency of the target behavior and reduce the representativeness of the target behavior (e.g., Paty et al., 1992). Finally, understanding reactivity to EMA is crucial because it raises one of the most important and fundamental questions concerning self-report – How does monitoring behavior, affect, and cognitions affect these phenomena?

Present Study

Understanding reactivity to EMA is an important methodological issue for the behavioral sciences (Affleck, Zautra, Tennen, & Armeli, 1999; Korotitsch & Nelson-Gray, 1999). This study sought to explore reactivity to EMA by engaging hazardous drinkers in an EMA study of their drinking behavior. Five possible domains of reactivity were assessed: drinking behavior, levels of motivation to change alcohol use patterns, self-efficacy reports, drinking urges, and self-reported reactivity.

Behavioral reactivity. This study examined the magnitude of change in self-reported drinking behavior between pre- to post-EMA monitoring. The historic literature suggests that monitoring an addictive behavior will decrease the frequency of its occurrence (e.g., Abrams & Wilson, 1979; Fremouw & Brown, 1980). Collins et al. (1998) found support for this hypothesis in a study of heavy drinkers recruited from the
community. Study participants tended to report slight decreases in their weekly consumption of alcohol (Collins et al., 1998). Based on this evidence, it was hypothesized that EMA participants in the present study would report decreased quantity and frequency of alcohol consumption during monitoring compared to their pre-monitoring, baseline levels of consumption.

Motivational reactivity. This study also examined changes in motivation to change. Motivation to change has been found to correlate with self-awareness (Prochaska, Velicer, DiClemente, & Fava, 1988). The Transtheoretical Model of Change (Prochaska et al., 1992) suggests that increasing information about self and one’s problem and assessing how one feels and thinks about oneself with respect to a problem (e.g., self-monitoring) can be an important predictor of change. Based on this premise, it was hypothesized that participants in an EMA protocol would demonstrate forward movement on pre- to post-reports of their readiness to change. Specifically, it was hypothesized that EMA participants would increase their Contemplation scores, and decrease their Precontemplation scores, as a function of real-time monitoring. Furthermore, it was predicted that there would be a significant increase in mean Readiness to Change (RtC) scores, a composite index of motivation for change, from pre- to post-monitoring.

Self-Efficacy Reactivity. This study also explored change in alcohol self-efficacy expectancies; in other words, one’s confidence in one’s ability to resist heavy drinking. Based on the empirical literature relating self-monitoring to increases in problem awareness (Prochaska et al., 1992; Litt et al., 1998), it was hypothesized that participants in an EMA protocol would report decreased self-efficacy expectancies after completing
the real-time monitoring protocol. That is, once participants monitor their drinking patterns in near real-time, they will be more likely to report decreased confidence in their ability to resist heavy drinking.

**Urge reactivity.** The magnitude and direction of change in alcohol urge as a function of completing the real-time assessments is also of interest. Current theories of urges suggest that they are supported by effortful, controlled cognitive processing (Tiffany, 1990). Researchers have confirmed that urges require attentional resources (e.g., Sayette & Hufford, 1994; Sayette & Parrott, 1999). Furthermore, Litt et al. (1998) reported that their alcoholic participants reported decreased urges to drink while self-monitoring. By completing a real-time interview it was hypothesized that EMA participants would sustain a reallocation of their attentional resources from the urge itself to completing the real-time reports. That is, participants were signaled to complete an urge report at the beginning and end of five daily assessments. It was hypothesized that urge magnitude would decrease between the 1st and 2nd urge report within each assessment.

**Self-reported reactivity.** This study explored participants' accounts of the impact monitoring had on their drinking behavior. Litt et al. (1998) explored self-reported reactivity to EMA among drinkers. In post-monitoring interviews they found that alcoholic participants reported that the monitoring protocol made them more aware of their drinking and consequently made them drink less. However, when compared with a no-monitoring control group, reported drinking behavior did not significantly change. Based on their data, it was hypothesized that after completing the EMA protocol
Participants would self-report decreased drinking quantity and frequency relative to their pre-monitoring baseline alcohol consumption.

Methods

Participants

A total of 389 (39.6% female and 20.7 [SD = 4.4] years of age) Psychology 100 students were screened for this study. Participants scoring 10 or greater (N = 148, 38%) on the Alcohol Use Disorders Inventory Test were eligible for participation (AUDIT; Allen, Litten, Fertig, & Babor, 1997). AUDIT scores can range from 0 to 40. Participants meeting inclusionary criteria were contacted by phone to inquire about their participation in the real-time monitoring study. During this initial telephone contact participants were given information about the study, including the use of the electronic diaries (ED’s) for assessment. Thirty-eight participants were invited to attend the initial interview and EMA training session. Thirty-three participants completed this study. Five participants withdrew or were removed from the study as a result of not showing up for the training session (n = 2), an inability to integrate ED into their lives (n = 2), or losing the Palm Pilot (n = 1). All participants received full experimental credit for their Psychology 100 class requirement.

