Assessment of Alcohol Use in the Natural Environment

Thomas M. Piasecki, Ph.D.

Department of Psychological Sciences, University of Missouri, 210 McAlester Hall, Columbia, MO 65211.

Abstract

The current article critically reviews three methodological options for assessing drinking episodes in the natural environment. *Ecological momentary assessment* (EMA) typically involves using mobile devices to collect self-report data from participants in daily life. This technique is now widely used in alcohol research, but investigators have implemented diverse assessment strategies. This article focuses on 'high resolution' EMA protocols that oversample experiences and behaviors within individual drinking episodes. A number of approaches have been used to accomplish this, including using signaled follow ups tied to drinking initiation, asking participants to log entries before and after individual drinks or drinking episodes, and delivering frequent signaled assessments during periods of the day when alcohol use is most common. *Transdermal alcohol sensors* (TAS) are devices that are worn continuously and are capable of detecting alcohol eliminated through the skin. These methods are appealing because they do not rely upon drinkers' self-report. Studies using TAS have been appearing with greater frequency over the past several years. New methods are making the use of TAS more tractable by permitting back-translation of transdermal alcohol concentration data to more familiar estimates of blood- or breath alcohol concentration. However, the current generation of devices can have problems with missing data and tend to be relatively insensitive to low-level drinking. An emerging area of research investigates the possibility of using *mobile device data and machine learning* to passively detect the user's drinking, with promising early findings. EMA, TAS, and sensor-based approaches are all valid, and tend to produce convergent information when used in conjunction with one another. Each has a unique profile of advantages, disadvantages, and threats to validity. Therefore, the nature of the underlying research question must dictate the method(s) investigators select.

Introduction

Field studies of alcohol use generate detailed descriptive information characterizing naturalistic drinking behaviors and their correlates, which can in turn inform theory generation, theory testing, and clinical decision making (e.g., Maisto, et al., 2014). Such studies can capture important, naturally occurring alcohol–related phenomena that would otherwise be difficult or unethical to study in laboratory conditions, such as underage drinking (e.g., Miranda et al., 2014a), the near-term correlates and consequences of high-dose alcohol exposures (e.g., Piasecki, et al. 2012a; Trela et al. 2016), and drinking lapses during treatment (e.g., Holt et al., 2012). Monitoring of alcohol use in naturalistic settings

Correspondence to: Thomas M. Piasecki.

can also be a valuable tool for evaluating treatment outcomes and delivering timely interventions (e.g., Dougherty, et al. 2015a; Suffoletto et al., 2014).

Alcohol researchers are now commonly using ecological momentary assessment (EMA) to investigate drinking behaviors in the natural environment (Beckjord & Shiffman, 2014; Wray et al. 2014). This approach has become increasingly tractable owing to the widespread adoption of smartphones and a proliferation of commercially available applications for implementing ambulatory studies on mobile operating systems. At the same time, alcohol researchers have been actively developing and testing wearable biosensors permitting passive, objective, and quasi-continuous monitoring of alcohol use over time (Swift, 2000; Leffingwell, 2013). The maturation of these lines of work now provides alcohol researchers with a set of powerful tools for studying alcohol use in the field. Emerging approaches, such as the development of computer algorithms that can passively detect drinking episodes from smartphone sensor data (Bae, et al., 2018, McAfee et al, 2017), promise to further expand the options for field studies of alcohol use.

The primary purpose of the current article is to review available methodological options for assessing drinking episodes in the natural environment and to consider their particular strengths and weaknesses. A central thesis is that there is no gold standard for field studies of drinking, so the research question at hand must play an important role in method selection.

Ecological Momentary Assessment

EMA is a measurement strategy that involves collecting frequent repeated measures from participants over time in their natural environments (Shiffman, et al 2008). Investigators have used a number of different terms to designate related research strategies, including experience sampling (Csikszentmihalyi and Larson, 1987; Myin-Germeys et al 2009), ambulatory assessment (Trull & Ebner-Priemer, 2014), and intensive longitudinal methods (Bolger and Laurenceau, 2013; Walls and Schafer, 2006). Rather than asking a participant to recall and reconstruct the past from memory, the EMA investigator seeks to collect frequent, presumably unbiased samples from the participant's stream of experience. This results in a more objective record of events that can later be interrogated by the researcher using rigorous data analytic techniques to address a wide array of research questions (Shiffman, 2014). A typical EMA study involves collection of self-report data, but it is possible to incorporate additional modes of assessment, including neurocognitive tasks (e.g., Jones et al., 2018), geolocation data (e.g., Epstein, et al., 2014), external physiological sensors (Bertz et al. 2018), and photos or videos (e.g., Alessi and Petry 2012; Luczak et al., 2015) to create more complex protocols.

Time-based protocols ask participants to complete assessments in response to scheduled prompts delivered at fixed times of day, at regular intervals, or at randomly determined times. In *event-based* protocols, participants are trained to initiate recording of assessments when a target event occurs. This is valuable when investigators are interested in capturing events (e.g., alcohol use episodes) that are central to the goals of the investigation, may occur relatively infrequently, and could be missed by time sampling alone. Investigators

often combine both modes of sampling, permitting comparisons of observations from focal events to those collected during time-sampled background states (Paty et al, 1992).

Application of EMA to Alcohol Research

There are a many resources dealing with broad topics related to conceptualization and implementation of EMA studies (e.g., Bolger and Lauranceau, 2013; Shiffman et al. 2008; Mehl and Connor, 2012), but a systematic treatment of issues relevant to assessment of alcohol use is lacking. Consequently, I highlight assessment considerations unique to the alcohol field and review a variety of strategies EMA investigators have adopted for assessing naturally occurring drinking episodes. The focus is on assessment of alcohol use per se. Some important topics (e.g., selection and psychometric evaluation of supporting self-report measures; data analysis) that are covered in the broader EMA literature fall outside the scope of the current review.

Researchers are commonly interested in knowing not just whether or not drinking occurred, but also in evaluating the peak levels and trajectories of effective alcohol doses. Formulas have been developed for translating the number of drinks consumed and time spent drinking into estimated blood alcohol concentrations (eBACs; Hustad and Carey, 2005). Though somewhat imprecise (Clapp, et al., 2006; Kraus, et al., 2005), eBAC calculations are useful tools for field investigations that rely on self-report for tracking alcohol use. Because they take into account sex differences in body composition, body weight, and typical alcohol elimination rate, eBACs are often preferable to simple drink totals, providing a more meaningful biological gradient against which to study subjective and behavioral responses to drinking. Imprecision in estimates may arise from a variety of factors, including discrepancies between consumed and standard drinks (e.g., Barnett et al., 2009; Kerr et al., 2005), drinkers' difficulty of reporting in standard drink units (e.g., White, et al. 2005), failure to account for factors associated with speed of alcohol absorption (e.g., whether drinks were consumed on an empty stomach), and individual differences in rates of alcohol elimination (e.g., Boyd and Corbin, 2018). There are several smartphone apps available that generate eBAC estimates directly based on user-entered information about personal characteristics and reported alcohol use (Hoeppner, et al., 2017). These can be productively incorporated into EMA studies (e.g., Luczak, et al., 2018). However, eBAC calculation apps should be selected cautiously and tested prior to use as commercially available options vary greatly in terms of the types data collected and accuracy of the estimates they output (Weaver, et al., 2013). Personal breathalyzers that connect to the user's phone via Bluetooth offer another promising approach to collecting information about levels of alcohol exposure in EMA designs (Riordan, et al. 2017).

Identifying the rising and falling limbs of the blood alcohol concentration (BAC) curve is important for many research questions. Alcohol has biphasic effects, with subjective stimulation being relatively more prominent on the ascending BAC limb and sedative effects more pronounced when BAC is falling (Martin, et al., 1993; Morean et al., 2013). Acute tolerance is observed for many alcohol effects, such that response are blunted on the descending limb relative to those observed at an equivalent BAC on the rising limb (e.g., Fillmore and Weafer, 2012).

