Published in final edited form as:
Psychol Assess. 2009 December ; 21(4): 486–497. doi:10.1037/a0017074.

Ecological Momentary Assessment (EMA) in Studies of Substance Use

Saul Shiffman
University of Pittsburgh

Abstract

Ecological Momentary Assessment (EMA) is particularly suitable for studying substance use, because use is episodic and thought to be related to mood and context. This paper reviews EMA methods in substance use research, focusing on tobacco and alcohol use and relapse, where EMA has most been applied. Common EMA designs combine event-based reports of substance use with time-based assessments. Approaches to data organization and analysis have been very diverse, particularly regarding their treatment of time. Compliance with signaled assessments is often high. Compliance with recording of substance use appears good, but is harder to validate. Treatment applications of EMA are emerging. EMA captures substance use patterns not measured by questionnaires or retrospective data, and hold promise for substance use research.

Keywords

Ecological momentary assessment; substance use; drug use; tobacco; alcohol

Ecological Momentary Assessment (EMA; Shiffman, Stone, & Hufford, 2008; Stone & Shiffman, 1994) methods are designed to help researchers get ecologically valid data about behavior, thoughts, and feelings over time, while avoiding the pitfalls of retrospective recall. As described elsewhere (Shiffman et al., 2008), EMA involves repeated administration of assessments, in real time (or close to it), in subjects' natural environments.

The prototypical EMA study collects data about subjects' states and behavior at particular moments throughout the day, while subjects go about their daily lives in their natural environments (see Shiffman, 2007; Shiffman et al., 2008; Stone & Shiffman, 1994). Often, particular moments are targeted for assessment because they represent events of interest, such as drug use, which the subject is asked to monitor and record. Other moments are sampled for assessment based on a regular or random time schedule, so as to obtain a representative sample of their state. Event- and time-based assessments are often combined and compared. This paper addresses the application of EMA to studies of drug use and abuse. (I will refer to "drug use" as a general term, incorporating both legal and illicit drugs, and both use and abuse.)

Saul Shiffman, c/o Pinney Associates, 210 N. Craig St (Suite 320), Pittsburgh, PA 15213-1516, shiffman@pitt.edu, (412) 687 5677.

Suitability of EMA for drug use research

It is no accident that EMA methods have seen particularly wide adoption in studies of drug use, because EMA methods are particularly well-suited to studying drug use. Drug use itself is a discrete, episodic behavior that lends itself to event-oriented recording, making EMA a useful method for tracking its frequency and distribution over time. Moreover, many theories of drug use emphasize the role of the immediate situation in drug use, with emphasis on immediate internal experience (e.g., the user's mood, craving, or withdrawal state) and external situational factors (e.g., the presence of the target substance, substance-related cues, social pressures to use), making a situational and momentary focus natural for evaluating theory. Theory has similarly emphasized the role of the acute effects of drugs (i.e., reinforcement, euphoria, relief of stress), which also lend themselves to momentary assessment. Even theories that have not specifically emphasized conditions at the moment of drug use have often emphasized dynamic processes playing out over time (e.g., stress, abstinence violation effects) as important influences on drug use. These factors make EMA particularly well-suited to study drug use and abuse.

Perhaps because of the importance of situational cues and social context, drug use can also be hard to study in the laboratory, which typically does not recreate the contexts associated with use. Alan Marlatt's creation of a "bar lab" (Marlatt & Gordon, 1985) represents one creative attempt to overcome this limitation, but points to the difficulty of modeling drug use in the lab. That difficulty is likely magnified when studying illicit drug use, since subjects may be reluctant to display such socially-sanctioned behavior while under close scrutiny in a laboratory context. The challenges of modeling drug use in the lab have increased the impetus to study drug use in the field, using EMA.

Challenges for EMA in drug use research

Just as EMA seems particularly suited to drug use research, drug use research also raises particular and sometimes unique issues for the application of EMA. One hypothetical objection to using EMA methods to study drug use is the expectation that drug users will not be willing or able to comply with the substantial demands of EMA protocols. Drug abusers and addicts tend to lead disorganized lives, and to suffer from comorbid psychopathology and social pathology, and are not known for responsibility and adherence to schedules or instructions. This has raised skepticism that drug users will be able or willing to comply with EMA protocols, and a related concern that such subjects would either steal or break the PDAs or cell phones provided to them for EMA data collection. These concerns have impeded adoption of EMA methods in studies of drug use, particularly for illicit drugs. (In contrast, there has been substantial adoption in tobacco and alcohol research.)

As understandable as these concerns are, the empirical experience is reassuring. In perhaps the most startling demonstration, Freedman, Lester, McNamara, Milby, and Schumacher (2006) were able to get good compliance from homeless crack-cocaine addicts enrolled in a study using cell phones. The study is all the more striking because this difficult population was subject to a particularly demanding protocol, scheduled to receive telephone calls every 3 hours all through the day and night for 14 days. Exact compliance could not be reported because the system did not track missed calls or distinguish them from failures of the cell phone system or calls made while people were asleep, but compliance appears to have been good: Even with the caveat that more data were missed because of failures of the cell-phone company than because of participants' non-compliance, 77% of calls between 6 am and midnight (which undoubtedly includes hours that subjects were sleeping) received responses. Moreover, only 10% of participants dropped out, and an average of only 1 day early, and only one of 30 cell phones was lost. In another key demonstration, Epstein et al. (2009) had subjects in treatment

for heroin and cocaine use track drug use and respond to random prompts for up to 6 months. Subjects were retained for a median of 162 days, compliance was good, and PDAs were only rarely lost or damaged. Similarly, Johnson, Barrault, Nadeau, & Swendsen (in press) demonstrated the feasibility of EMA data collection with female opiate addicts, and Hopper et al. (2006) with regular ecstasy users who also engaged in use of alcohol, marijuana, cocaine, and hallucinogens. The fact that the illicit drug users in the above studies reported frequent episodes of drug use also addresses the concern that subjects might be reluctant to record and report illegal behavior. These studies suggest that illicit drug users do actually report copious use.

A further challenge in collecting data about episodes of psychotropic drug use is the concern that subjects' intoxication might make it difficult for subjects to complete data entry or to do so in a valid way. I know of no published study that addresses this question. However, in an unpublished study, Lorraine Collins and I established that subjects could enter data accurately on a PDA while intoxicated with alcohol. However, the question of whether perceptions are blurred or biased by intoxication with alcohol or other drugs remains for EMA as well as for laboratory studies where subjects are intoxicated.

Illustrating application of EMA methods to studies of drug use

The application of EMA methods, designs, and analyses to drug use is perhaps best conveyed through illustrative examples. Accordingly, this section reviews select EMA studies of drug use, focusing on tobacco and alcohol, where much of the EMA literature has focused. (I do not aim for comprehensive coverage of the EMA literature.) Seeing the methods applied in the context of specific studies and papers will place the methodological issues properly in the context of substantive research questions the investigators are trying to address, and the substantive insights they generate. It is a truism that research methods must be dictated by objectives and research questions, and the review will hopefully demonstrate the linkage between the two.

Although, as we shall see, EMA studies of drug use are quite varied, the designs used for the EMA data collection actually fall into just a few approaches. The most common design uses event-based recording to capture drug use occasions, and time-based prompted assessments (usually randomly-scheduled) to capture subjects' experience around those events. That is, subjects are asked to initiate an entry and assessment when they engage in drug use, and in parallel, they are prompted at random times to complete similar assessments when not using drugs. This approach has been used in studies of both ongoing drug use and relapse.

A second design does not ask users to initiate event entries when drugs are used, but relies solely on time-based assessments, at which time subjects are asked retrospectively about drug use and other variables. The less common variation, which I'll call the interval approach, tries to narrow the interval for reporting of drug use by assessing subjects multiple times per day, either at regular intervals or at random times within pre-specified time blocks. The most common variation is the daily diary, where subjects complete assessments just once a day, and report both drug use and other experiences retrospectively for the whole day. Daily diary designs are not really "real time," and rely substantially on retrospection, which raises issues of retrospective bias. However, daily diaries are very common, particularly in alcohol research, and they do share some of the other essential features of EMA methods, that is, repeated assessments over time, in the natural environment, and a consequent focus on within-subject variation related to changing psychological states or environmental circumstances. Accordingly, I will include some daily diary studies in this review, while emphasizing more real-time methods.

Within this relatively narrow range of designs, EMA studies display a wide array of approaches to answering their respective questions. Much of the variability and creativity lies in how the investigator organizes and analyzes the data to address the study's particular research question. EMA data provide the investigator with a rich series of observations arrayed over time. This data stream then allows investigators a wide range of options for how to conceptualize their research question and frame, structure, and analyze the data in a way that fits their research question. I will try to highlight some of these variations in approach.