Design and Procedures

Participants at the psychology 100 screening who met inclusionary criteria for the present study were asked to the Addictive Behaviors Research Laboratory (ABRL) for an interview and EMA training session. Prior to completing any questionnaires, participants signed an informed consent form. At that time, they completed a series of questionnaires (see Table 1) and participated in an interview session with either a graduate student or
Ph.D. in Clinical Psychology. Two interviews, the Addiction Severity Index (ASI; McLellan et al., 1992) and the mood disorders and alcohol use disorders modules of the Structured Clinical Interview for DSM-IV Disorders-I (SCID-I; First, Spitzer, Gibbon, & Williams, 1997) were administered. Participants received 1 to 1.5 hours of group EMA training. The training included instruction on the use of the electronic diary (ED), including how to record drinks and temptations to drink, how to respond to signal-contingent assessments, and how to use the sleep/wake, nap, and other functions. Participants were given the opportunity to practice the electronic interviews and to ask questions regarding the ED and monitoring protocol.

Once mastery of the ED had been demonstrated by the participant’s successfully navigating the electronic questionnaires, each participant was given his or her own ED and instructed to return to the ABRL after 3 days of self-monitoring. This session was scheduled to assess initial compliance with monitoring and answer questions that arose as a function of monitoring in the field. Two subsequent meetings were scheduled with each participant. These were conducted to continue monitoring compliance and upload data from each ED to a main computer. Sessions were scheduled for one and two weeks after intake into the study. During the final upload, participants completed post-monitoring questionnaires (see Table 1). Participants were then debriefed and given appropriate referrals if they were interested in talking with someone about their drinking. Finally, participants were given their experimental credits.

Timeline. This study ran participants in four waves of monitoring. Each participant’s timeline approximated the following (see Table 1):
Table 1. Timeline and Self-Report Measures.

<table>
<thead>
<tr>
<th>DAY</th>
<th>STUDY EVENT</th>
<th>DESCRIPTION OF EVENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>---</td>
<td>Psychology 100 Screening</td>
<td>1) Psychology 100 students were group screened with the Alcohol Use Disorders Inventory Test (AUDIT).</td>
</tr>
</tbody>
</table>
| --- | Phone confirmation with prospective study participants | 1) Review demographics and data from screening interview.  
2) Describe study including EMA protocol.  
3) Inquire about their continued participation in the study.  
4) Schedule for intake interview and EMA training. |
| 1   | Intake Interview and EMA Training | 1) Participants arrive at the Addictive Behaviors Research Laboratory (ABRL).  
2) Participants sign informed consent form and complete Timeline Follow-Back (TLFB), Situational Confidence Questionnaire (SCQ), Alcohol Expectancy Questionnaire III (AEQ), the University of Rhode Island Change Assessment Scale (URICA), and the Social Desirability Scale.  
3) Participants interviewed with the Structured Clinical Interview for DSM-IV Axis I Disorders (mood and substance use modules; SCID-I) and the Alcohol Severity Index (ASI).  
4) Participants receive training and given ED. |
| 4   | Shakedown | 1) Participants return to ABRL and meet with research assistant to monitor compliance with ED.  
2) Participants sent back into the field and instructed to continue self-monitoring. |
| 8   | 1st Data Upload | 1) Participants return to ABRL and meet with research assistant to upload data from ED to main computer.  
2) Participants return to the field for final week of monitoring. |
Measures

The study employed standard measures of substance use, psychiatric symptoms, and trait characteristics important in the study of self-report. These data allowed us to examine individual characteristics that may influence reactivity.

Alcohol Use Disorders Inventory Test. The Alcohol Use Disorders Inventory Test-Core Instrument (AUDIT; Connigrave, Saunders, & Reznik, 1995) is a widely used self-report measure of alcohol consumption. The measure requires 5 minutes to administer and was used to screen participants for hazardous alcohol use (Allen et al., 1997). This study used a cutoff score of 10 (range = 0 to 40), which has been found to be a valid and reliable predictor of harmful and hazardous alcohol use. Saunders, Ashland, Babor, de la Fuente, & Grant (1993) demonstrated that when a cutoff score of 10 was used, the overall sensitivity of the AUDIT for hazardous drinking was 80% and the overall specificity was 98%. Other researchers have replicated these findings (Fleming, Barry, & MacDonald, 1991; Claussen & Aasland, 1993). The measure contains 10 items that ask, for example, “How often do you have a drink containing alcohol?” and “Has a relative or friend or doctor or other health worker been concerned about your drinking or suggested you cut down?”

Demographic Data Sheet. This form gathered basic demographic data including age, sex, ethnicity, and year in school.