An early impediment to the adoption of EMA in alcohol research was concern that acute intoxication could interfere with drinkers' ability to log user-initiated assessments, respond to signaled prompts in a timely fashion, and provide reliable self-report data (Collins, et al., 1998; Tidey, et al., 2008). Accumulating experience suggests drinkers' can achieve good compliance with signaled prompts and log reports at high drink totals and eBACs (e.g., Carpenter, et al., 2017; Trela, et al. 2016). Nonetheless, this issue deserves systematic study. For instance, researchers might examine whether assessment completion, coverage of alcohol consumption, and validity of real-time self-reports collected during drinking episodes vary as a function of total consumption indexed by next-day retrospective reports regarding the same episode. Such analyses could provide alcohol researchers more tangible information concerning the boundary conditions for tractable EMA monitoring.

High Resolution EMA Designs

In this article, I focus on a subset of intensive EMA approaches from the existing literature that I venture to call "high resolution" strategies for investigating alcohol use (Figure 1). The defining feature of these strategies is that they involve oversampling experiences during ongoing drinking, allowing investigators to capture intra-episode variations in states and behaviors. Notably, these approaches generally incorporate assessments of ongoing alcohol consumption, permitting calculation of momentary eBACs (assuming availability of necessary supporting information about sex and body weight). Repeated measures of eBAC within an episode can be used to characterize drinking patterns (e.g., speed of consumption; Carpenter, et al., 2017), to investigate dose-dependent subjective responses to alcohol (e.g., Miranda, et al., 2014a; Ray, et al., 2010; Trela, et al., 2016), and to investigate potential limb-specific effects (Piasecki, et al., 2012b).

The *timed follow-ups* approach involves instructing participants to initiate a recording when they begin drinking, and then delivering alarms periodically in the wake of this first drink log to prompt additional assessments (e.g., Luczak, et al., 2015; Piasecki, et al., 2012b; Swendsen, et al., 2000; Trela, et al., 2016). At follow-up entries, respondents can be asked to report any new drinks since the last report, building a cumulative record of consumption.

In the *drink bracketing* approach, participants are instructed to initiate recordings before and after each drink in an episode. This strategy has commonly been limited to the first few drinks due to concern that recording quality might deteriorate with extended drinking. One variant of this approach has been to collect a begin-drink report prior to and an end-drink report following each of the first two drinks (Miranda, et al., 2018; Tidey, et al., 2008; Ray, et al., 2010). A second variant involves training participants to initiate a log before and after each drink consumed, but configuring study software so that a full begin-drink assessment is administered only prior to the first drink and end-drink assessments are administered following each of the first three drinks in an episode (Miranda, et al., 2014a; Miranda, et al., 2014b; Miranda, et al., 2016; Ramirez and Miranda, 2014; Treloar and Miranda, 2017). This produces a thorough, time-stamped accounting of all drinks consumed within an episode, but self-report data are only collected for the early drinks in extended episodes.

The *high-frequency sampling* approach uses numerous, closely-spaced prompted assessments to capture drinking behaviors. Because the sampling strategy is intensive,

prompt scheduling is often concentrated during periods when drinking is especially likely. Investigators have employed several variants of this approach. For example, Dvorak, et al., (2014; 2016) collected prompted assessments at random times within each of 9 two-hour windows between 8 am and 2 am. At each assessment, participants reported how many drinks they had consumed since the last report. Gautreau, et al. (2015) prompted participants at random times within 2-hour periods from 4 pm to 4 am, asking participants to provide a running tally of drinks for the current day at each prompt. Suffoletto, et al. (2018) prompted participants hourly between 8 pm and 12 am on Friday and Saturday nights, collecting reports of the number of drinks consumed since the preceding assessment. I count these approaches as high resolution designs because, in principle, they permit tracking alcohol consumption, eBAC, and associated states and contexts across the entire drinking episode. However, these goals were typically not those driving the design of these studies. The existing studies sought to establish the intraday temporal precedence of theoretically relevant triggers in the pre-drinking period (e.g., mood states, motives) and characterize the occurrence and intensity of subsequent drinking (Dvorak, et al., 2014; 2016; Gautreau, et al., 2015) or to capture variations in eBAC as criterion data against which to validate algorithms for passive detection of drinking (Gharani, et al., 2017; Suffoletto, et al., 2018). The high-frequency sampling designs were well-suited to addressing these aims.

Finally, the *episode bracketing* approach instructs participants to initiate assessments at the beginning of a drinking episode and at the completion of the episode (Collins, et al., 1998; 2003; Kashdan, et al., 2010; Mayhugh, et al., 2018 Muraven, et al., 2005). This simple pre-post design allows assessment of baseline conditions in the begin drinking report and the amount consumed and effects of drinking at the close of the episode.

These high-resolution EMA designs have been used to study a variety of populations, including adolescents (e.g., Miranda, et al., 2014), college students and young adults (e.g., Dvorak, et al., 2014; Kashdan, et al., 2010), psychiatric outpatients (e.g., Carpenter et al., 2017), and treatment-seeking adult heavy drinkers (e.g., Collins et al., 1998; Tidey, et al., 2008). The modal duration of EMA monitoring has been about 2–3 weeks, but this varies considerably across studies, ranging from just a few days (e.g., Mayhugh, et al., 2018; Suffoletto, et al., 2018) up to eight weeks of daily recording (Collins et al., 1998).

Lower Resolution EMA Designs

Most existing EMA studies assessing alcohol have used designs that do not involve repeated assessments within episodes. Many employ what I would call "medium resolution" designs that capture intra-day variation in alcohol use. A common approach is to collect periodic assessments (at random times or fixed intervals) across the day, assessing alcohol consumption in each survey. Some studies use a "dipstick" strategy, asking whether any alcohol use occurred during some period (e.g., past 30 or 60 minutes) prior to the survey (e.g., Gorka et al., 2017; Simons, et al., 2015). Others attempt to more completely capture drinking occurrence by asking respondents to indicate any alcohol use since the last EMA report (e.g., Fatseas et al., 2018; Trull et al., 2016).

A large number of studies use comparatively "low resolution" assessments that capture variations at the between-day level. A common technique is to use a once-daily morning

report assessing details of any prior night drinking (e.g., Crooke et al., 2013; Morgenstern, et al., 2016). Sparser designs are possible, such as delivering weekly prompts asking participants to report drinking for each day over the past week (e.g., Voogt et al., 2014).

These designs all have important applications in alcohol-focused research, and are frequently used to incorporate assessment of alcohol into investigations primarily targeting other content domains, such as smoking or symptoms of psychiatric disorders (e.g., Black, et al., 2016; Tomko et al., 2017). These lower resolution approaches are not the focus of the current article because the relevant methodological considerations are well covered in the broader literature on EMA in substance use (e.g., Ferguson and Shiffman, 2011; Shiffman, 2009; 2014). However, these strategies bear mentioning because they can complement high-resolution approaches and are frequently used in combination with them (e.g., Dvorak et al., 2016; Epler, et al., 2014; Treloar and Miranda, 2017).

Comparing High Resolution EMA Designs

Initiation of Recording and Available Baseline Data.—The timed follow-ups, drink-bracketing, and episode bracketing approaches rely upon users to initiate an EMA assessment when drinking begins. Studies using timed follow-ups have differed with respect to instructions, with some asking drinkers to initiate the timed follow-up sequence either just before (e.g., Luczak, et al., 2015; Swendsen, 2000) or just after the first drink (e.g., Carpenter, et al., 2017; Piasecki, et al. 2012b). In the drink- and episode-bracketing designs, drinkers have generally been instructed to initiate a recording prior to the first drink.