To address the use of EMA methods in drug abuse research, I will focus on two research topics: the role of affect in smoking and drinking, and the process of relapse in smoking and drinking. These topics were selected because they represent the biggest mass of EMA research on drug use, and because they allow the review to illustrate the diverse ways EMA has been applied to understanding drug use.

The role of affect and mood in ad libitum smoking and drinking

The enduring interest in the relationship between mood and drug use – and mood and smoking or drinking, specifically – derives from important theories that suggest that users smoke and drink in order to diminish or control negative affect. This idea is typically supported by users' global questionnaire reports, where a majority of smokers (Shiffman, 1993) and a majority of problem drinkers (Cooper, Agocha, & Sheldon, 2000) report using tobacco and alcohol to manage distress, and cite this as a major motive for smoking or drinking. Given the difficulties drawing valid inferences from subjects' global retrospective reports (in a separate section, I address the validity of these reports), it is natural that researchers have applied EMA methods to study the relationship between mood and smoking/drinking in the field using EMA.

Smoking and mood

Diaries have been used in smoking research and treatment for a long time, but EMA studies of smoking present the investigator with a number of challenges. Older studies (Paty, Kassel, & Shiffman, 1992; Surawy & Cox, 1987) represent a traditional, perhaps naïve, use of diary methods. These studies used simple event diaries, giving smokers paper diaries on which to complete mood assessments whenever they smoked, and correlating these data with individual difference variables to assess different smoking patterns. However, studies that are limited to event monitoring and thus generate data only about smoking occasions, cannot really address the associations between smoking and mood – e.g., they cannot distinguish a subject who is pervasively depressed from a subject who specifically smokes when depressed. As Paty et al. (1992) argue, one cannot establish associations between an event and its antecedents without data on non-events, just as one cannot gain insight into the causes of disease without studying non-diseased controls in a case-control design. Thus, these studies are relatively uninformative regarding the relationship between mood and smoking.

To address this limitation, investigators have developed case-control or, more accurately, case-crossover (because it's within-subjects; Maclure & Mittleman, 2000) designs that also collect data on non-smoking moments, by using time-based assessments, prompting subjects for assessments at randomly-scheduled times. For example, Shiffman et al. (2002) had smokers carry PDA-based diaries, on which they recorded their smoking occasions. In parallel, the PDAs were programmed to "beep" subjects at non-smoking times (between cigarettes, as it were) around 5 times per day, at which time they completed similar assessments. Similar designs were implemented by Carter et al. (2008) and others. These designs allow the analysis to analyze the situation surrounding smoking occasions by contrasting it to the subjects' general state when they are not just about to smoke.

A variety of schedules and arrangements have been used for assessing non-smoking moments. One study (Moghaddam & Ferguson, 2007) just asked smokers to record their non-smoking state twice a day at vaguely specified times, "morning" and "evening." It seems unlikely that these subject-selected occasions would provide representative coverage of subjects' non-smoking states. Most studies have prompted throughout the day, with some applying fixed schedules of prompting, e.g., every 45 minutes (Shapiro, Jamner, Davtdov, & James, 2002). The most common and strongest design uses randomly-scheduled assessments to get a random time sample of subjects' states.

Studies have varied widely in the frequency of such time-based assessments. While several studies prompted subjects about 5 times a day (Shiffman et al., 2002; Shapiro et al., 2002), others (Delfino, Jamner, & Whalen, 2001) have prompted subjects 30 or more times per day, usually for short periods. Event-monitoring of smoking also presents a substantial subject burden, because of the sheer frequency of smoking in most subject populations (often 15–30 or more cigarettes per day in many studies). Whereas simply recording the event imposes a minimal burden, completing an assessment on each occasion can add up to a substantial burden. One solution has been to ask subjects to record all events, but only administer assessments for a random sample of smoking occasions. For example, several studies (Beckham et al., 2008; Shiffman et al., 2002; Shiffman, Paty, Gwaltney, & Dang, 2004) have randomly sampled about 5 smoking occasions per day. Having the EMA computer sample events at random ensures that the resulting assessments are representative. Other studies have selected occasions in other ways: e.g., Cooney et al. (2007) only assessed the first cigarette reported in each of four assessment intervals. It is less clear whether this sample of smoking occasions is representative.

While most studies have relied on event-based entries to track smoking, there are exceptions: Todd (2004) assessed smokers at regular intervals every four hours and asked about their smoking, as well as negative events and stress, over the preceding four-hour time-block. Thus, the sampling was based on fixed intervals and adopted a retrospective "coverage" strategy, that tries to cover the whole day by using recall to collect data for non-overlapping intervals (see Shiffman et al., 2008). This approach relies on subject recall, which can be biased, and specifically allows for confounding between recalled smoking and affect – i.e., subjects may "recall" more smoking when they recall feeling stressed.

The diverse studies on smoking and mood have also demonstrated different approaches to analysis. Most have simply compared mood between collections of smoking and non-smoking moments (Shiffman et al., 2002; Shiffman et al., 2004; Carter et al., 2008) to assess the within-subject associations between mood and smoking. Some authors have used the longitudinal aspect of EMA data to perform truly prospective analyses. For example, Delfino et al. (2001) used the dense EMA data they collected (assessments every 20 minutes) to assess whether mood ratings from one assessment predicted the likelihood of smoking in a subsequent assessment. (See also Epstein et al, 2009). Interestingly, when Todd et al. (2005) performed both cross-sectional and prospective analyses, they found that the associations between stress and smoking that were seen in cross-section disappeared in the prospective analyses, suggesting the utility of using an EMA data stream to gain a prospective perspective. Beckham et al. (2008) illustrated the utility of examining between-subject factors as moderators of the within-subject associations assessed by EMA: they showed that the relationships between mood and smoking differed among patients with PTSD versus normal controls.

Substantively, the striking thing about the EMA studies of smoking and mood is that they have not broadly supported the hypothesized robust link between negative moods and smoking. Several studies show no relationship at all between antecedent mood and smoking (Carter et al., 2008; Shiffman et al., 2002; Shiffman et al., 2004). Others have reported mood effects, but these have been very small (e.g., 0.07 points on a 5-point scale – or about 1% of the scale span

– and contradictory (Shapiro et al., 2002). The dominant finding that negative moods are not prominent triggers for smoking stands in stark contrast both to smokers' global reports on questionnaires (Shiffman, 1993) and to the large body of theory (based largely on those questionnaire data) that gives mood a prominent role in motivating and triggering smoking (Baker, Piper, McCarthy, Majeskie, & Fiore, 2004). While more sophisticated models of the relationship between mood and smoking may yet find the expected link, this illustrates the capacity of EMA methods to test in the real world, and in great detail, relationships that are posited on the basis of theory and/or on the basis of subjects' global reports.

Alcohol use

Interest in the role of affect in alcohol use has also been a key driver of EMA methods in alcohol use. The idea that people drink in order to reduce stress or mitigate negative affect, has been theoretically prominent, and has been addressed in multiple EMA studies.

It is interesting to note the different ways in which EMA studies of drinking seem to differ from those of smoking. One striking difference is that daily end-of-day diaries predominate over real-time momentary assessment. Perhaps this reflects the lower frequency of drinking, and a concomitant expectation that drinking episodes are more easily recalled without bias. Whether this is true remains to be established. In any case, even if drinking itself could be accurately recalled, alcohol studies often rely on recall of affective experience over the day; given the volatility of affect, and the sensitivity of recall to state-specific recall bias, there is good reason to be concerned about the accuracy of recall of one's mood over the past day. In any case, here I concentrate on momentary assessments, while including some daily diary work.

In studies using event-based monitoring, investigators have defined the 'event' or unit of recording in different ways. Some studies (e.g., Hufford, Shields, Shiffman, Paty, & Balabanis, 2002; Litt, Cooney, & Morse, 1998) have asked subjects to record each drink as a distinct event, paralleling the smoking studies. Others, recognizing that drinks often cluster into drinking episodes, and perhaps being concerned about the subject burden of recording each drink, have had subjects record data for each episode of drinking, foregoing recording of individual drinks within the episode (Collins et al., 1998). Obviously, these two approaches address different questions: the latter design can address what initiates an episode of drinking, but not what triggers each successive drink within an episode. How relevant events are defined is also sometimes determined at the time of analysis rather than in data collection: for example, Collins et al. collected data on all drinking episodes, but focused their analysis on the antecedents of episodes of "excessive drinking," defined as having ≥ 5 drinks in an episode. To provide contextual or comparative data for these event entries, several studies have concomitantly collected data on non-drinking moments through random sampling, implementing a case-crossover design (e.g., Litt, Cooney, & Morse, 2000; Collins et al.).