Addiction Severity Index. The Addiction Severity Index (ASI; McLellan, et al., 1992) is a widely used measure in addictive behavior research (Ball & Corty, 1988;
Gawin et al., 1989; Kadden, Cooney, Getter, & Litt, 1990). Average overall reliability of trained technicians administering the ASI in person is .89 and .83 or higher for each of the seven scales (McLellan et al., 1985). The ASI uses 57 items to assess an individuals’ recent status (the 30 days just prior to the interview) in 7 categories relating to addiction severity. These categories include medical (e.g., “How many days have you experienced medical problems in the past 30 days?”), employment (e.g., “How many days were you paid for work in the past 30 days?”), alcohol problems (e.g., “Alcohol-any use in the past 30 days?”), drug use (e.g., “Heroin, Cocaine, etc., any use at all in the past 30 days?”), legal issues (e.g., “Are you currently awaiting charges; trial or sentence?”), family problems (e.g., “How many days in the past 30 have you had serious conflicts with your family?”), and psychiatric status (e.g., “In the past 30 days, have you had a significant period [that was not a direct result of your drug/alcohol use] in which you have experienced serious depression?”).

Timeline Follow-Back. The Timeline Follow-Back (TLFB; Sobell & Sobell, 1992) is a sensitive retrospective measure of drinking behavior. It has well established psychometric qualities as a method for evaluating drinking quantity and frequency. For example, Sobell, Sobell, Klajner, Pavan, & Basian (1986) demonstrated a test-retest reliability of .92 to .96 among a college age population. Similar findings have been demonstrated across a variety of populations including inpatient and outpatient alcoholics (Maisto, Sobell, Connors, & Sobell, 1979). O’Farrell, Cutter, Bayog, Dentch, & Fortgang (1984) compared collateral reports of a subjects drinking and found very high correlations for total drinking days ($r = .88$). Finally, Sobell et al. (1986) demonstrated strong concurrent validity of the TLFB by comparing it with established measures of
alcohol related disabilities (Alcohol Dependence Scale, ADS; Short Michigan Alcohol Screening Test, SMAST). As predicted, higher scores on the ADS and SMAST were positively and significantly correlated with TLFB alcohol quantity and frequency variables.

Using the TLFB method, participants are asked to provide estimates of their drinking quantity and frequency over a designated length of time (the 30 days prior to the interview in this study). The measure uses a calendar and a standard drink card to allow the researchers to equate different alcoholic beverages with ethanol content. In this way, researchers are able to standardize the amount of alcohol consumed across participants.

**Situational Confidence Questionnaire.** The Situational Confidence Questionnaire (SCQ; Annis & Davis, 1988) is a 39-item measure of how confident an individual feels regarding their ability to successfully resist the urge to drink alcohol in response to a variety of stimuli (e.g., stress at work, fight with spouse, etc.). Annis & Graham (1988) provide support for the construct validity of the SCQ. These authors correlated SCQ subscale scores (situation-specific confidence levels; e.g., dealing with unpleasant emotions/frustrations) with measures of alcohol consumption. They concluded that the more drinking reported by individuals, the less likely they will be to endorse high confidence levels in their ability to resist the urge to drink across all categories of drinking situations. Furthermore, Annis & Graham (1988) provide evidence for the reliability of the SCQ and conclude that individuals tend to respond consistently to the items within each subscale. SCQ global scores have been found to be strong predictors of alcohol treatment outcome (Greenfield, Hufford et al., 2000). Examples of SCQ items include, “I would be able to resist the urge to drink heavily if other people didn’t seem to
like me” and “I would be able to resist the urge to drink heavily if I felt that I had let myself down.” Participants respond using a 6-point scale that ranges from 0% (not at all confident) to 100% (very confident).

**Alcohol Expectancy Questionnaire-III.** Alcohol expectancies have been found to be powerful predictors of alcohol use and abuse (Marlatt & Rohsenow, 1980). The Alcohol Expectancy Questionnaire (AEQ-III; George et al., 1995) is a 39-item measurement of alcohol related outcome expectancies. Although the AEQ can be broken down into 8 expectancy subscales (global positive, social and physical pleasure, social expressiveness, sexual enhancement, power and aggression, tension reduction and relaxation, cognitive and physical impairment, and careless unconcern), this study focused on the two general categories (positive and negative expectancies) as predictor variables. These two categories provide a more conceptually distinct representation of beliefs about alcohol and avoid analytic and psychometric difficulties (e.g., multicollinearity) that research with the eight subscales generally encounters (George et al., 1995). This measure has participants respond on a 6-point scale ranging from “Agree Strongly” to “Disagree Strongly.” For example, participants are asked to respond to items such as, “Drinking makes the future seem brighter to me” or “Alcohol makes me careless about my actions.”

**University of Rhode Island Change Assessment Scale.** The University of Rhode Island Change Assessment Scale (URICA; McConnaughy, Prochaska, & Velicer, 1983) is a 32-item measure of motivation to change a problem behavior. A follow-up study to this original research (McConnaughy, DiClemente, Prochaska, & Velicer, 1989) confirmed that the URICA reliably measures 4 distinct stages. These stages include
Precontemplation, Contemplation, Action, and Maintenance. Participants are asked questions like, “As far as I’m concerned, I don’t have any alcohol problems that need changing” and “It might be worthwhile to work on my alcohol problem.” Participants respond on a 5-point scale (1 = Strongly Disagree to 5 = Strongly Agree). Factor scores from the URICA place individuals into one of these stages of change which are thought to improve prediction of the probability that one is likely to change the identified problem behavior (e.g., alcohol use; Prochaska et al., 1992). Recently, researchers have been using a readiness to change (RtC) index score by adding scores from the Contemplation, Action, and Maintenance subscales of the URICA and subtracting Precontemplation factor scores. The resulting RtC index provides a single continuous measure of readiness to change (DiClemente, et al., 1999).