Asking users to initiate a diary entry *before* drinking allows the investigator to capture clean data concerning pre-drinking states and contexts. A potential downside of this strategy is that it assumes participants are willing and able to interrupt drinking initiation to make a pre-drink report. If ability or inclination to stop is systematically related to drinking behaviors or important individual difference factors, pre-drinking reports could be biased. Empirically addressing this issue would be informative for designing new studies. A tractable approach might involve instructing participants to initiate drink reports prior to the first drink and also including items assessing whether or not this instruction was followed in the corresponding EMA report. Between-person analyses could then test whether delayed responding is associated with individual difference variables. Within-person comparisons of timely vs. delayed drinking initiation reports could potentially identify whether delayed reports are characterized by a biased set of contextual factors or subjective states.

Instructing participants to initiate a recording *after* the first drink does not require interruption of the early self-administration sequence, but sacrifices collection of clean baseline data for a drinking episode. If this strategy is used in conjunction with other kinds of assessments, it may be possible to compensate for the lack of a true pre-drink baseline observation. For example, if random prompts are administered throughout the day, the nearest pre-drinking random prompt assessment may serve as an approximate baseline. This may or may not prove to be a meaningful benchmark depending upon the spacing of assessments and the stability of the construct(s) of interest. An alternative approach is to incorporate items in the first drink log that ask participants to report both their current experiences and to retrospectively report experiences just before beginning the drink. Such

retrospective pre-drink reports should be interpreted cautiously because they could be influenced by the early effects of alcohol or some of the same kinds of retrospective biases EMA seeks to bypass.

The timed follow-ups strategy usually depends upon the participant to log the first drink in order to start the sequence. As a potential failsafe, some studies have incorporated questions about recent alcohol use in other ongoing assessments, using affirmative responses to trigger the follow-up sequence. This design feature may be a valuable way to increase coverage, as findings indicate that a substantial proportion (30–45%) of the drinking episodes captured in these studies were triggered by mechanisms other than user-initiated first drink logs (Carpenter, et al., 2017; Trela, et al, 2016). The prevalence and interpretation of this phenomenon is likely to differ across studies owing to variation in the frequency and characteristics of other assessments and whether participants are tasked with logging the first drink before or after its completion. When using backup mechanisms to detect drinking and trigger timed follow-ups, it is valuable to include assessments of the number of drinks consumed and the time since drinking initiation prior to the triggering record to permit adjustments to the timeline and cumulative drink total for the episode.

The high frequency sampling design lacks a user-initiated report flagging the start of the drinking episode. This means that drinking occurring outside the time period selected for intensive sampling will generally not be captured. Because drinking episodes are detected only after they are underway, questions about time since drink initiation are needed to identify the start time and duration of the bout. Self-report data collected at the time drinking is first detected may be influenced by prior alcohol consumption and large portions of the drinking episode could be missed, particularly when long and variable inter-assessment intervals are used.

Termination of Recording.—A key choice point in using the timed follow-ups strategy concerns how the sequence of assessments should be terminated once it has been triggered. One could choose to deliver a fixed number of follow-ups covering a set amount of time after every drinking episode is detected. The potential downside is that this could fail to fully capture prolonged drinking episodes. To address this, investigators have implemented flexible schemes in which the length of the follow-up interval can adapt based on drinking behavior (Lane, et al., 2016; Luczak, et al., 2015; Piasecki, et al., 2012b).

The desire to capture as much of the post-drinking experience as possible with follow-ups has to be balanced against the "livability" of the assessment protocol. Because many drinking episodes occur late in the evening, some follow-up alarms are likely to be scheduled for times after participants have gone to bed. To reduce inconvenience, participants can be trained to make user-initiated bedtime reports that immediately terminate the follow-up sequence (Luczak, et al., 2015; Piasecki, et al., 2012b) or allowed to mute alarms.

In typical usage, the drink bracketing strategy only incorporates detailed assessments surrounding the first few drinks, but many drinking episodes will continue beyond this focal window. Participants averaged approximately 5 drinks per episode in drink bracketing

studies targeting the first two (Ray et al., 2010) and first three drinks (Miranda, et al., 2014b), meaning the self-report data tended to be biased toward the lower end of the BAC range. Coverage could be increased by adding additional end-drink assessments. Unfortunately, there is little available empirical evidence concerning how far this process can be extended successfully. Experimental studies in which participants are randomized to differing numbers of end-drink assessments would permit comparative evaluation of compliance, episode coverage, and user satisfaction.

The episode bracketing approach requires that drinkers initiate a recording at the end of a drinking episode, a task that is somewhat more ambiguous than the drink-bracketing instruction tied to completion of a drink. To aid participants, investigators have provided guidance for defining discrete episodes using criteria such as passage of time and changes in location and activity (Collins, et al., 1998; 2003; Muraven, et al, 2005).

The high frequency sampling approach is terminated when the period of pre-determined concentrated prompts is reached on a given day. Some studies using this approach include scheduled prompts extending into the late night or early morning to capture extended drinking. In these studies, participants are instructed to ignore prompts occurring after bedtime (Gautreau, et al., 2015) or provided with a method for muting alarms (Dvorak, et al., 2014; 2016).

A problem common to all the high-resolution designs is that termination of the data stream may not reflect the true end of drinking. For example, follow-ups, random prompts, and end-drink reports can be missed, bedtime reports can be intentionally or unintentionally misused to terminate prompts, or a change of plans may extend drinking longer than initially intended. Missing end-drinking reports may be especially problematic for the episode bracketing strategy because all information about acute alcohol effects and amount consumed will be missed. For these reasons, use of next-day morning reports as a secondary source of information about the entire prior-night's drinking episode is common in EMA studies of alcohol (e.g., Dvorak, et al., 2016; Miranda, et al., 2016; Muraven, et al., 2005; Piasecki, et al., 2012a;).

Intra-Episode Temporal Resolution.—In the timed follow-ups design, the scheduling of assessments is tied to the first drink report with follow-ups occurring at investigator-determined intervals. A simple approach is to use regular intervals, such as prompting assessments every hour following the initial drink report (Swendsen, et al., 2000). A variant of this strategy uses shorter intervals (e.g., 30 or 45 minutes) early in the episode, then switches to longer intervals (e.g., 60 minutes) after the first few prompts (e.g., Carpenter, et al., 2017; Luczak, et al., 2015; Piasecki, et al., 2012b). Conceptually, the front-loading of timed follow-ups is appealing for studies of subjective response to alcohol, as earlier assessments are more apt to capture the shorter-lived rising limb of the blood alcohol concentration curve.

High frequency sampling also follows an investigator-determined schedule, and more concentrated assessments have the potential to yield many intra-episode data points. If randomly-timed prompts are used, more variability in inter-assessment intervals will result.

In the drink bracketing approach, reports are tied to completing individual drinks, leading to some variation in the intervals between successive assessments. Temporal resolution and episode coverage are limited by the number of end-drink reports the investigator decides to collect, with fewer data points available when a more conservative total is selected. The amount of data collected in an episode is also dictated by the number of drinks consumed. Lighter drinking episodes may only yield a few data points even when many end-drink reports are possible.

Compared to typical drink-bracketing designs, episode bracketing is more likely to cover the entire drinking bout. However, episode bracketing offers the lowest temporal resolution. The two data points collected per drinking episode permit only pre- vs. post-drinking comparisons, and cannot resolve any nonlinear change processes between the two assessments.

Coverage of Descending BAC Limb.—Assessments in the timed follow-ups and high-frequency sampling approaches may extend past the end of a drinking episode, capturing moments on the descending BAC limb (Piasecki, et al., 2012b; Suffoletto, et al., 2018). Descending limb observations may be censored by users terminating signaled assessments at bedtime. For instance, despite the fact that the descending limb tends to be longer in duration than the ascending limb and our use of an extensible follow-up scheme, we found that only about 25% of our post-drinking assessments could be categorized as descending limb moments (Piasecki, et al., 2012b; Piasecki, et al., 2014). It may simply be inherently difficult to study biphasic effects of alcohol using EMA because drinkers often sleep during the descending limb.