However, even within similar data collection methods, different studies have framed and analyzed the data differently. The studies differ particularly in how they summarize observations over time and how they deal with time in the analysis. Perhaps because drinking (unlike smoking) tends to be concentrated on particular days and times of day, alcohol studies have sometimes divided the day (evening vs. daytime) and week (weekend vs. weekday) into blocks, and to use these blocks as a basis for prospective analyses that address whether mood in time block A influences drinking at a subsequent block B. Starting with data collected and summarized at the day level, Hussong, Hicks, Levy, & Curran (2001) blocked data into weekday and weekend observations, and analyzed how weekday mood predicted weekend drinking (and vice versa). This implicitly collapsed the influence of time except at the weekday-weekend boundary, essentially assuming that subjects' Saturday drinking is influenced roughly equally by their Monday mood as by their Wednesday mood. In contrast, a daily diary study

by Armeli, Todd, Conner, and Tennen (2008), who were also interested in how drinking and mood interrelated over the week, treated time (days, in this case) as a continuous quantity that exercises contiguous influence. Armeli et al. focused on how daily mood influenced when subjects started drinking as the week wore on towards and into the weekend. In a survival (or time-to-event) analysis, they modeled the initiation of drinking, predicted from the cumulative mood ratings over the preceding days (treated as a time-varying covariate). Notably, Hussong et al. analyzed the quantity of drinking, while Armeli et al. analyzed the time of onset of drinking.

In a parallel to this in within-day data, Swendsen et al. (2000) collected mood data during three time blocks each day, with the final one scheduled between 8:00 – 9:30 pm. Evening (8:00–9:30) mood was then used to predict the volume of later drinking. They found that both happiness and nervousness predicted more drinking. In contrast, Todd, Armeli, and Tennen (in press) collected data in an identical way, but used morning mood (10:00–11:30 am) to predict the time of onset of later drinking, finding that morning nervousness and anger were associated with earlier onset of drinking (e.g., around 5 PM in high-nervous days, 9 pm on low-nervous days). These varied approaches illustrate that the unit of data collection need not be the unit of analysis, and exemplify how a stream of time-bound EMA data can be organized and analyzed in a variety of ways that treat time differently, and thereby focus the analysis on slightly different questions dictated by different theoretical models being tested.

The alcohol literature also illustrates an issue that concerns all EMA studies, but is an especially acute concern for alcohol and drug use studies, namely the hours of coverage for EMA assessments. Investigators sometimes restrict the hours during which the EMA protocol is active, usually in order to avoid “beeping” subjects when they are asleep. For example, Hussong et al. (2001) assessed mood and drinking only between the hours of 10 am and 10 pm, and Cooney et al. (2007) only between 8 am and 10 pm. However, these times of day may not be representative, particularly as alcohol consumption tends to occur later, and mood, activity and social setting also varies by time. Thus, it seems important for studies to assess the full range of a subject’s waking hours. Even having the protocol active during the subject’s “typical waking hours” (McCarthy, Piasecki, Fiore, & Baker, 2006; Todd, 2004) is not an adequate solution, because a subject’s bedtime varies, and is likely to be later on the days they use drugs. A more flexible and complete solution is to provide a way for subjects to turn off the computer when they are going to sleep, essentially treating it like an alarm clock (e.g., Shiffman et al., 2002).

Whereas much of the drug research has focused on assessments at the time of drinking, or at random times, there is sometimes reason to schedule assessments at other times. For example, Muraven, Collins, Morshimer, Shiffman, and Paty (2005) administered an assessment each morning to capture hang-over effects from the previous night’s drinking, as well as regret or guilt about that drinking. The point is that the appropriate timing of EMA assessments is determined by the research questions and theoretical and empirical considerations of the best time to measure the relevant effects.

Processes of relapse

The study of relapse processes involves a very different set of challenges from those encountered in studying ongoing use, but is nevertheless also particularly well-suited to EMA methods. Like ongoing drug use, relapse involves some specific, concrete events–lapse episodes, in which the drug is used after a period of abstinence. Unlike ongoing drug use, where drug use is at steady state and each observed episode of use might be regarded as simply another sample of a use event, the relapse process includes milestone events, such as the very first lapse after a period of abstinence or the threshold of full relapse. The relapse process also plays out

longitudinally over time, as the participants initiate abstinence, then experience an initial lapse, then further lapses, and eventually complete relapse. In examining this unfolding process over time, investigators often look both at fast-moving, local processes, like the immediate contexts that precipitate lapse episodes, and also slower-moving "background" processes, like the buildup of stress or the progressive exhaustion of coping resources that may set subjects up for lapsing (Shiffman, 1989). For both of these, the detailed record-over-time provided by EMA data provides a powerful analytic tool, and its use for studies of relapse has been advocated by McKay, Franklin, Patapis, and Lynch (2006).

Smoking lapses and relapse

Investigators of relapse have long been interested in the details of lapse episodes and the situations in which they occur (Marlatt & Gordon, 1985). Such data were traditionally collected retrospectively, by distant recall, often months after the event, after relapse had already occurred (e.g., Marlatt & Gordon). The long-recognized potential for recall biases to creep into these reports, along with an analysis (discussed later) showing that such reports were inaccurate and biased (Shiffman et al., 1997) have moved investigators towards the use of EMA to capture data about initial lapse episodes.

Exemplifying this approach, Shiffman, Paty, Gnys, Kassel, and Hickcox (1996) had 108 smokers who had quit smoking (>24 hours without smoking) initiate an event recording if they lapsed (i.e., smoked at all), upon which the electronic diary they were carrying administered an extensive assessment of the context: location, activity, mood, consumption of food and drink, and so on. Paty et al.'s (1992) argument that drug use episodes cannot be understood without contrast to moments of non-use applies equally well to lapse episodes – without a comparator or "control," we cannot know what is particular to lapse episodes, versus being typical settings for the person or the person's experience during withdrawal and a struggle to maintain abstinence. As in studies of ongoing use, assessed moments of non-use provide useful control data.

In the case of lapse episodes, investigators have also used another control: "Temptations," or moments when the subjects experienced strong temptation to use or came close to using but did not actually lapse (Shiffman, 1982; O'Connell et al., 1998). Shiffman et al. (1996) illustrated the use of both comparators, reporting that lapse episodes differed in mood, activity, and setting from both random moments and temptation episodes sampled in the days preceding an initial smoking lapse. In these studies, subjects initiate recording of temptations as "events." This can be challenging, because temptations are not defined as objectively as episodes of use. Moreover, such subjective events are even less susceptible than episodes of use to objective verification and confirmation of compliance. In any case, subjects in such studies do record multiple temptations (sometime > 200; Tindle, Shiffman, Paty, & Dang, 2006). Analyses that compare lapse episodes to temporally proximal control observations have the virtue that they isolate factors that seem unique to lapses, controlling for between-subject differences (because the comparators come from the same subject) and also for relatively slow-changing states within a subject (e.g., withdrawal), because the comparator observations are close in time to the focal lapse episode. In the Shiffman et al. study, for example, temptation episode and random observations typically fell within a day prior to the lapse, making it unlikely that the observed elevation of negative affect in lapse episodes could be explained by nicotine withdrawal (which changes slowly and decreases over time). Thus, EMA assessments allow the lapse episode to be understood in context.

Relapse investigators have been particularly interested in the *initial* lapse to drug use, as it represents a pivotal transition from abstinence back to use. This imposes particular challenges, because the initial lapse episode is a unique event – there is no second first lapse. As exemplified

in these studies, the strategy in EMA studies is to engage subjects in ongoing monitoring, so that they are poised to record the first lapse if and when it does occur. This process is by no means fool-proof: for example, in Shiffman et al. (1996) 16% of subjects failed to record their first lapse in real time, and thus could not be used in analyses of first lapses. Nevertheless, EMA data provide a method for collecting timely information about the unique circumstances in which subjects' first lapse.

"Pulling back the camera" - Lead-up to lapses and relapse

These episode-level analyses train close-up lens on initial lapse episodes, with informative results. However, their very focus is also limiting, in that it attends only to the most immediate circumstances of the lapse. Other EMA studies have "pulled back the camera" to examine broader time-frames, both leading up to the initial lapse, and following forward from it as users progress towards relapse.