Structured Clinical Interview for DSM-IV Disorders-I. The Structured Clinical Interview for DSM-IV Disorders-I (SCID-I; First, Spitzer, Gibbon, & Williams, 1997) is a semi-structured clinical interview that allows for assessment of DSM-IV Axis I clinical disorders. Test-retest reliability studies have indicated that the SCID-I is at least as good as other diagnostic instruments in obtaining stable results (Williams et al. 1992). Results on the validity of the SCID-I remain elusive because of the difficulty in the procedures for comparing diagnoses across measures and standards. However, Kranzler, Kadden, Babor, Tennen, & Rounsaville (1995) were able to demonstrate excellent validity in a study of SCID-I diagnosed substance abusers compared with a standard clinical intake interview. Participants in the present study were administered the alcohol use and mood disorders modules of the SCID-I.
Social Desirability Scale. The Social Desirability scale (SD; Edwards, 1970) is a 39-item measure thought to discriminate individuals who are, more or less, attempting to represent themselves in socially desirable ways. Items are endorsed as either true or false and include such statements as, “I find it hard to keep my mind on a task or job” and “I am happy most of the time.”

Electronic Diary Debriefing Sheet. The Electronic Diary (ED) debriefing sheet was utilized post-EMA monitoring in order to obtain information on how participants perceived the EMA monitoring protocol and its effect on their behavior.

Real-time monitoring

Participants were trained to use the ED, a Palm Pilot Professional Edition (by 3COM) palmtop computer, for data collection. The ED is small (11.8cm X 7.7cm X 1.5cm), lightweight (234 grams including batteries and case), and includes an illuminated liquid crystal display (6cm X 6cm), speaker, real-time clock, and calendar. It runs on 2 AAA batteries and uploads its data to a notebook computer via a Palm Pilot Cradle attached to a serial port. All data receives a date and time stamp. In addition, the software flags non-responses to prompts, allowing for checks regarding participants’ compliance with monitoring. The software has proven itself to be stable in past studies (Shiffman, Enberg, et al., 1997), and the hardware itself is durable (e.g., one participant accidentally dropped her Palm Pilot into a cooler of water; after drying out the data was captured and the unit was sent back out into the field!).

The use of EMA requires careful attention to the type of sampling used (Stone & Shiffman, 1994). Therefore, participants were instructed on how to respond to signal-
and time-contingent sampling schedules as well as on how to initiate event-contingent records.

Signal-contingent sampling included computer-generated random prompts that signal the participant via a "beep" to complete a report an average of 5 times per day. The signal-contingent interviews included items assessing activities (e.g., work, leisure, interacting with others), setting (e.g., home, work-place), mood (10 positive and negative affect mood adjectives drawn from the Positive and Negative Affect Schedule; Watson, Clark, & Tellegen, 1988), and alcohol urges. The urge assessment at the beginning of the signal-contingent assessment asked participants "Urge to drink?" The end assessment of alcohol urge was drawn from the Alcohol Urge Questionnaire, Bohn, Krahn, & Staehler, 1995, and asked, "All I want to do now is have a drink?" This item had the highest factor loading on the AUQ. Participants could drag a visual marker up and down a 1-11 scale to record their responses to both urge questions.

Time-contingent sampling included regular morning and evening reports. Morning reports included items assessing sleep quality and current mood. Evening reports included items summarizing their mood for the day and the occurrence of any positive or negative life events. An event-contingent schedule was also implemented to encourage participants to report any subjective change in their temptation to drink alcohol or if any alcohol is consumed (including the quantity consumed, triggers for drinking, and mood pre- and post-drinking). Participants were instructed to complete the temptation or drink reports after these episodes were over. These near real-time assessments included activity, setting, mood, and coping questions, as well as questions
Regarding triggers for the temptation or drinking episode. The types of sampling contingencies and domains assessed are presented in Table 2.

Table 2: Assessment Contingencies and Domains.

<table>
<thead>
<tr>
<th>TYPE OF CONTINGENCY</th>
<th>MOOD</th>
<th>ACTIVITIES</th>
<th>TRIGGERS</th>
<th>URGE</th>
<th>RCT</th>
<th>COPING</th>
<th>CONSUME</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signal</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Event: Temptation</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Event: Drinking</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Time (morning &amp; evening report)</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1 Readiness to change items

As part of the EMA monitoring, participants were expected to return to the laboratory on three separate occasions to upload the real-time data onto a main computer and receive feedback from research assistants regarding their compliance. Following the completion of the EMA protocol, each participant received a short debriefing questionnaire focused on obtaining subjective experience data on the real-time monitoring.