In the drink-bracketing approach data collection is tethered to the consumption of new drinks. As a practical consequence, self-report data will tend to be largely limited to moments from the rising limb of the blood alcohol concentration curve. Available findings indicate that fewer than 2% of the self-report assessments captured in drink-bracketing studies occurred when eBAC was lower relative to the preceding observation in the same episode (Miranda et al. 2014a; 2014b; 2016). Drinkers in episode bracketing studies typically log the end-of-episode assessment within a few minutes of finishing the last drink (Collins, et al., 1998; 2003), suggesting these designs also primarily capture self-reports on the ascending limb.

Burden.—Collecting more assessments per episode, decreasing inter-assessment intervals, and lengthening the duration of the total study monitoring period yield more data, but at the cost of increased burden for participants. It is possible that frequent interruption of the drinking episode intrudes upon and alters the experience, perhaps diminishing ecological validity. Although one can readily appreciate these kinds of trade-offs in the abstract, at present there is little empirical evidence addressing whether choices about assessment intensity make a substantial difference or pointing to a set of best practices for alcohol studies. Experimental studies in which participants are randomized to protocols of varying intensity are needed to better characterize how different protocol features sum or interact to potentially influence participants' real-time subjective reports, perception of the study experience, and willingness to enroll in similar studies in the future.

In a 14-day study using timed follow-ups (Luczak, et al., 2015), 74% of college students deemed monitoring the drinking experience as neutral or a little fun or interesting. Only 3% of the morning reports indicated that the assessment was very unappealing. Similarly, young adult social drinkers who participated in a 14-day study using an episode bracketing scheme administered by an interactive voice response system reported high levels of satisfaction and ease of use (Collins, et al., 2003).

One study of community-recruited frequent drinkers allowed participants to choose whether to use a timed follow-ups scheme with hourly prompts or to log each drink as it was consumed (Swendsen, et al. 2000). Drink bracketing was vastly preferred, with only 18.5% of drinkers selecting the timed follow-ups option. However, there was a significant correlation between the number of drinks per drinking day and selection of the timed follow-ups strategy. This suggests that the drinking habits of the sample may be relevant for determining which strategy will be most acceptable. For heavier drinkers, responding to occasional alarms may be less burdensome than logging every drink. For light drinkers, responding to a fixed series of follow-ups may require more work than simply logging one or two drinks.

In a drink bracketing study focused on the first two drinks in a treatment-seeking sample of heavy drinkers, Ray et al. (2010) sought to mitigate burden by assessing contextual and environmental aspects of the drinking episode at the first post-drink report and mood items at the second post-drink assessment. This approach results in less work for participants (relative to assessing all constructs at each post-drink assessment). A potential cost is that BAC level may be confounded with measurement of particular constructs and the ability to characterize changes occurring between drinks may be compromised. Similarly, abbreviated follow-up surveys prioritizing key constructs expected to change in response to alcohol have been used (e.g., Trela, et al., 2016). This has meant that intra-episode changes in measures omitted from follow-ups (e.g., location) could not be tracked. The nature of the focal research question will dictate whether these kinds of tradeoffs are acceptable.

When designing a protocol combining multiple sampling schemes, it is useful to consider areas of overlap that might result in unnecessary assessment burden. In our work, we have employed the timed drinking follow-ups strategy in combination with several random prompts per day and user-initiated reports of cigarette smoking (Piasecki, et al., 2011; Piasecki, et al., 2012b). To avoid overwhelming participants during drinking episodes, activation of the drinking follow-up sequence automatically pre-empted delivery of any previously scheduled random prompts. Additionally, the button participants would have used to log cigarettes was removed from the device screen during drinking episodes. Instead, participants were asked how many cigarettes they had smoked since the last report as part of each drinking follow-up assessment.

High frequency sampling requires that participants respond to numerous prompts. If the focus of the research is on intra-episode change processes, much of the burden associated with high frequency sampling may be unnecessary. This is because many (perhaps most) prompts will be delivered when participants are not actively drinking. On the other hand,

frequent assessment of non- or pre-drinking states may be desirable for other research questions (e.g., Dvorak, et al., 2014; Gautreau, et al., 2015).

Compliance.—Completion rates for timed follow-ups are typically good, with estimates ranging from 84–93% in samples of young adults and psychiatric outpatients (e.g., Luczak, et al, 2015; Lane, et al., 2016; Piasecki, et al., 2012b). High frequency sampling studies have reported response rates to signaled prompts ranging from 63–90% in college student samples (Dvorak, et al., 2014; 2016; Gautreau, et al., 2015).

The available evidence from drink bracketing studies suggests these approaches capture most targeted drinks. For example, in a study using begin- and end-drink reports for the first two drinks, treatment-seeking heavy drinkers indicated they had forgotten to enter a prior-day drink in only 3.9% of daily morning assessments (Ray, et al., 2010; Tidey et al, 2008). Expected end-drink reports were logged for between 92–99% of first begin-drink reports and 92–100% of second begin-drink reports (Ray, et al., 2010; Tidey et al, 2008). Comparisons of EMA drinking data from adolescent drinkers to timeline follow-back assessments of the same period have found strong correlations (rs = .72 to .90) for total number of drinks, number of drinking days, and number of heavy drinking days (Miranda, et al., 2014a; 2014b).

Questions about the timeliness of reporting have also been used to characterize compliance. Requiring that begin-drink assessments be logged within 5 minutes of starting a drink has resulted in discarding 9–25% of drinking episodes in samples of adolescent drinkers (Miranda, et al. 2014b; 2016). Tidey et al, (2008) observed an average reported delay of approximately 3 minutes between starting a drink and logging begin-drink reports in a sample of treatment-seeking heavy drinkers. Excluding user-initiated reports with delays of 10 minutes or more (and applying some additional compliance criteria) required dropping 16–19% of begin- and end-drink reports (Tidey, et al., 2008). Miranda and Treloar (2017) reported that 48% of the begin-drink assessments from a sample of adolescents and young adults were logged after participants had started to consume the first drink. Limiting analyses to assessments within 10 minutes of starting a drink resulted in exclusion of 26% of the begin-drink reports. Together, these findings suggest applying conservative constraints on the timeliness of drink logging may result in substantial data loss. However, there is little available empirical data evaluating whether imposing these constraints affects analytic conclusions. In some cases, investigators have indicated that findings were unchanged (Miranda, et al., 2016; Treloar and Miranda, 2017) but this question has not been addressed in many reports using timing-related exclusions.

Episode bracketing studies have primarily used timeliness of reporting to index compliance. Young adult and treatment-seeking older drinkers have generally reported small delays relative to the target event, with 93%- 95% logging begin-drink reports within 5 minutes of starting to drink, and 48–68% logging end drinking reports within 1 minute of finishing the last drink (Collins, et al., 1998; 2003; Muraven, et al., 2005). Less information is available concerning completed pairs of assessments. In one study, college drinkers logged a begin- or end-drink report for 47 total episodes, but both reports were available for only 75% of these drinking occasions (Collins, et al., 2003).

EMA researchers often provide incentives and incorporate other procedures into studies in attempts to foster participant retention and compliance with signaled prompts. Examples include providing escalating payments over time in the monitoring period (Piasecki, et al., 2011), decreasing compensation as a function of the amount of missing diary entries (Lane et al., 2016), paying participants a fixed amount for every signaled prompt that is completed (Dvorak, et al., 2016; Suffoletto, et al., 2018), paying participants according to the number of days they record field data (Treloar and Miranda, 2017), discontinuing participants who fail to achieve a minimal compliance standard during a trial period (Tidey et al., 2008), and providing periodic feedback on compliance rates to participants (e.g., Collins, et al., 1998; Piasecki, et al., 2011). Systematic empirical investigations characterizing the effectiveness of various procedures would be valuable for researchers planning future studies, but are currently lacking.