An example of broader examination of antecedents of the first lapse is given by Shiffman and Waters (2004), who focused on mood in the days and hours preceding the initial lapse. Shiffman et al. (1996) had shown that smokers reported excess negative affect during the initial lapse, but it was not known whether affective distress actually preceded the episode and thus precipitated it. Further, the data on affect during the lapse were potentially subject to recall bias, since the episode was recorded after the smoking had already occurred, so subjects could confound post-lapse affect with their report if pre-lapse affect. Accordingly Shiffman and Waters examined affect preceding the initial lapse, in analyses that illustrate the use of different time-frames for aggregation and analysis of EMA data. First, Shiffman and Waters analyzed data aggregated at the day level, estimating affect in the days preceding the lapse day (from 5–6 randomly-scheduled momentary prompts), and also examining once-a-day stress ratings. The risk of lapsing on a given day was, in fact, not related to the prior day's average negative affect or stress. However, a more fine-grained and more proximal analysis examined individual momentary ratings of affect on the same day, in the hours preceding the lapse, to determine the trajectory of mood preceding the lapse episode. That analysis showed steadily increasing negative affect in the 6–7 hours preceding stress-related lapse episodes. In a variation of this approach, Cooney et al. (2007) compared the assessment immediately preceding a lapse to prior assessments on the same day, and similarly concluded that negative mood foreshadowed first lapses to smoking. Substantively, these findings illustrate that events like lapses are often affected by very proximal influences, rather than distal ones. Methodologically, they demonstrate the importance of attention to the time frames used in analyses of EMA data. The findings were also methodologically telling in that they validated from prospective data the momentary ratings made right after the lapse episode itself.

The finding that daily stress does not predict lapse, but more proximal stress does, is also particularly interesting in contrast to the findings of Hall, Havassy, and Wasserman (1990), who undertook an analysis of the same question using non-EMA retrospective data. Hall et al. collected mood and stress data weekly, and reported increased negative mood and stress on the week of a lapse (retrospectively), but no prospective association. Hall et al. concluded that stress was unrelated to lapsing (based on the prospective data), and that the retrospective data simply reflected recall bias. However, in light of Shiffman & Waters' (2004) finding that negative affect and stress do not exercise prospective influence on lapse risk even over days (much less weeks), but only over hours, the Hall study may exemplify Collins and Graham's (2002) caution that when assessment intervals are too widely spaced, important relationships can be missed. Collins and Graham note that cause-effect relationships tend to decay over time, and that more frequent assessment allows one to capture such "immediate" associations, while also allowing for assessment of associations at longer lags if those prove important.

While these analyses focus on individual initial lapse episodes, McCarthy et al. (2006) presented a broader view of smokers' experiences during the quit process. They had smokers record craving, negative affect, and withdrawal 4 times a day for three weeks before and three weeks after quitting. Analyses treated the series as a set of piece-wise trajectories, with an inflection on the quit date. All three measures increased on the quit date, and, surprisingly, did not decrease significantly over the next three weeks post-quit. An interesting aspect of McCarthy's analysis is that it looked at between-person variability as well as average trends, and found that between-subject variability increased after the quit date, both in the level of symptoms and in how they evolved over time. This suggests that quitting unmasks some individual differences not seen during ongoing smoking.

Just as EMA data have been used to track processes over time preceding a lapse, they have also been used to track processes following a lapse, to understand how subjects subsequently progress towards relapse. Some EMA studies have used records of subsequent lapses simply as end-points to be predicted from independent variables: For example, Shiffman, Ferguson & Gwaltney (2006) showed that how pleasant smokers found an initial lapse affected how quickly smokers progressed to second lapse or to relapse, which were assessed via EMA event records. ("Relapse" was defined as smoking at least 5 cigarettes a day for three consecutive days, and was computed by the PDA from event-based entries of individual smoking episodes.)

Other studies have used ongoing monitoring to examine the trajectory of smoking and psychological experience that leads to relapse. For example, Gwaltney, Shiffman, and Sayette (2005) used data on abstinence self-efficacy (ASE), collected 4–5 times a day via randomly-scheduled assessments, along with smoking data, collected by event recording, to track the interplay between ASE and smoking in progression from an initial lapse to relapse. Averaged daily ASE on a given day predicted lapsing the following day, and progression towards relapse, even after accounting for smoking on the prior day. This analysis did not examine individual lapse episodes at all, but used the day as the aggregate unit of analysis for both ASE and smoking.

A contrasting approach is illustrated by Kirchner, Shiffman, and Cheong (2007), who also examined ASE in a nearly-identical design and, but, conversely, did not use the time-sampled data at all, and did not aggregate the data by day, but rather focused exclusively on data collected at lapse events themselves. Kirchner et al.'s analysis conceptualized each lapse episode as the beginning of a period of risk for a further subsequent lapse, and examined whether ASE reported after the index lapse predicted the risk/rate of progression to a subsequent lapse (as hypothesized by Marlatt & Gordon, 1985). It did. The effect was stronger when the analysis also considered the trajectory of ASE coming into the index lapse: Independent of its level, if ASE decreased from the previous lapse to the current one, the risk of progression to another lapse increased. In a different vein, but again illustrating different ways to organize and analyze EMA data, O'Connell, Schwartz, and Shiffman (2008) computed the number and duration of temptation episodes (from event-based entries) over time between successive lapses, and found, contrary to hypothesis, that experiencing more temptation episodes was associated with decreased lapse risk. The range of analyses recounted here illustrates the variety of the ways in which a longitudinal series of event- and time-based EMA observations can be used to answer theory-based questions.

Alcohol

There have been surprisingly few studies using EMA to track the process of relapse to drinking. Litt et al. (1998) performed an early study using programmable wristwatches to prompt post-treatment alcoholic subjects for assessments. Most of the findings were methodological, and are discussed elsewhere. A subsequent study by Cooney et al. (2007) used electronic diaries

with alcoholics who were also treated for smoking cessation, and who were assessed four times daily after being released from an inpatient program. Cooney et al. focused the analysis prospectively on the proximal antecedents of alcohol lapses, by comparing the assessment just before the lapse to other assessments that same day. The findings were surprising in that alcohol lapses were predicted by urge to smoke, but not urge to drink. Less surprisingly, lower ASE in the immediately-preceding observation predicted the occurrence of a drinking lapse. This is important, because it demonstrates that there are meaningful changes in self-efficacy – a variable that is often measured just once and assumed to be stable – over the space of minutes or hours. (See also Gwaltney et al. 2005, who demonstrated that ASE can vary by situational context.) Thus, the study illustrates the utility of EMA for demonstrating change where none was expected, and showing the effects of such acute changes on meaningful outcomes such as lapses to drinking in abstinent alcoholics.

Cooney et al.'s (2007) analysis demonstrates the potential utility of EMA methods for understanding the relapse process in alcoholism, making it seem all the more surprising that EMA is not more often used to study alcohol relapse. As with illicit drug use, the assumed or actual willingness and ability of alcoholics or alcohol abusers to perform EMA may be a barrier. Cooney et al. also comment on how few subjects in their treated sample lapsed within the two weeks they were under EMA observation: Cooney et al. followed 102 drinkers, but only had data on 13 first lapses. Longer follow-ups with different samples may make observation of lapses more common.

Cooney et al. (2007), and other investigators studying drinkers who are trying to stop (Litt et al., 1998) have also commented on the low levels of craving reported in EMA; e.g., Litt et al. report that alcoholics recorded *any* craving in only 8% of time-based assessments. It is not yet clear whether this is a real finding, which suggests that our models of alcohol relapse need to be revised, or whether it might represent a methodological limitation, i.e., that treated alcoholics are reluctant to "admit" to craving. The latter seems less plausible because the same subjects report moderate craving in the laboratory (Litt et al.). Further exploration of this issue, and of EMA approaches to alcohol relapse, is warranted.

Even more so than the studies of ongoing use, analyses of the relapse process illustrate different organizations of the data and different forms of statistical analysis that can be applied to a rich EMA data stream to answer a spectrum of different research questions. Some analyses focused on particular episodes or moments and contrasted them to others. Some have looked at the trajectory of psychological processes over time, either anchored at a fixed, predictable point in time (such as quitting or end of treatment) or anchored after-the-fact by initially unpredictable events such as lapses. Still others have used time-to-event (event history, survival) analyses to examine how one event or process shapes the risk of a second event, over time. The role of time in the relapse process, and its representation in EMA analyses, is particularly important, as relapse represents a progression of events over time.