Data Management

Use of EMA methodology creates very large and complex data sets. The 33 participants who completed this 2-week monitoring protocol yielded a data set with 167 drinking episodes, 248 temptations to drink, the completion of 1926 signal-contingent assessments, and a total of 462 person-days of monitoring. A data interpreter developed by Saul Shiffman and the Smoking Research Group at the University of Pittsburgh was used to upload and store the ED data. This interpreter creates text files ready to be imported into SAS for Windows (v6.12) for analysis.

Data Analysis

This examination of reactivity focused on three types of data. First, within-participant changes over the course of the two-week EMA monitoring protocol were
explored in the areas of drinking behavior, motivation for change, and alcohol self-efficacy expectancies. Evidence for mean reactivity across participants was tested by using common parametric within-participant analyses. Second, the conceptual implications of participant characteristics influencing reactivity in the absence of mean group-level reactive effects are also of interest. This exploratory study sought better understanding of which participant characteristics predicted reactivity even in the absence of group level effects (Fremouw & Brown, 1980). The following sets of predictors were used in stepwise multiple regression equations to examine participant characteristics that could influence reactivity: addiction severity (ASI alcohol use subscale and number of DSM-IV alcohol abuse and dependence criteria met), alcohol outcome expectancies (global positive and negative alcohol expectancies from AEQ-III at intake), alcohol self-efficacy expectancies (global scores from SCQ), readiness to change (each factor score-Precontemplation, Contemplation, Action, and Maintenance-from the URICA at Time 1), and social desirability. For each regression equation, demographic variables (age and gender) were forced into the equation prior to the stepwise procedure in order to hold constant the variability in the criterion variable associated with age and gender.

Finally, reactivity to each signal-contingent assessment of alcohol urge was addressed. Consistent with contemporary cognitive models of urges suggesting that they are governed by effortful cognitive processing (Tiffany, 1990; Sayette & Hufford, 1994), it was hypothesized that the magnitude of urge at Time 1 within each signal-contingent assessment would predict a greater decrease in urge magnitude. That is, responding to the prompt would require an expenditure of attentional cognitive resources leading to a
decrease in reported urge. Therefore, the higher the urge measurement in Time 1 the
greater the reduction between Time 1 and Time 2 alcohol urges.

Generalized estimating equations (GEE) are ideally suited for data analysis
involving repeated measures designs with correlated data. A primary convenience of the
GEE method is that it is capable of appropriately accommodating an unequal number of
repeated observations that are correlated over time (see, e.g., Zeger & Liang, 1986).

Before using GEE, it is important to determine which correlation structure best
approximates the relationship between the within-participant variable of interest. Three
correlation structures have been considered in EMA studies: independence, exchangeable
correlation, and discretization of an exponentially decaying model or autoregressive
correlation structure. An independence correlation structure makes the assumption that
every individual observation is completely uncorrelated with all other observations within
that same individual. That is, all within-individual observations are independent of all
others. Within the current analysis, it is unreasonable to assume that observations within
participants will be independent of each other. An exchangeable correlation structure
assumes that every observation within an individual is equally correlated with every other
observation in that same individual, regardless of the amount of time separating the
observations. Finally, the autoregressive model makes the assumption that any two
observations within an individual will have a stronger correlation the closer they are
taken over time. Furthermore, this model assumes that the strength of the relationship
between within-individual observations will decay over time. Similar to other
researchers, we used an exchangeable correlation model (e.g., Collins et al., 1998;
Shiffman, Enberg, et al., 1997). Conceptually, an exponentially decaying model appears
to better approximate within-person variables over time; however, the model assumes observations are equally spaced in time, which simplifies the model to a first-order autoregressive (AR1) correlation structure. Our study employed computer generated random signal-contingent assessments, so we expected to observe a wide spectrum of time duration between assessments. The AR1 model cannot account for these continuous-time variables. It is important to note that findings from GEE have been found to be robust even when different types of correlational models are specified (Zegar & Liang, 1992).

Results

Participants

Participants who completed monitoring (n = 33, 53% female) averaged 19.8 (SD = 2.5) years of age. Ethnicity was 97% Caucasian and 3% Native American. Participants reported consuming 48.8 (SD = 38.7) drinks in 9.4 (SD = 5.6) days of drinking in the 30 days prior to the interview. Participants endorsed an average of 2.1 (SD = 1.4) alcohol abuse symptoms and 1.7 (SD = 1.6) alcohol dependence symptoms from the SCID-I interview. Participants scored .16 (SD = .10) on the alcoholism subscale of the ASI. On the depression module of the SCID, participants endorsed an average of 2.0 (SD = 2.4) symptoms. Twenty-four participants met DSM-IV (APA, 1994) criteria for alcohol abuse, four met DSM-IV criteria for alcohol dependence, and five met DSM-IV criteria for major depression.

Compliance with Monitoring

Participants were remarkably compliant with the monitoring. For example, participants responded to 86% of the 2248 random prompts administered over the 2-week
monitoring protocol. These results are parallel to recent EMA studies by Collins et al. (1998) and Stone et al. (1997) that report overall participant compliance to random prompts as 85% and 90%, respectively. In the present study, social desirability scores did not predict compliance as measured by the total number of missed random prompts ($r = .18$, $p = .35$).