Reactivity.—The data on possible reactivity to high resolution EMA designs is very limited. Using an episode bracketing design to study a drinking moderation training program in a sample of heavy drinking adults, Collins et al. (1998) found that excessive drinking (defined as consuming 5+ drinks in an episode) decreased significantly over an 8-week EMA monitoring period in both the active treatment condition and an assessment-only control group. This may indicate some reactivity of heavy drinking to EMA occurred in the control condition. However, participation required expressing an interest in cutting down one's drinking and the EMA recording was presented to control group participants as an important first step in this process.

On mornings following drinking episodes, Luczak and colleagues (2015) asked college students to evaluate the effect of the monitoring on their drinking using time-based follow-ups. Most (71%) indicated they would have consumed the same amount had they not been monitoring their drinking, 17% said they would have consumed more, and 8% said they would have drunk less. In a 21-day EMA study using timed follow-ups, timeline follow-back data collected at baseline were used to estimate how many drinks would be expected during the monitoring period in sample of community-recruited frequent drinkers (Piasecki, et al., 2011). The EMA data captured approximately 4% more drinks than projected from baseline drinking reports, suggesting limited reactivity of drinking.

Transdermal Alcohol Sensors

A small portion of ingested alcohol is excreted through the skin (Swift, 2003). Specialized devices worn on the wrist or ankle can be used to frequently analyze the alcohol concentration in samples of insensible perspiration, providing a method for assessing the wearer's drinking (Greenfield, et al., 2014; Leffingwell, et al., 2013). Because transdermal assessment is passive, data collection does not depend upon participants' effortful compliance with recording instructions or ability to report consumption in standard drink units.

One barrier to adoption of TAS is that transdermal alcohol concentration (TAC) data are complexly related to measures of BAC or breath alcohol concentration (BrAC) to which alcohol researchers are more accustomed (Luczak, et al, 2018). TAC curves tend to be right-

shifted and have lower peaks relative to BAC and BrAC (Sakai, et al., 2006). Additional complexity arises because factors such as individual differences in skin thickness and variation in device characteristics can moderate the relationship between TAC and BAC/ BrAC (Luczak and Rosen, 2014; Swift, 2000). Accounting for these factors has often required an initial laboratory-based calibration session to establish the TAC output produced by a particular device-user coupling in response to a standardized alcohol challenge input. A number of statistical models and software tools are being developed to facilitate back-translation of observed TAC field data to estimates of BAC, BrAC, or standard drinks (Dougherty, et al., 2015b; Hill-Kapturczak, et al., 2014; Luczak and Rosen, 2014). As more TAC datasets are amassed, it may become increasingly feasible to use these tools without a time-consuming in-lab calibration session for each participant using the TAS (e.g., Sirlanci, et al., 2018).

Variants of two TAS devices have been used in the bulk of the available literature (Leffingwell, et al. 2013). The Secure Continuous Remote Alcohol Monitor (SCRAM, Alcohol Monitoring Systems, Inc., Littleton, CO) is a bracelet that is locked to the ankle. The Wrist Transdermal Alcohol Sensor (WrisTAS; Giner, Inc., Newton, MA) is a wrist-worn bracelet. Both devices contain electrochemical sensors to determine alcohol concentration in insensible perspiration in an enclosed area above the skin and have additional sensors measuring skin temperature and contact to confirm the device is being worn. Smaller and more efficient devices are being developed, but currently have limited commercial availability and lack supporting data (Wang, et al., 2018).

TAS field studies have been conducted in various populations, including college students and young adults (e.g., Luczak, et al. 2015; Neville, et al. 2013), treatment-seeking (e.g., Barnett, et al., 2014) and non-treatment-seeking heavy drinkers recruited from the community (Dougherty, et al., 2015), outpatients from alcohol treatment facilities (Alessi, et al., 2017), and adults recently arrested for DUI (e.g., Mathias, et al. 2018). Investigations using WrisTAS have typically involved wearing the device for 1–2 weeks (Luczak, et al., 2015; Marques and McKnight, 2009; Simons, et al., 2015). SCRAM studies have been more variable in length, ranging from a single day (Clapp, et al., 2017) to 16 weeks (Dougherty, et al. 2015).

Below, I review the available evidence concerning TAS assessment using some of the same broad criteria previously applied to the EMA designs.

Initiation and Termination.

Because transdermal sensors may be worn continuously, they should capture the entirety of a drinking episode. Ultimately, however, the collected data stream must be inspected in order to identify drinking episodes. Uncertainty is introduced due to factors such as exposures to environmental alcohol, device removal, and sensor failure (Barnett, et al., 2014; Bond, et al., 2014; Clapp, et al., 2017).

The SCRAM device manufacturer applies a set of proprietary criteria for determining that a pattern of TAC observations represents a likely drinking event. These criteria are inherently conservative, as appropriate for the forensic applications for which they were developed.

This results in many false negatives when compared to self-report of drinking in community samples, particularly for light drinking episodes (Barnett, et al., 2014; Dougherty, et al., 2012). Accordingly, researchers have begun to develop and evaluate more liberal criteria sets to more completely capture drinking events in basic research (Barnett, et al., 2014; Roache et al., 2018) and tools for processing and interpreting SCRAM data (Barnett, et al., 2015). Investigations using WrisTAS have employed varying approaches, including a mathematical filtering scheme (Rosen et al., 2014) and differing thresholds on peak TAC (Marques and McKnight, 2009; Simons, et al., 2015) to identify drinking events.

Intra-Episode Temporal Resolution.

TAS devices offer good to excellent temporal resolution of TAC. Studies using SCRAM devices typically collect samples every 30 minutes, whereas WrisTAS studies have been used to collect samples every 5 minutes. (e.g., Barnett, et al., 2014; Luczak, et al., 2015; Marques and McKnight, 2009). Empirically derived equations have been developed to translate TAC parameters, such as peak level and time-to-peak TAC into estimates of BrAC or standard drink equivalents (Dougherty, et al., 2012; 2015b; Hill-Kapturczak, et al., 2014; 2015). Although useful for many purposes, these techniques yield single point estimates, sacrificing some of the temporal information contained in the original TAC record. Innovations in mathematical modeling and software development now allow for back-translating individual TAC data points into a corresponding set of points forming an estimated BrAC or BAC curve that can characterize the dynamics of intoxication during naturalistic drinking episodes in greater detail (Luczak and Rosen 2014; Luczak et al, 2018; Rosen, et al., 2014).

Coverage of Descending BAC Limb.

TAS devices are worn continuously, so they capture information from the entire episode including the descending limb. Because data collection is not interrupted by sleep, descending limb observations will not be censored as is typical in EMA data. Advances in TAC back-translation permit identification of the timing of the estimated BrAC/BAC peak as a point of demarcation between ascending and descending limbs (Luczak and Rosen 2014; Luczak et al, 2018; Rosen, et al., 2014).

Compliance, Burden, and Reactivity.

Compliance with transdermal monitoring can be assessed by examining sensor data indicating whether or not the device is currently worn against the skin. One study using WrisTAS reported that 6% of college student participants failed to consistently wear the sensor (Luczak, et al, 2015), and another study involving students found that the sensor was removed on approximately 10% of all study days (Simons, et al., 2015). In a clinical sample using SCRAM devices, 34% of participants tampered with the sensor, and 56% of these tamper events were linked to drinking (Alessi, et al., 2017).

The most common complaints reported in studies using SCRAM have been discomfort, interference with physical activities, and stigma associated with device usage (Alessi, et al., 2017; Barnett, et al., 2011; 2017; Sakai, et al., 2006). Approximately 9% of an alcohol outpatient treatment sample refused study participation due to the bracelet (Alessi, et al,

2017). Treatment-seeking heavy drinkers reported noticing the device approximately once per hour during waking and twice per night during sleep (Barnett, et al., 2011). Despite some discomfort, most participants indicate a willingness to wear the device longer (Alessi, et al., 2017; Barnett, et al., 2017).