EMA as an outcome measure

Given EMA's unique ability to support analyses of process, it is not surprising that most of the studies reviewed have been observational studies with a process focus. However, EMA methods have also been used to assess outcome in randomized clinical trials. Usually, event-based entries of drug use are used to compute some pre-defined outcome measure. For example, Tidey et al. (2008) used electronic diary EMA reports of drinking events to assess drinking in heavy drinkers treated with naltrexone or placebo: Naltrexone resulted in significantly fewer drinking days, though it had no effect on how much drinking took place on those days. In a randomized trial of combinations of bupropion and counseling, McCarthy et al. (2008) used EMA records of individual smoking occasions to assess abstinence; bupropion increased

abstinence, but counseling had no effect. EMA has also been used to examine the effects of treatment on symptoms, such as craving and withdrawal, which are often regarded as outcomes in themselves. For example, Shiffman et al. (2000) used data collected during 12 random prompts daily over three weeks to demonstrate that a 24-hour patch for smoking cessation yielded lower craving than a 16-hour patch. This study was unusual in that the random prompting schedule was weighted towards assessments in the first few hours of the day, when the differences between the patch formulations were expected to be greatest. This illustrates that sampling schemes need not be based on simple random sampling, but can over- or under-sample when the hypothesis warrants.

As a contrast, in some studies, EMA data are not used to assess the outcome – or even drug use- at all, but are instead used to assess predictors of drug use or outcome. For example, Weinstein, Mermelstein, Shiffman, and Flay (2008) used time-based EMA assessments to assess the variability in teen smokers' moods at baseline, and then demonstrated that the teens with the most volatile mood showed accelerated progression towards heavy smoking, based on a non-EMA longitudinal measure of cigarette consumption. In a similar vein, McCarthy et al. (2006) used EMA data about smokers' changes in negative affect and craving leading up to and on a quit date to predict (non-EMA) abstinence status three months later. Thus, EMA analyses focused on drug use need not include EMA measures of drug use.

Perhaps a more important use of EMA data has been to assess and analyze putative processes that are hypothesized to mediate the effects of treatment. For example, in a randomized trial of bupropion vs placebo for smoking cessation, McCarthy et al. (2008) used data collected during 3 random assessments per day to estimate both the level of craving when subjects quit and also the rate at which craving declined over days. Bupropion had been hypothesized to help smokers maintain abstinence by reducing craving. McCarthy et al. report that bupropion did not affect the initial level of post-quit craving, but did increase its rate of decline, and that this effect partially mediated the effect if the drug in enhancing abstinence. In a more complex model, Kranzler, Armeli, Feinn, and Tennen (2004), building on findings suggesting that positive and negative affect promote drinking, examined the effect of naltrexone on the association between affect and heavy drinking in a group of heavy drinkers. Using daily diaries, they found that naltrexone buffered drinkers from the effects of both positive and negative affect: for subjects on placebo both positive and negative affect were associated with heavy drinking; for subjects on naltrexone, affect did not influence heavy drinking. A later analysis (Armeli, Feinn, Tennen, & Kranzler, 2006) concentrating specifically on positive or negative social interactions produced more complex (and somewhat contradictory) results (e.g., subjects on naltrexone were *more* likely to drink and to drink heavily on days with celebratory social occasions, perhaps to overcome naltrexone's blunting effect on alcohol reinforcement). In any case, these analyses illustrate how EMA data can be used not only to evaluate *whether* a treatment works, but also *how* it works. Collins and Graham (2002) have emphasized the importance of longitudinal data for evaluating mediational processes. Thus, EMA methods have considerable utility in drug treatment trials, not only for assessing outcome, but for assessing processes that may mediate or moderate treatment effects.

Methodological analyses

Compliance

EMA methods obviously depend on subjects' compliance with instructions to respond to prompts, record episodes of drug use, and so on. Non-compliance cannot only lead to missing data, but introduce bias in the data that are collected. Compliance with prompted assessments can be assessed, because we know when assessments were to be entered. In studies using paper diaries, compliance cannot really be assessed, because it is not known when assessments were completed, and studies have demonstrated that subjects frequently fake compliance by

completing diaries after the fact (Stone, Shiffman, Schwartz, Broderick, & Hufford, 2002; Broderick, Schwartz, Shiffman, Hufford, & Stone, 2003). Litt et al. (1998) reported this specifically in a study of alcohol treatment.

In contrast, studies using electronic diaries can confirm compliance, because the timing of the entry can be confirmed. Compliance has varied widely across studies, ranging from over 90% in some studies (Litt et al., 2000; Shiffman et al., 2002; Todd et al., in press), to 75%–80% in many studies (e.g., Carter et al., 2008); Epstein et al. (2009)), to as low as 50% (Litt et al., 1998; Otsuki, Tinsley, Chao, & Unger, 2008; Shapiro et al., 2002). It is not clear what causes this variation. Beckham et al. (2008) report that compliance improved when an incentive for completing assessments was introduced, suggesting that subject-management procedures such as incentives (and likely also training, feedback, etc.) can help boost compliance.

Assessing compliance with recording of events (e.g., drug use) is much more challenging, because we usually have no way of knowing how many events actually occurred. Global self-reports of drug use are suspect as representations of the truth. But using these as a rough benchmark, the proportion of cigarettes estimated to have been entered ranges from 22% (Rowan et al., 2007), to 50% (Delfino et al., 2001; Shapiro et al., 2002) and up to 90% (Shiffman, in press). Available evidence, which is scant, suggests that compliance is also variable when subjects are asked to record lapses. For example, Litt et al. (2000) report that recovering alcoholics recorded much more drinking on EMA than on a later retrospective assessment, suggesting good compliance. However, Litt et al. (1998) report that recovering alcoholics only recorded a minority of their lapses, and that alcoholics sometimes suspended recording for a few days after a lapse, though this was observed in less than a third of alcohol subjects (Litt et al., 2000). Thus compliance is variable, but good subject management procedures can yield high compliance.

Comparisons with biochemical markers

Although we lack a true Gold Standard, drug use, more than other behaviors studied with EMA, may lend itself to objective verification, because it leaves biological traces. In Freedman et al.'s (2006) study of homeless crack addicts, only 7% of participants showed positive urine tests while failing to report cocaine use in their EMA data, suggesting reasonable compliance, but not verifying the quantity of drugs used. Smoking is associated with more quantitative biochemical markers. For example, smoking elevates carbon monoxide (CO), which can be assessed in the breath. A recent study (Shiffman, in press) found that EMA records of cigarette consumption correlated reasonably well with observed levels of CO monitored at clinic visits, and that changes in CO across multiple measurement occasions correlated well with corresponding changes in EMA-recorded cigarettes. CO data were also used to verify that episodes of use are being recorded in a timely way. In contrast, Toll, Cooney, McKee, and O'Malley (2006) report poor compliance with recording of lapses after treatment: daily IVRS reports of claimed abstinence were contradicted by CO measures 9% of the time.

In the case of alcohol, blood alcohol concentrations (BAC) provide an imperfect indicator of alcohol consumption. Perrine, Mundt, Searles, and Lester (1995) report that daily IVRS reports of drinks consumed correlated 0.72 with observed evening BAC. Within-subject correlations across days were generally high (average r=0.61), but varied considerably across subjects (−0.07 –0.92). All currently available biochemical markers of substance use are imprecise, but may still be of some use for validating EMA, until improved markers are developed.

Comparisons with Time-Line Follow-Back measures

Another common comparator for EMA or daily-diary estimates of drug use has been retrospective self-reports collected using the Time-Line Follow-Back (TLFB) method (Sobell,

1995), in which subjects are asked to recall their drug use for a past period (usually 7 to 30 days). One detailed analysis of TLFB and EMA records of cigarette smoking (Shiffman, in press) noted that TLFB records were subject to high levels of digit bias – the tendency of people to report quantities in round numbers – demonstrating rounding to units of 10 on almost half the days, rendering the TLFB data suspect. On average, subjects recorded 11% fewer cigarettes (2.6) than they later recalled on TLFB, but on about one third of days, subjects had entered more cigarettes than they later recalled smoking, again suggesting that the TLFB data might not be valid. When the data were aggregated across days to generate subject-level estimates, the association between EMA and TLFB was quite strong, β=0.77 (based on multiple regression, controlling for subjects). However, the within-subject association across days was modest, β=0.29: the patterns of variation in cigarette consumption over days did not match very well. Studies examining correspondence between daily IVRS and EMA data on drinking similarly find strong between-subject correlations, but weaker, and markedly variable, correspondence at the day level (Carney, Tennen, Affleck, Del Boca, & Kranzler, 1998; Perrine et al., 1995; Toll, Cooney, McKee, & O'Malley, 2005; Toll et al., 2006; Hufford et al., 2002; Tucker, Foushee, Black, & Roth, 2007). This suggests that when estimating patterns or changes in consumption, TLFB may not be adequate and EMA may be necessary.