**Behavioral Reactivity**

Behavioral reactivity was examined by contrasting participants' real-time reports of drinking behavior to their data from the 30-day TLFB administered pre-monitoring. The direction of change suggested a decrease in both mean number of drinks per week ($M = 1.8$, $SD = 8.1$) and drinking days ($M = 1$, $SD = 1.2$); however, a paired comparison $t$-test yielded no significant differences for either drinks per week or days drinking per week. Real-time reports of drinks per week ($r = .86$, $p < .001$) and drinking days ($r = .84$, $p < .001$) were significantly correlated with the two-week timeline follow-back data.

We examined whether the participant characteristics outlined above would predict behavioral reactivity using a stepwise multiple regression analysis. After controlling for age and gender, participants' Action scores from the URICA were inversely related to behavioral reactivity (partial $r^2 = .44$, $t = 3.1$, $p = .01$). Participants who reported they were actively trying to ameliorate an alcohol problem were less likely to demonstrate behavioral reactivity by reducing their average drinks per week. Alcohol use subscale scores from the ASI were marginally related to behavioral reactivity (partial $r^2 = .30$, $t = 2.2$, $p = .054$) such that participants reporting more alcohol problems in the 30 days prior to the onset of the study were somewhat less likely to reduce their drinking during monitoring than those participants reporting less of a problem with alcohol.
Motivational Reactivity

At intake to the study, participants’ scores on the URICA indicated that 46% were in Precontemplation, 33% were in Contemplation, and 21% were in the Action stages of change. This distribution remained largely unchanged at the end of the study with 49%, 42%, and 9% of participants falling into the Precontemplation, Contemplation, and Action stages of change, respectively. No participants were identified as being in the Maintenance stage of change at either assessment.

Motivational reactivity was examined in two ways. First, each of the four factors from the URICA were analyzed using a repeated measures MANOVA. This 2 (Time) X 4 (Factor: Precontemplation, Contemplation, Action, and Maintenance) repeated measures MANOVA demonstrated that the main effect for Time approached significance ($F [1,32] = 3.95$, $p = .055$). The analysis also illustrated a main effect for Factor ($F [3,30] = 39.96$, $p < .0001$). The Time X Factor interaction was not significant ($F [3,30] = 2.0$, $p = .13$).

Next, the URICA data was analyzed using the readiness to change (RtC) index score. A repeated measures ANOVA with the RtC index as the dependent measure revealed a main effect for Time ($F [1,32] = 5.5$, $p = .03$). Participants RtC index score decreased between pre- ($M = 40.7$, $SD = 15.2$) and post-monitoring ($M = 32.1$, $SD = 20.3$).

In order to examine whether individual differences related to motivational reactivity, a stepwise multiple regression analysis was conducted using the predictors outlined above (excluding the readiness to change factors themselves) and change on the RtC index as the dependent variable. This analysis revealed no significant effects.
**Self-Efficacy Reactivity**

First, SCQ total scores were examined for change between pre- and post-monitoring. This within-participant ANOVA revealed no effect for Time ($F[1,24] = .56, p = .46$). Participants' global alcohol self-efficacy scores did not differ between Time 1 ($M = 65.9, SD = 17.8$) and Time 2 ($M = 64.3, SD = 16.6$).

A stepwise multiple regression analysis with the variables outlined above as predictors and change score on the SCQ as the dependent measure revealed that, after controlling for gender and age, the **Precontemplation** factor from the URICA predicted change in alcohol self-efficacy expectancies (partial $r^2 = .64, t = 6.8, p = .001$). The stronger the participants identified themselves as having no desire to change an alcohol problem, the more likely they were to increase their self-efficacy expectancies as a function of the real-time monitoring. Total negative alcohol expectancies approached significance in predicting change in alcohol self-efficacy expectancies (partial $r^2 = .33, t = 2.2, p = .052$). Participants with strong negative alcohol outcome expectancies were more likely to demonstrate a decrease in their alcohol self-efficacy expectancies over the course of the study.

**Urge Reactivity**

For each participant, we created a change score computed as the difference between their self-reported urge at the beginning (T1; $M = 3.86, SD = 1.3$) and end (T2; $M = 3.57, SD = 1.3$) of the signal-contingent assessments. Participants urge scores significantly decreased during the signal-contingent assessment ($M = .29, SD = .60, t[32] = 2.8, p = .01$). In addition, an Urge Change Ratio (UCR) was created by using the T1 urge in the numerator and the T2 urge in the denominator. For example, if a participant
endorsed a T1 urge magnitude of 10 and T2 urge magnitude of 5, then the UCR would be 2. Similarly, if a participant endorsed an urge magnitude of 5 at both T1 and T2, the UCR would then be 1. To provide a more stable test of change, a t-test comparing the mean UCR with a test value = 1 was performed. This also resulted in a significant change score (\(M = 1.28, SD = .33, t [32] = 4.9, p < .001\)).