Evidence for reactivity to transdermal monitoring is mixed, and may depend on sample characteristics and drinking goals. In an observational study of community-recruited young adult drinkers, there was no evidence that wearing the WrisTAS sensor influenced either the likelihood of drinking or the amount consumed during drinking episodes (Simons, et al., 2015). Use of the SCRAM sensor improved compliance with instructed abstinence in a sample of college students (Neville, et al., 2013). Participants attributed this effect to the knowledge that any drinking would be detected and the ability to use the device as a way to defuse social pressure to drink. Similarly, treatment-seeking drinkers have reported that wearing SCRAM helped them to avoid drinking (Alessi, et al., 2017).

Device Malfunction.

Erratic output and device failure have been commonly reported in TAS studies (Bond, et al., 2014; Greenfield, et al., 2014; Leffingwell, et al., 2013; Marques & McKnight, 2009). In a pooled analysis of two clinical studies using SCRAM devices, data from approximately 5% of study days were affected by device malfunction (Barnett, et al., 2014). Studies using WrisTAS in observational studies of young adults have reported device malfunction on approximately 18% of study days (Simons, et al., 2015) and that only 38% of participants had a complete, interpretable data series over a two-week period (Luczak, et al., 2015). Similar problems have been observed with newer sensor devices (Wang, et al., 2018).

Comparisons with Laboratory and Diary Measures.

Various parameters of TAS data, such as peak TAC level, time to peak, and area under the TAC curve, are robustly related to BrAC measures and number of drinks consumed in laboratory challenge conditions (e.g., Dougherty, et al., 2012; Hill-Kapturczak, et al., 2014; 2015) and in EMA reports (Barnett, et al., 2014; 2017; Fairbairn, et al., 2018; Simons, et al., 2015). Available evidence is mixed, however, concerning the ability to detect drinking events on the basis of TAC readings alone. Studies applying varying episode detection rules to SCRAM data have reported correctly identifying between 39.9 to 87.7% of drinking events observed in the laboratory or reported by drinkers in the field (Barnett, et al., 2011; 2014; Karns-Wright, et al., 2018; Marques and McKnight, 2009; Neville, et al., 2013; Roache, et al., 2018). Drinking events undetected by SCRAM tend to be those involving lower levels of consumption (e.g., 1–4 drinks; Barnett, et al., 2014; Karns-Wright, et al., 2018; Roache, et al., 2015; 2018). In studies using WrisTAS, TAC data have detected between 23.6 and 85.6% of drinking events identified by other means (Bond, et al., 2014; Marques and McKnight, 2009; Simons, et al., 2015).

Passive Detection Using Sensor Data from Mobile Devices

An emerging line of inquiry has attempted to use machine learning algorithms to passively detect alcohol use based upon sensor data collected from mobile technologies. Investigators

have used measures of gait and posture derived from inertial sensors in smartphones and watches (Aiello and Agu, 2016; Arnold, et al., 2015; Gharani, et al., 2017; McAfee, et al., 2017; Suffoletto, et al., 2018) and other smartphone data such as communications, typing, and detected physical activity to make predictions concerning the alcohol consumption of the user (Bae, et al., 2017; 2018). In these studies, the stream of sensor data is segmented and each segment is labeled according to indices of alcohol use collected in the lab (e.g., real or simulated BAC) or in the field (e.g., number of drinks or estimated BAC based on EMA reports). Most of the data is used to train machine learning algorithms to classify data segments according to alcohol use criteria, and performance of the resulting prediction models are evaluated in a subset of the data reserved for validation testing.

Early findings from field studies are encouraging. For example, 70% accuracy was found for predicting the approximate number of drinks consumed using gait data from mobile phones (Arnold, et al., 2015). A study using a standardized phone-based walking task found that model-predicted estimates of BAC based on gait features correlated at .90 with eBAC estimates generated from hourly EMA assessments of drinking during the same episodes, with most errors in predicted eBAC falling between +/− .012 g/dl (Suffoletto, et al., 2018). Device usage and sensor data were able to identify high-risk drinking occasions (i.e., >4/5 drinks for women/men) with 90% accuracy within 30 minutes of drinking initiation (Bae, et al., 2018). More work is needed to evaluate the replicability of these findings and develop and disseminate these methods for usage in larger field studies.

At present, there is limited data for comparing these sensor-based passive detection schemes against other assessment alternatives. One model system has been set up to continuously sample sensor data in small segments (e.g., 5 sec), feed them to prediction models as available, and then provide feedback from the model results to the user (e.g., McAfee, et al., 2017). Using this kind of system, it may be possible to achieve very good temporal resolution of eBAC or other alcohol-related estimates in the field. Sensor data can be collected on both the ascending and descending limbs of the BAC curve (Gharani, et al, 2017). However, as these systems typically depend upon user interaction with the device(s) for input, sleep will tend to censor many falling limb observations.

Burden, reactivity and compliance may depend upon system features. In a study of young adults using an installed app to collect data on phone usage, sensor data were only captured on 57% of study days (Bae, et al., 2018). Participants cited concerns about data privacy and perceptions of battery drain as reasons for having disabled the study app. Protocols that require participants to follow a set of specific instructions (e.g., providing a standardized walking sample; Gharani, et al., 2017; Suffoletto, et al., 2018) are likely to be more burdensome compared to approaches that passively collect data from ongoing activities. Noisy data may result from a number of sources, such as failure to carry the mobile device(s), carrying the device(s) in a bag rather than on one's person, and use of the device by non-participants (e.g., Arnold, et al., 2015).

Overall, these mobile sensor-based approaches are in an early stage of development. They offer the promise of new sources of information for capturing alcohol use parameters in field work that may complement or be alternatives to EMA and TAS approaches. Their ultimate

value remains to be seen as the underlying technology and methods develop, experience with their use accumulates, and the practical hurdles are confronted. Developing these approaches requires expertise in mobile computing and predictive analytics. Accordingly, these projects will generally require collaboration between computer scientists and alcohol researchers with biomedical or social science backgrounds.

Comparing and Combining Assessment Options

Researchers now have a number of options for assessing 'real world' drinking. The supporting literature indicates that EMA, TAS, and sensor-based approaches are all valid, and tend to produce convergent information when used in conjunction with one another. Each has a unique profile of advantages, disadvantages, and threats to validity.

TAS methods are appealing because they permit excellent temporal resolution and use objective input rather than relying upon users' self-report. A major limitation is that TAS data only captures information about alcohol concentration or dose. This information is sufficient for many applications. For example, TAS is well suited to contingency management interventions, allowing biomonitoring to determine whether patients attained abstinence goals required to earn reinforcers (Barnett, et al., 2017; Dougherty, et al., 2015a). Because the current generation of sensors may fail to detect lighter drinking occasions and can sometimes malfunction, it is often desirable to collect supporting data on drinking behaviors during the TAS monitoring period using EMA or other self-report strategies.

A great many important questions in alcohol research, however, require some form of self-report data be collected during drinking episodes. Alcohol researchers often want to address questions such as how factors such as mood and social context relate to drinking, what motivates individual drinking events, and how drinkers experience the acute effects of alcohol. When the investigation requires collecting self-report data, some form of EMA must be incorporated into the design. Simultaneous TAS may be useful for corroborating participants' drinking reports and providing higher-resolution estimates of an episode's BrAC/BAC curve. However, once self-report data are needed to address a particular research question, the inherent limitations of EMA and all of the associated decisions concerning EMA protocol design must be confronted. In a combined EMA/TAS protocol, the more restrictive constraints of the EMA on temporal resolution, descending limb coverage, compliance, and so on, effectively apply in the final analysis of self-report outcomes. Because the EMA and TAS data are robustly associated with one another, researchers who require self-reports from the field should consider whether they have a compelling additional uses for the TAS data to justify a combined approach.

The passive detection of drinking from mobile device data generally requires independent information about alcohol use to label segments of sensor data and enable machine learning classifications. Studies developing these systems can benefit from incorporating EMA or TAS data collection to obtain criteria against which to identify predictive features of sensor data during system development and to validate trial systems in independent samples.