Reactivity

Reactivity – the possibility that the research methods themselves affect the behavior under study, and thus distort the findings – is a concern for EMA studies of drug abuse. EMA methods seem particularly vulnerable to reactivity, because the assessments are completed repeatedly and in close proximity to the behavior of interest, putting them in a prime position to affect behavior. Reactivity is also enhanced when subjects are asked to record undesirable target events before they are completed (Rozensky, 1974), probably because this gives subjects a chance to reconsider the behavior. (This suggests that event recording could be scheduled to occur after an episode to minimize reactivity; see e.g., Shiffman, Paty, et al., 1996). Drug use may also be particularly vulnerable to reactivity because it is typically considered undesirable, and reactivity increases when subjects are motivated to change their behavior (McFall & Stuart, 1977). Thus, EMA assessment of drug use seems to have great potential for reactivity.

However, thus far, studies assessing reactivity of drug-use recording have not demonstrated strong reactivity (Hufford et al., 2002). A study of smokers preparing to quit smoking (Shiffman et al., 2002) – which should maximize reactivity – found only a modest decrease in cigarette consumption (a reduction of 0.3 cigarettes per day), and no change on CO levels, suggesting any changes in consumption were modest. Rowan et al. (2007) reported some changes in smoking-related cognitions due to monitoring, but their monitoring was accompanied by an intervention. Collins et al. (1998) report decreases in drinking among treatment-seeking heavy drinkers, while neither Hufford et al. nor Carney et al. (1998) report changes in drinking. In a study with a no-monitoring control group, which provides the clearest interpretation, Simpson, Kivlahan, Bush, and McFall (2005) reported no effect on drinking, despite subjects' impressions that monitoring affected their drinking. The absence of robust changes in drug use due to monitoring seems surprising. Moos (2008) speculates that the very intensity of EMA assessment may blunt reactivity, by causing habituation. Thus, although reactivity remains a concern, the literature to date has not indicated strong reactivity effects.

Validation of EMA data

Comparison against questionnaires

While data gathered by tracking each episode of drug use seem a valid way to assess individual differences in the circumstances of drug use, and to infer motives for drug use, the question arises whether these patterns might also be captured by questionnaires, which are far easier to

administer. Data on smoking patterns suggest that questionnaires do not adequately capture variations in smoking settings and moments. Among heavy smokers in a quit clinic (Shiffman, 1993), correlations between EMA-assessed patterns and questionnaire-reported patterns were near zero. Similarly, in a study of light-smoking Asian college students Otsuki et al. (2008) found no correlation between questionnaire-assessed social motives for smoking and EMA-observed social smoking. Coping motives were similarly unrelated to smoking settings.

Piasecki et al. (2007) also report a study of college-student smoking, but with a different strategy, in which smokers who had completed a questionnaire of smoking motives also explicitly recorded their motives for smoking each cigarette, rather than just reporting the circumstances of their smoking (and inferring a motive). Associations were generally weak or absent. Notably, questionnaire reports of smoking to reduce negative emotions – which are often cited as a key motive for smoking – were unrelated to real-time reports of smoking for this reason. Indeed, whereas smoking to reduce negative affect is almost always cited as the leading motive for smoking in questionnaire measures (Shiffman, 1993), it was the *least* often cited motive when motives were tracked in near real time. Questionnaires do not validly capture smoking patterns or motives as assessed by EMA.

Much of the literature on drinking has focused on a similar set of hypotheses about individual differences in drinking motives, focusing on drinkers who score high on Drinking to Cope (DTC; Cooper, Russell, Skinner, Frone, & Mudar, 1992) – i.e., who report drinking in order to reduce affective distress. Many studies have used EMA data or daily diaries to test whether smokers who report DTC motives demonstrate a stronger link between stress and drinking. The results have been mixed and complex. For example, Armeli et al. (2008) found that for subjects with high DTC, accumulating anxiety during the week led to earlier onset of drinking, whereas among those low in DTC, lower anxiety led to earlier drinking onset. However, the relationship was not consistent: anger showed reverse effects. In another study, Todd et al. (in press) found that DTC did not moderate the relationship between mood and the onset of drinking throughout the week, but only during the work-week, and only for certain low-arousal negative moods. In other studies, there was no relationship between stress and drinking among high-DTC subjects, though stress may reduce drinking among low DTC subjects (Todd, Armeli, Tennen, Carney, & Affleck, 2003). In sum, the results suggest that DTC questionnaires do tap some individual differences in drinking motives, but that the relationships, as revealed by EMA, appear to be far more complex than is captured in the DTC questionnaires.

In studies of relapse, investigators have tried to capture the details of initial lapse situations. Typically, this had been done by debriefing subjects months later (e.g., Marlatt & Gordon, 1985; Shiffman et al., 1997). Shiffman et al. evaluated the accuracy of such retrospective recall by comparing data gathered that way to EMA data recorded in real time by subjects using a PDA to record lapses. Recall accuracy was quite poor, with κ statistics for agreement in the . 20s. Moreover, the retrospective reports appeared to be influenced by subjects' smoking status at the time of the retrospective report, which could introduce substantial bias. Finally, Shiffman et al. suggested that much of what passed for "recall" might actually be based on general declarative knowledge about typical lapse situations. (Because people who had never smoked could characterize the typical lapse situation about as well as the actual smokers could do retrospectively.)

Not only does recall of lapses appear to be inaccurate, it also appears to yield biased results. Whereas retrospective data made it appear that how smokers felt about their first lapse influenced subsequent relapse (Curry, Marlatt, & Gordon, 1987) – probably because subsequent smoking influenced recall – prospective analyses from real-time data found little support for this notion (Shiffman, Hickcox et al., 1996). This illustrates how EMA data can provide new insights that go beyond those from recall methods.

Predictive and construct validity

Relatively few studies have examined the predictive validity of EMA data in drug use. One recent analysis showed that individual differences in patterns of smoking over the day, as assessed by EMA, predicted outcomes in subsequent smoking cessation (Chandra et al., 2007). (This also helps validate the timely recording of smoking events.) Another analysis from the same dataset (Shiffman et al., 2007) directly compared the predictive power of EMA and questionnaire data on negative affect smoking. Only the EMA-based measure (the individual subject's correlation between mood and smoking) predicted individual differences in relapse risk; the questionnaires did not. Additional comparisons of predictive and construct validity of questionnaire and EMA methods would be useful.

Conclusions

Much work remains to be done to extend the EMA study of drug use. As already discussed, EMA methods need to be applied to the full range of drug use and drug users. Methodological analyses of reactivity and compliance, particularly with event-based records of drug use, will help improve our interpretation of EMA data. Also, although assessment and analysis of momentary states is the forte of EMA methods, research that also incorporates assessment of slower moving "background" processes (Shiffman, 1989), such as commitment to behavior change, access to social support, or changes in psychiatric status is likely to extend the theoretical reach of EMA data. The contribution of EMA data is also likely to be enhanced by technological developments, such as real-time biochemical measures of drug intake, integrated measures of physical and physiological parameters, automated measures of environmental exposures, social network measures, and GPS-based geographical information. Perhaps most important of all will be theoretical developments that incorporate a key emergent implication of EMA data – that cognition, affect, and behavior are dynamic, changing over relatively short time periods, and are highly responsive to environmental stimuli and cues. This deeper level of databased insight will drive new thinking, in the same way the invention of the microscope revealed new worlds and required new thinking.

EMA methods have helped illuminate on the processes that drive drug use, cessation, and relapse, and have the promise to push our understanding further in the future. By shedding light on how drug use plays out across time and across contexts, EMA has helped us understand drug use and relapse as never before. The detailed analyses we have seen of mood variations and their relation to drug use, and of the unfolding of relapse processes over minutes, hours, and days would simply not be possible without the rich stream of repeated assessments afforded by EMA methods. There is also great potential in extending EMA from assessment to clinical intervention. Even with current technology, EMA seems to hold great potential to bring both assessment and intervention into the drug user's daily life, in the moment. EMA methods will yield new ways to better understand and change drug use and drug users.

Acknowledgments

This work was supported by National Institutes of Health, National Institute on Drug Abuse grant # DA02074. Thanks are due to Chantele Mitchell-Miland and Sarah Scholl for help with preparation of the manuscript. The author is a co-founder of invivodata, inc., which provides electronic diary services for research.