Moreover, generalized estimating equations (Zeger & Liang, 1986) using an exchangeable correlational model found that the magnitude of the urge at T1 was positively related to urge reactivity (\(\beta = .34, Z = 13.7, p < .001\)). That is, the greater the initial urge to drink at the outset of the signal-contingent assessment, the greater the urge reduction at T2. It was also examined whether urge reactivity would change over time in the study and found this not to be the case (\(\beta = -.003, Z = -.68, p = .50\)). Participants’ urge reactivity showed no signs of increasing or decreasing over the 2-week real-time monitoring protocol.

In order to examine whether participant characteristics influenced the mean level of urge reduction across participants, we ran a stepwise multiple regression analysis with the mean change urge for each participant across the random prompts as the dependent measure. This analysis revealed no significant predictors for change in urge. However, when using the UCR as the criterion variable and controlling for age and gender, total negative outcome expectancies from the AEQ-III (partial \(\tau^2 = .46, t = 3.0, p = .011\)) and the Maintenance stage of change (partial \(\tau^2 = .38, t = -2.58, p = .026\)) predicted urge reactivity. Participants who strongly endorsed negative outcome expectancies were more likely to demonstrate urge reactivity while items on the URICA indicating past successful changes in alcohol behavior predicted a decrease in urge reactivity.
Self-Reported Reactivity

At the end of the study, participants completed a measure assessing their subjective impressions of whether EMA affected their drinking. Twenty-nine percent of participants (n = 9) believed that the real-time monitoring caused them to decrease their drinking while 10% (n = 3) reported that monitoring caused them to increase their drinking. Sixty-one percent (n = 19) of participants reported that the real-time monitoring did not affect their drinking behavior. For those reporting decreased drinking, the monitoring had an estimated average impact of 5.7 (SD = 1.8) on an 11-point scale (0 = no impact to 10 = a great deal of impact). The participants who indicated monitoring caused them to increase their drinking estimated the impact to be a 4.7 (SD = 1.5) on an 11-point scale. Of all the participants who thought that real-time monitoring affected their drinking, either an increase or decrease (n = 12, 39%), also demonstrated behavioral reactivity for both drinks per week (r = -.70, p = .01) and days drinking per week (r = -.69, p = .01). The more the participant acknowledged a reactive decrease in their alcohol consumption as a result of monitoring, the greater the decrease in self-reported drinking over the course of the study.

Discussion

This study examined five types of reactivity to EMA using a sample of undergraduate problem drinkers. It was hypothesized that participants' real-time reports would produce significantly less days drinking per week and number of drinks consumed per week compared to their retrospective reports for the 30-days before the study (Collins et al., 1998; Litt et al., 1998). Real-time monitoring of drinking behavior for short periods (i.e., 2-weeks) failed to produce significant overall changes in drinking quantity.
and frequency. Evidence for behavioral reactivity producing decreased drinking during real-time monitoring has been shown in studies using longer monitoring protocols (e.g., Collins et al., 1998). This suggests that the short monitoring protocol used in this study may have been partially responsible for the lack of behavioral reactivity.

The majority of our participants did not report that the monitoring affected their drinking behavior. The minority (39%) of participants who believed that monitoring affected their drinking also showed changes in their EMA-assessed drinking. This is a marked departure from the findings of Litt et al. (1998), who found that fully 70% of their alcoholic participants self-reported decrease in drinking due to monitoring but demonstrated no significant change when their actual drinking data was compared to a non-monitoring group. This contrast in findings is likely a function of the significant differences between the two samples (e.g., addiction severity, motivation for change). However, researchers interested in reactivity should consider these divergent findings prior to assessing various alcohol using populations. Further research into the agreements and discrepancies between participant perceptions versus the actual impact EMA has on drinking behavior as measured in real-time is warranted.

Contrary to the hypothesized relationship between monitoring and readiness to change, participants were slightly less likely to endorse being ready to change their drinking behavior after completing the real-time monitoring protocol. This was most evident in the analysis using the RtC index score (DiClemente et al., 1999). The majority of our undergraduate sample, however, viewed their drinking as non-problematic, despite scoring as hazardous drinkers on the AUDIT and consuming approximately five drinks per drinking occasion. It is unknown whether participants seeking help for an alcohol
problem would show a similar pattern of motivational reactivity to EMA. It could be the case that participants in the contemplation or action stages of change would manifest greater therapeutic reactivity to EMA (Abrams & Wilson, 1979).

We found no evidence of self-efficacy reactivity as a function of the real-time monitoring. Alcohol self-efficacy expectancies have been shown to change over the course of 4-week treatment (e.g., Brown, Carrello, Vik, & Porter, 1998). Somewhat surprisingly, this group of undergraduate drinkers scored relatively low on the SCQ, averaging approximately 65% confidence to resist heavy drinking in a variety of situations. This is remarkable in that SCQ global scores less than 50% have been found to predict relapse among inpatient alcoholics (Greenfield, Hufford et al., 2000). However, less is known about the stability of alcohol self-efficacy expectancies among participants not involved in alcoholism treatment.