Conclusion

Different research aims will require varying kinds of data, and available methods for field studies of alcohol use differ with respect to important parameters, such as temporal resolution, precision of alcohol dose estimates, coverage of drinking episodes, and assessment burden. Accordingly, there is no one-size-fits-all solution to measuring drinking in field studies. Instead, investigators must make considered choices from available options that best suit their unique research needs. Continued research systematically addressing the effects of different EMA protocol features (e.g., frequency and timing of follow-ups, reliability and validity of self-report at high levels of alcohol consumption) and exploring new methods of passive monitoring of alcohol use is needed to assist researchers to make informed choices.

Acknowledgments

References

Aiello C, Agu E (2016) Investigating postural sway features, normalization and personalization in detecting blood alcohol levels of smartphone users. Proc Wirel Health 73–80.

Alessi SM, Barnett NP, Petry NM (2017) Experiences with SCRAMx alcohol monitoring technology in 100 alcohol treatment outpatients. Drug Alcohol Depend 178: 417–424. [PubMed: 28709081]

Alessi SM, Petry NM (2012) A randomized study of cellphone technology to reinforce alcohol abstinence in the natural environment. Addiction 108: 900–909.

Arnold Z, LaRose D, Agu E (2015) Smartphone inference of alcohol consumption levels from gait. IEEE ICHI 417–426.

Bae S, Chung T, Ferreira D, Dey AK, Suffoletto B (2018) Mobile phone sensors and supervised machine learning to identify alcohol use events in young adults: Implications for just-in-time adaptive interventions. Addict Behav 83: 42–47. [PubMed: 29217132]

Bae S, Ferreira D, Suffoletto B, Puyana JC, Kurtz R, Chung T, Dey AK. (2017) Detecting drinking episodes in young adults using smartphone-based sensors. PACM Interact Mob Wearable Ubiquitous Technol 1: Article 5

Barnett NP, Celio MA, Tidey JW, Murphy JG, Colby SM, Swift RM (2017). A preliminary randomized controlled trial of contingency management for alcohol use reduction using a transdermal alcohol sensor. Addiction 112: 1025–1035. [PubMed: 28107772]

Barnett NP, Meade EB, Glynn TR (2014) Predictors of detection of alcohol use episodes using a transdermal alcohol sensor. Exp Clin Psychopharmacol 22: 86–96. [PubMed: 24490713]

Barnett NP, Souza T, Rosen IG, Luczak SE, Glynn TR, Swift R (2015). Transdermal Alcohol Sensor Data Macro (Version 1.3) [software]. Brown University.

Barnett NP, Tidey J, Murphy JG, Swift R, Colby SM (2011) Contingency management for alcohol use reduction. A pilot study using a transdermal alcohol sensor. Drug Alcohol Depend 118: 391–399. [PubMed: 21665385]

Barnett NP, Wei J, Czachowski C (2009) Measured alcohol content in college party mixed drinks. Psychol Add Behav 23: 152–156.

Beckjord E, Shiffman S (2014) Background for real-time monitoring and intervention related to alcohol use. Alcohol Res 36: 9–18. [PubMed: 26258996]

Bertz JW, Epstein DH, Preston KL (2018) Combining ecological momentary assessment with objective, ambulatory measures of behavior and physiology in substance use research. Addict Beh 83: 5–17.

Black AC, Cooney NL, Justice AC, Fiellen LE, Pietrzak RH, Lazar CM, Rosen MI (2016) Momentary assessment of PTSD symptoms and sexual risk behavior in male OEF/OIF/OND Veterans. J Affect Disord 190: 424–428. [PubMed: 26551400]

Bolger N, Laurenceau J-P (2013) Intensive longitudinal methods: An introduction to diary and experience sampling research. New York: Guilford.

Bond JC, Greenfield TK, Patterson D, Kerr WC (2014) Adjustments for drink size and ethanol content: New results from a self-report diary study and transdermal sensor validation study. Alcohol Clin Exp Res 38:3060–6037. [PubMed: 25581661]

Boyd SJ, Corbin WR (2018) Faster alcohol metabolism is associated with increased stimulation and within session consumption. Exp Clin Psychopharm 26: 168–176.

Carpenter RW, Trela CJ, Lane SP, Wood PK, Piasecki TM, & Trull TJ (2017). Elevated rate of alcohol consumption in borderline personality disorder patients in daily life. Psychopharmacol 234:3395–3406.

Clapp JD, Madden DR, Mooney DD, Dahlquist KE (2017) Examining the social ecology of a bar crawl: An exploratory pilot study. PLOS One 12: e0185238. [PubMed: 28953932]

Clapp JD, Min JW, Shillington AM, Reed MB, Lange JE, Holmes MR (2006). Environmental and individual predictors of error in field estimates of blood alcohol concentration: A multilevel analysis. J Stud Alcohol 67: 620–627. [PubMed: 16736083]

Collins RL, Kashdan TB, Gollnisch G (2003) The feasibility of using cellular phones to collect ecological momentary assessment data: Application to alcohol consumption. Exp Clin Psychopharmacol 11, 73–78. [PubMed: 12622345]

Collins RL, Mosheimer ET, Shiffman S, Paty JA, Gnys M, Papandonatos GD (1998) Ecological momentary assessment in a behavioral drinking moderation training program. Exp Clin Psychopharmacol 6: 306–315. [PubMed: 9725114]

Crooke AH, Reid SC, Kauer SD, McKenzie DP, Hearps SJ, Khor AS, Forbes AB. (2013) Temporal mood changes associated with different levels of adolescent drinking: Using mobile phones and experience sampling methods to explore motivations for adolescent alcohol use. Drug Alcohol Rev 32: 262–268. [PubMed: 23432520]

Csikszentmihalyi M, Larson R (1987) Validity and reliability of the experience-sampling method. J Nerv Ment Dis 175: 526–536 [PubMed: 3655778]

Dougherty DM, Charles NE, Acheson A, John S, Furr RM, Hill-Kapturczak N (2012) Comparing the detection of transdermal and breath alcohol concentrations during periods of alcohol consumption ranging from moderate drinking to binge drinking. Exp Clin Psychopharmacol 20:373–381. [PubMed: 22708608]

Dougherty DM, Hill-Kapturczak N, Liang Y, Karns TE, Lake SL, Cates SE, Roache JD (2015) The potential clinical utility of transdermal alcohol monitoring data to estimate the number of alcoholic drinks consumed. Addict Disord Their Treat 14: 124–130. [PubMed: 26500459]

Dougherty DM, Lake SL, Hill-Kapturczak N, Liang Y, Karns TE, Mullen J, Roache JD (2015) Using contingency management practices to reduce at-risk drinking in heavy drinkers. Alcohol Clin Exp Res 39: 743–751. [PubMed: 25833033]

Dvorak RD, Pearson MR, Day AM (2014) Ecological momentary assessment of acute alcohol use disorder symptoms: Associations with mood, motives, and use on planned drinking days. Exp Clin Psychopharmacol 22: 285–297. [PubMed: 24932896]

Dvorak RD, Pearson MR, Sargent EM, Stevenson BL (2016). Daily associations between emotional functioning and alcohol involvement: Moderating effects of response inhibition and gender. Drug Alcohol Depend 163: S46–S53. [PubMed: 27306731]

Epler AJ, Tomko RL, Piasecki TM, Wood PK, Sher KJ, Shiffman S, Heath AC (2014) Does hangover influence the time to next drink? An investigation using ecological momentary assessment. Alcohol Clin Exp Res 38: 1461–1469. [PubMed: 24588377]

Epstein DH, Tyburski M, Craig IM, Phillips KA, Jobes ML, Vahabzadeh M, Mezghanni M, Lin J-L, Furr-Holden CDM, Preston KL (2014) Real-time tracking of neighborhood surroundings and mood in urban drug misusers: Application of a new method to study behavior in its geographical context. Drug Alcohol Depend 134: 22–29. [PubMed: 24332365]

Fatseas M, Serre F, Swendsen J, Auriacombe M (2018). Effects of anxiety and mood disorders on craving and substance use among patients with substance use disorder: An ecological momentary assessment study. Drug Alcohol Depend 187: 242–248. [PubMed: 29684892]

Ferguson SG, Shiffman S (2011) Using the methods of ecological momentary assessment in substance dependence research -- Smoking cessation as a case study. Subst Use Misuse 46: 87–95. [PubMed: 21190409]

Fairbairn CE, Rosen IG, Luczak SE, Venerable WJ (2018 Epub) Estimating the quantity and time course of alcohol consumption from transdermal alcohol sensor data: A combined laboratory-ambulatory study. Alcohol.