References

Armeli S, Feinn R, Tennen H, Kranzler HR. The effects of naltrexone on alcohol consumption and affect reactivity to daily interpersonal events among heavy drinkers. Experimental and Clinical Psychopharmacology 2006;14(2):199–208. [PubMed: 16756424]

Armeli S, Todd M, Conner TS, Tennen H. Drinking to cope with negative moods and the immediacy of drinking within the weekly cycle among college students. Journal of Studies on Alcohol and Drugs 2008;69(2):313–322. [PubMed: 18299774]

Baker TB, Piper ME, McCarthy DE, Majeskie MR, Fiore MC. Addiction motivation reformulated: An affective processing model of negative reinforcement. Psychological Review 2004;111(1):33–51. [PubMed: 14756584]

Beckham JC, Wiley MT, Miller SC, Dennis MF, Wilson SM, McClernon FJ, et al. Ad lib smoking in post-traumatic stress disorder: An electronic diary study. Nicotine & Tobacco Research 2008;10(7): 1149–1157. [PubMed: 18629724]

Broderick JE, Schwartz JE, Shiffman S, Hufford MR, Stone AA. Signaling does not adequately improve diary compliance. Annals of Behavioral Medicine 2003;26(2):139–148. [PubMed: 14534031]

Carney MA, Tennen H, Affleck G, Del Boca FK, Kranzler HR. Levels and patterns of alcohol consumption using timeline follow-back, daily diaries and real-time “electronic interviews”. Journal of Studies on Alcohol 1998;59(4):447–454. [PubMed: 9647427]

Carter BL, Lam CY, Robinson JD, Paris MM, Waters AJ, Wetter DW, et al. Real-time craving and mood assessments before and after smoking. Nicotine & Tobacco Research 2008;10(7):1165–1169. [PubMed: 18629726]

Chandra S, Shiffman S, Scharf DM, Dang Q, Shadel WG. Daily smoking patterns, their determinants, and implications for quitting. Experimental and Clinical Psychopharmacology 2007;15:67–80. [PubMed: 17295586]

Collins LM, Graham JW. The effect of the timing and spacing of observations in longitudinal studies of tobacco and other drug use: Temporal design considerations. Drug Alcohol Dependence 2002;68 (Suppl 1):85–96.

Collins RL, Morsheimer ET, Shiffman S, Paty JA, Gnys M, Papandonatos GD. Ecological momentary assessment in a behavioral drinking moderation training program. Experimental and Clinical Psychopharmacology 1998;6:306–315. [PubMed: 9725114]

Cooney NL, Litt MD, Cooney JL, Pilkey DT, Steinberg HR, Oncken CA. Alcohol and tobacco cessation in alcohol-dependent smokers: Analysis of real-time reports. Psychology of Addictive Behaviors 2007;21(3):277–286. [PubMed: 17874878]

Cooper ML, Agocha VB, Sheldon MS. A motiv ational perspective on risky behaviors: The role of personality and affect regulatory processes. Journal of Personality 2000;68(6):1059–1088. [PubMed: 11130732]

Cooper ML, Russell M, Skinner JB, Frone MR, Mudar P. Stress and alcohol use: Moderating effects of gender, coping, and alcohol expectancies. Journal of Abnormal Psychology 1992;101(1):139–152. [PubMed: 1537960]

Curry S, Marlatt GA, Gordon JR. Abstinence violation effect: Validation of an attributional construct with smoking cessation. Journal of Consulting and Clinical Psychology 1987;55(2):145–149. [PubMed: 3571666]

Delfino RJ, Jamner LD, Whalen CK. Temporal analysis of the relationship of smoking behavior and urges to mood states in men versus women. Nicotine & Tobacco Research 2001;3(3):235–248. [PubMed: 11506767]

Epstein DH, Willner-Reid J, Vahabzadeh M, Mezghanni M, Lin JL, Preston KL. Real-time electronic diary reports of cue exposure and mood in the hours before cocaine and heroin craving and use. Archives of General Psychiatry 2009;66:88–94. [PubMed: 19124692]

Freedman M, Lester K, McNamara C, Milby J, Schumacher J. Cell phones for ecological momentary assessment with cocaine-addicted homeless patients in treatment. Journal of Substance Abuse Treatment 2006;30:105–111. [PubMed: 16490673]

Gwaltney CJ, Shiffman S, Sayette MA. Situational correlates of abstinence self-efficacy. Journal of Abnormal Psychology 2005;114(4):649–660. [PubMed: 16351386]

Hall SM, Havassy BE, Wasserman DA. Commitment to abstinence and acute stress in relapse to alcohol, opiates, and nicotine. Journal of Consulting and Clinical Psychology 1990;58(2):175–181. [PubMed: 2335634]

Hopper JW, Su Z, Looby AR, Ryan ET, Penetar DM, Palmer CM, et al. Incidence and patterns of polydrug use and craving for ecstasy in regular ecstasy users: An ecological momentary assessment study. Drug and Alcohol Dependence 2006;85(3):221–235. [PubMed: 16730923]

Hufford MR, Shields AL, Shiffman S, Paty J, Balabanis M. Reactivity to ecological momentary assessment: An example using undergraduate problem drinkers. Psychology of Addictive Behaviors 2002;16(3):205–211. [PubMed: 12236455]

Hussong AM, Hicks RE, Levy SA, Curran PJ. Specifying the relations between affect and heavy alcohol use among young adults. Journal of Abnormal Psychology 2001;110(3):449–461. [PubMed: 11502088]

Johnson EI, Barrault M, Nadeau L, Swendsen J. Feasibility and validity of computerized ambulatory monitoring in drug-dependent women? Drug and Alcohol Dependence. (in press).

Kirchner, TR.; Shiffman, S.; Cheong, J. Modeling the abstinence-violation effect during smoking cessation. Paper presented at the Annual Meeting of the Society for Research on Nicotine and Tobacco; Austin, TX. 2007 Feb.

Kranzler H, Armeli S, Feinn R, Tennen H. Targeted naltrexone treatment moderates the relations between mood and drinking behavior in early problem drinkers. Journal of Consulting and Clinical Psychology 2004;72(2):317–327. [PubMed: 15065964]

Litt MD, Cooney NL, Morse P. Ecological momentary assessment (EMA) with treated alcoholics: Methodological problems and potential solutions. Health Psychology 1998;17(1):48–52. [PubMed: 9459069]

Litt MD, Cooney NL, Morse P. Reactivity to alcohol-related stimuli in the laboratory and in the field: Predictors of craving in treated alcoholics. Addiction 2000;95(6):889–900. [PubMed: 10946438]

Maclure M, Mittleman MA. Should we use a case-crossover design? Annual Review of Public Health 2000;21(1):193–221.

Marlatt, GA.; Gordon, JR. Relapse prevention. New York: Guilford Press; 1985.

McCarthy DE, Piasecki TM, Fiore MC, Baker TB. Life before and after quitting smoking: An electronic diary study. Journal of Abnormal Psychology 2006;115(3):454–466. [PubMed: 16866586]

McCarthy DE, Piasecki TM, Lawrence DL, Jorenby DE, Shiffman S, Baker TB. Psychological mediators of bupropion sustained-release treatment for smoking cessation. Addiction 2008;103(9):1521–1533. [PubMed: 18783504]

McFall, RM.; Stuart, RB. Behavioral self-management: Strategies, techniques, and outcome. New York: Brunner/Mazel; 1977. Parameters of self-monitoring; p. 196-214.

McKay JR, Franklin TR, Patapis N, Lynch KG. Conceptual, methodological, and analytical issues in the study of relapse. Clinical Psychology Review 2006;26(2):109–127. [PubMed: 16371242]

Moghaddam NG, Ferguson E. Smoking, mood regulation, and personality: An event-sampling exploration of potential models and moderation. Journal of Personality 2007;75(3):451–478. [PubMed: 17489888]

Moos RH. Context and mechanisms of reactivity to assessment and treatment. Addiction 2008;103(2): 249–250. [PubMed: 18199303]

Muraven M, Collins RL, Morshimer ET, Shiffman S, Paty JA. The morning after: Limit violations and the self-regulation of alcohol consumption. Psychology of Addictive Behaviors 2005;19(3):253–262. [PubMed: 16187803]

O'Connell KA, Gerkovich MM, Cook MR, Shiffman S, Hickcox M, Kakolewski KE. Coping in real time: Using ecological momentary assessment techniques to assess coping with the urge to smoke. Research in Nursing & Health 1998;21(6):487–497. [PubMed: 9839794]

O'Connell KA, Schwartz JE, Shiffman S. Do resisted temptations during smoking cessation deplete or augment self-control resources? Psychology of Addictive Behaviors 2008;22(4):486–495. [PubMed: 19071973]

Otsuki M, Tinsley BJ, Chao RK, Unger JB. An ecological perspective on smoking among Asian American college students: The roles of social smoking and smoking motives. Psychology of Addictive Behaviors 2008;22(4):514–523. [PubMed: 19071976]

Paty, JA.; Kassel, JD.; Shiffman, S. Assessing stimulus control of smoking: The importance of base rates. In: deVries, MW., editor. The Experience of Psychopathology. Cambridge, England: Cambridge University Press; 1992. p. 347-353.