Our findings regarding urge reactivity were consistent with cognitive models of urges that suggest they require effortful cognitive processing (Tiffany, 1990; Sayette & Hufford, 1994). The data revealed that the magnitude of alcohol urge assessed at the beginning of each signal-contingent assessment decreased over the course of the assessment. Furthermore, this lower level in self-reported alcohol urge was greater in subjects with greater levels in pre-assessment urge. This finding may have clinical implications for studies using EMA to produce and/or maintain therapeutic effects. It may be possible to program palmtop computers to magnify the reductions in urge magnitude observed in the present study. However, this finding is also consistent with regression to the mean effects; as alcohol urges at T1 increased, greater opportunity existed for them to decrease at T2.
Like other addiction researchers with an interest in reactivity (Fremouw & Brown, 1980), we examined a variety of individual difference characteristics, including demographic factors, addiction severity, readiness to change, alcohol outcome and self-efficacy expectancies, and social desirability as influences on reactivity. The results demonstrated that even in the absence of mean group-level reactive effects, certain participant characteristics were related to reactivity.

Factor scores on the URICA (i.e., stages) predicted behavioral, self-efficacy, and urge reactivity. Participants who reported they were taking some Action to change an alcohol problem were less likely to decrease their drinking behavior simply as a function of monitoring. Similarly, participants who resisted acknowledging a problem with alcohol (Precontemplation) were more likely to show increased confidence in their ability to resist heavy drinking in high-risk situations after the real-time monitoring. This suggests that real-time monitoring may serve as an arena for precontemplators to test (in this case, successfully) their alcohol self-efficacy expectancies. Further, individuals endorsing URICA items associated with the Maintenance stage of change (indicative of successful application of past alcohol behavior change techniques) were less likely to demonstrate urge reactivity. This implies that individuals with a history of alcohol behavior change attempts are less likely to have their urges affected by real-time monitoring. Taken together, these findings suggest that the motivation for change is an important influence on reactivity.

This study has a number of limitations. First, by not having a no-monitoring control group, these results should be viewed with caution. Typically, control groups have not been included in real-time studies, but even if they were it is difficult to
understand how their inclusion might yield strong evidence for or against reactive effects. In the present study, it is possible that coming to a research laboratory and filling out self-report measures and completing interviews could have been partially responsible for our findings. Several issues mitigate this concern, including that test-retest studies of questionnaires rarely find significant effects simply as a matter of administering the measure. Moreover, the lack of behavioral reactivity indicates that most participants did not significantly change their behavior as a result of completing the strongest test case: self-report measures and a real-time monitoring protocol. Lastly, there is the question of interpretation of data from a control group. Affleck et al. (1999) note that the inclusion of any single control group to understand reactivity to EMA is unlikely to be helpful. For example, what is the key variable being controlled: carrying the ED, being prompted by ED, receiving feedback on ED compliance; recording drinking versus mood states or other variables on ED?

Many studies using EMA in addictive behaviors research focus on treatment-seeking populations (e.g., Collins et al., 1998; Litt et al., 1998; Shiffman et al., 1997). It is unknown whether the general absence of reactive effects observed among these undergraduate problem drinkers would also be absent in a treatment-seeking sample. The relatively short monitoring protocol used in this study also differentiates it from studies using 4- to 8-week monitoring protocols (e.g., Collins et al., 1998; Shiffman et al., 1997b). It is possible that increased or decreased reactive effects could emerge in studies using longer monitoring protocols. However, time trends in the data can be explained by factors other than reactivity and must be examined carefully. For example, participants in this study were remarkably compliant to a demanding EMA protocol (e.g., participants
responded to 86% of the random prompts). However, longer monitoring periods could show decreased compliance over time. Still, the high correlation between the real-time and retrospective reports of drinking behavior on the TLFB suggest that our participants were not only compliant but also relatively accurate in terms of their self-report reported quantity and frequency of alcohol consumption (Carney et al., 1998). This pattern of compliance to a rigorous monitoring protocol is in agreement with finding from a number of studies suggesting that participants' momentary experiences can be reliably and accurately measured using EMA (Shiffman & Stone, 1998).

This research has important implications, both for EMA researchers and the study of self-report more generally. Although other EMA and self-monitoring studies have addressed reactivity (e.g., Collins et al., 1998; Cruise et al., 1996, Sobell et al., 1989), this is the first study to systematically examine reactivity as a multidimensional construct involving behavioral, motivational, and cognitive components. Our findings indicate that reactive effects to EMA are generally small in magnitude. What remains to be seen is whether reactivity can be enhanced by using EMA for therapeutic gain. Additional research is needed to more fully understand the reactive effects of EMA on treatment-seeking populations. More specifically, research explicitly addressing the nature of individual differences and how participant characteristics, such as motivation to change a behavior, can influence the reactivity to EMA. Instead of viewing reactivity as a unitary phenomenon, future research should continue to examine how intensive self-monitoring can impact behavioral, cognitive and motivational variables important to the study of addictive behaviors.
References