Perrine MW, Mundt JC, Searles JS, Lester LS. Validation of daily self-reported alcohol consumption using interactive voice response (IVR) technology. Journal of Studies on Alcohol 1995;56(5):487–490. [PubMed: 7475027]

Piasecki TM, Richardson AE, Smith SM. Self-monitored motives for smoking among college students. Psychology of Addictive Behaviors 2007;21(3):328–337. [PubMed: 17874883]

Rowan PJ, Cofta-Woerpel L, Mazas CA, Vidrine JI, Reitzel LR, Cinciripini PM, et al. Evaluating reactivity to ecological momentary assessment during smoking cessation. Experimental and Clinical Psychopharmacology 2007;15(4):382–389. [PubMed: 17696685]

Rozensky RH. The effect of timing of self-monitoring behavior on reducing cigarette consumption. Journal of Behavior Therapy and Experimental Psychiatry 1974;5:301–303.

Shapiro D, Jamner LD, Davtdov DM, James W. Situations and moods associated with smoking in everyday life. Psychology of Addictive Behavior 2002;16(4):342–345.

Shiffman S. Relapse following smoking cessation: A situational analysis. Journal of Consulting and Clinical Psychology 1982;50(1):71–86. [PubMed: 7056922]

Shiffman, S. Conceptual issues in the study of relapse. In: Gossop, M., editor. Relapse and addictive behaviour. Kent, England: Routledge; 1989. p. 149-179.

Shiffman S. Assessing smoking patterns and motives. Journal of Consulting and Clinical Psychology 1993;61(5):732–742. [PubMed: 8245271]

Shiffman, S. Designing protocols for ecological momentary assessment. In: Stone, AA.; Shiffman, S.; Atienza, A.; Nebeling, L., editors. The Science of real-time data capture: Self-reports in health research. New York: Oxford University Press; 2007. p. 27-53.

Shiffman S. How many cigarettes did you smoke? Assessing cigarette consumption by global report, time-line follow-back, and ecological momentary assessment? Health Psychology. (In Press).

Shiffman S, Balabanis MH, Gwaltney CJ, Paty JA, Gnys M, Kassel JD, et al. Prediction of lapse from associations between smoking and situational antecedents assessed by ecological momentary assessment. Drug and Alcohol Dependence 2007;91:159–168. [PubMed: 17628353]

Shiffman S, Elash CA, Paton SM, Gwaltney CJ, Paty JA, Clark DB, et al. Comparative efficacy of 24-hour and 16-hour transdermal nicotine patches for relief of morning craving. Addiction 2000;95(8): 1185–1195. [PubMed: 11092066]

Shiffman S, Ferguson SG, Gwaltney CJ. Immediate hedonic response to smoking lapses: Relationship to smoking relapse, and effects of nicotine replacement therapy. Psychopharmacology 2006;184:608–614. [PubMed: 16283258]

Shiffman S, Gwaltney CJ, Balabanis M, Liu KS, Paty JA, Kassel JD, et al. Immediate antecedents of cigarette smoking: An analysis from ecological momentary assessment. Journal of Abnormal Psychology 2002;111(4):531–545. [PubMed: 12428767]

Shiffman S, Hickcox M, Paty JA, Gnys M, Kassel JD, Richards T. Progression from a smoking lapse to relapse: Prediction from abstinence violation effects, nicotine dependence, and lapse characteristics. Journal of Consulting and Clinical Psychology 1996;64(5):993–1002. [PubMed: 8916628]

Shiffman S, Hufford M, Hickcox M, Paty JA, Gnys M, Kassel JD. Remember that? A comparison of real-time versus retrospective recall of smoking lapses. Journal of Consulting and Clinical Psychology 1997;65(2):292–300. [PubMed: 9086693]

Shiffman S, Paty JA, Gnys M, Kassel JD, Hickcox M. First lapses to smoking: Within-subjects analyses of real-time reports. Journal of Consulting and Clinical Psychology 1996;64(2):366–379. [PubMed: 8871421]

Shiffman S, Paty JA, Gwaltney CJ, Dang Q. Immediate antecedents of cigarette smoking: An analysis of unrestricted smoking patterns. Journal of Abnormal Psychology 2004;113(1):166–171. [PubMed: 14992670]

Shiffman S, Stone AA, Hufford M. Ecological momentary assessment. Annual Review of Clinical Psychology 2008;4:1–32.

Shiffman S, Waters AJ. Negative affect and smoking lapses: A prospective analysis. Journal of Consulting and Clinical Psychology 2004;72(2):192–201. [PubMed: 15065954]

Simpson TL, Kivlahan DR, Bush KR, McFall ME. Telephone self-monitoring among alcohol use disorder patients in early recovery: A randomized study of feasibility and measurement reactivity. Drug and Alcohol Dependence 2005;79(2):241–250. [PubMed: 16002033]

Sobell, LSM. Alcohol timeline followback users' manual. Toronto, Canada: Addiction Research Foundation; 1995.

Stone AA, Shiffman S. Ecological momentary assessment (EMA) in behavioral medicine. Annals of Behavioral Medicine 1994;16(3):199–202.

Stone AA, Shiffman S, Schwartz JE, Broderick JE, Hufford MR. Patient non-compliance with paper diaries. British Medical Journal 2002;324:1193–1194. [PubMed: 12016186]

Surawy C, Cox T. Smoking under natural conditions: A diary study. Personality and Individual Differences 1987;8(1):33–41.

Swendsen JD, Tennen H, Carney MA, Affleck G, Willard A, Hromi A. Mood and alcohol consumption: An experience sampling test of the self-medication hypothesis. Journal of Abnormal Psychology 2000;109(2):198–204. [PubMed: 10895557]

Tidey JW, Monti PM, Rohsenow DJ, Gwaltney CJ, Miranda R, McGeary JE, et al. Moderators of naltrexone's effects on drinking, urge and alcohol effects in non-treatment-seeking heavy drinkers in the natural environment. Alcohol: Clinical and Experimental Research 2008;32:58–66.

Tindle, H.; Shiffman, S.; Paty, J.; Dang, Q. Relapsing in real time: A prospective assessment of the natural history of abstinence failure. Paper presented at the Annual Meeting of the Society for Research on Nicotine and Tobacco; Orlando, FL. 2006 Feb.

Todd M. Daily processes in stress and smoking: Effects of negative events, nicotine dependence, and gender. Psychology of Addictive Behaviors 2004;18(1):31–39. [PubMed: 15008683]

Todd M, Armeli S, Tennen H, Carney M, Affleck G. A daily diary validity test of drinking to cope measures. Psychology of Addictive Behaviors 2003;17(4):303–311. [PubMed: 14640826]

Todd M, Armeli S, Tennen H, Carney MA, Ball SA, Kranzler HR, et al. Drinking to cope: A comparison of questionnaire and electronic diary reports. Journal of Studies on Alcohol 2005;66:121–129. [PubMed: 15830912]

Todd M, Armeli S, Tennen H. Interpersonal problems and negative mood as predictors of within-day time to drinking? Psychology of Addictive Behaviors. (in press).

Toll BA, Cooney NL, McKee SA, O'Malley SS. Do daily interactive voice response reports of smoking behavior correspond with retrospective reports? Psychology of Addictive Behaviors 2005;19(3): 291–295. [PubMed: 16187808]

Toll BA, Cooney NL, McKee SA, O'Malley SS. Correspondence between interactive voice response (IVR) and timeline followback (TLFB) reports of drinking behavior. Addictive Behaviors 2006;31 (4):726–731. [PubMed: 15975732]

Tucker JA, Foushee HR, Black BC, Roth DL. Agreement between prospective interactive voice response self-monitoring and structured retrospective reports of drinking and contextual variables during natural resolution attempts. Journal of Studies on Alcohol & Drugs 2007;68(4):538–542. [PubMed: 17568958]

Weinstein SM, Mermelstein R, Shiffman S, Flay B. Mood variability and cigarette smoking escalation among adolescents. Psychology of Addictive Behaviors 2008;22(4):504–513. [PubMed: 19071975]