Fillmore MT, Weafer J (2012) Acute tolerance to alcohol in at-risk binge drinkers. Psychol Addict Behav 26: 693–702. [PubMed: 22023021]

Gautreau C, Sherry S, Battista S, Goldstein A, Stewart S (2015). Enhancement motives moderate the relationship between high-arousal positive moods and drinking quantity: Evidence from a 22-day experience sampling study. Drug Alcohol Rev 34: 595–602. [PubMed: 25735401]

Gharani P, Suffoletto B, Chung T, Karimi HA (2017) An artificial neural network for movement pattern analysis to estimate blood alcohol content level. Sensors 17: 2897.

Gorka SM, Hedeker D, Piasecki TM, Mermelstein R (2017) Impact of alcohol use motives and internalizing symptoms on mood changes in response to drinking: An ecological momentary assessment investigation. Drug Alcohol Depend 173: 31–38. [PubMed: 28189925]

Greenfield TK, Bond J, Kerr WC (2014) Biomonitoring for improving alcohol consumption surveys: The new gold standard? Alcohol Res 36: 39–45. [PubMed: 26258999]

Gurvich EM, Kenna GA, Leggio L (2013) Use of novel technology-based techniques to improve alcohol-related outcomes in clinical trials. Alcohol Alcohol 48: 712–719. [PubMed: 23955872]

Gustafson DH, McTavish FM, Chih M-Y, Atwood AK, Johnson RA, Boyle MG, Levy MS, Driscoll H, Chisholm SM, Dillenburg L, Isham A, Shah D (2014) A smartphone application to support recovery from alcoholism: A randomized clinical trial. JAMA Psychiatry 71: 566–572. [PubMed: 24671165]

Hill-Kapturczak N, Lake SL, Roache JD, Cates SE, Liang Y, Dougherty DM (2014) Do variable rates of alcohol drinking alter the ability to use transdermal alcohol monitors to estimate peak breath alcohol and total number of drinks? Alcohol Clin Exp Res 38: 2517–2522. [PubMed: 25335857]

Hill-Kapturczak N, Roache JD, Liang Y, Karns TE, Cates SE, Dougherty DM (2015) Accounting for sex-related differences in the estimation of breath alcohol concentrations using transdermal alcohol monitoring. Psychopharmacol 232: 225–123.

Hoeppner BB, Schick MR, Kelly LM, Hoeppner SS, Bergman B, Kelly JF (2017). There is an app for that – Or is there? A content analysis of publicly available smartphone apps for managing alcohol use. J Subst Abuse Treat 82: 67–73. [PubMed: 29021117]

Holt LJ, Litt MD, Cooney NL (2012) Predictive analysis of early lapse to drinking and smoking among individuals in concurrent alcohol and tobacco treatment. Psychol Addict Behav 26: 561–572. [PubMed: 22023022]

Hustad JTP, Carey KB (2005) Using calculations to estimate blood alcohol concentrations for naturally occurring drinking episodes: A validity study. J Stud Alcohol 66: 130–138. [PubMed: 15830913]

Jones A, Tiplady B, Houben K, Nederkoorn C, Field M (2018) Do daily fluctuations in inhibitory control predict alcohol consumption? Psychopharmacol 235: 1487–1496.

Karns-Wright TE, Dougherty DM, Hill-Kapturczak N, Mathias CW, Roache JD (2018) The correspondence between transdermal alcohol monitoring and daily self-reported alcohol consumption. Addict Behav 85: 147–152. [PubMed: 29910035]

Kashdan TB, Ferssizidis P, Collins RL, Muraven M (2010) Emotion differentiation as resilience against excessive alcohol use: An ecological momentary assessment in underage social drinkers. Psychol Sci 12: 1341–1347.

Kerr WC, Greenfield TK, Tujage J, Brown SE (2005) A drink is a drink? Variation in the amount of alcohol contained in beer, wine and spirits drinks in a US methodological sample. Alcohol Clin Exp Res 29: 2015 – 2021. [PubMed: 16340459]

Kraus CL, Salazar NC, Mitchell JR, Florin WD, Guenther B, Brady D, Swartzwelder SH, White AM (2005) Inconsistencies between actual and estimated blood alcohol concentrations in a field study

of college students: Do students really know how much they drink? Alcohol Clin Exp Res, 29: 1672–1676. [PubMed: 16205367]

Lane SP, Carpenter RW, Sher KJ, Trull TJ (2016) Alcohol craving and consumption in borderline personality disorder: When, where, and with whom Clin Psychol Sci 4: 775–792. [PubMed: 28042520]

Leffingwell TR, Cooney NJ, Murphy JG, Luczak S, Rosen G, Dougherty DM, Barnett NP (2013) Continuous objective monitoring of alcohol use: Twenty-first century measurement using transdermal sensors. Alcohol Clin Exp Res 37: 16–22. [PubMed: 22823467]

Luczak SE, Hawkins AL, Dai Z, Wichmann R, Wang C, Rosen IG (2018) Obtaining continuous BrAC/BAC estimates in the field: A hybrid system integrating transdermal alcohol biosensor, Intellidrink smartphone app, and BrAC Estimator software tools. Addict Behav 83: 48–55. [PubMed: 29233567]

Luczak SE, Rosen IG (2014) Estimating BrAC from transdermal alcohol concentration data using the BrAC Estimator software program. Alcohol Clin Exp Res 38: 2243–2252. [PubMed: 25156615]

Luczak SE, Rosen IG, Wall TL (2015) Development of a real-time repeated-measures assessment protocol to capture change over the course of a drinking episode. Alcohol Alcohol 50: 180–187. [PubMed: 25568142]

Mehl MR Connor TS. (2012) Handbook of Research Methods for Studying Daily Life. Guilford Press, New York.

Maisto SA, Kirouac M, Witkiewitz K (2014) Alcohol use disorder clinical course research: Informing clinicians' treatment planning now and in the future. J Stud Alcohol Drugs 75: 799–807. [PubMed: 25208198]

Marques PR, McKnight AS (2009) Field and laboratory alcohol detection with 2 types of transdermal devices. Alcohol Clin Exp Res 33: 703–711. [PubMed: 19170663]

Martin CS, Earleywine M, Musty RE, Perrine MW, & Swift RM (1993) Development and validation of the Biphasic Alcohol Effects Scale. Alcohol Clin Exp Res 17: 140–146. [PubMed: 8452195]

Mathias CW, Hill-Kapturczak N, Karns-Wright TE, Mullen J, Roache JD, Fell JC, Dougherty DM (2018). Translating transdermal alcohol monitoring procedures for contingency management among adults recently arrested for DWI. Addict Behav 83: 56–63 [PubMed: 29397211]

Matthews DB, Miller WR (1979) Estimating blood alcohol concentration: Two computer programs and their applications in therapy and research. Addict Behav 4:55–60. [PubMed: 420046]

Mayhugh RE, Rejeski WJ, Petrie MR, Laurienti PJ, Gauvin L (2018) Differing patterns of stress and craving across the day in moderate-heavy alcohol consumers during their typical drinking routine and an imposed period of alcohol abstinence. PLOS One 13: e0195063. [PubMed: 29668736]

McAfee A, Watson J, Bianchi B, Aiello C. Agu E (2017) AlcoWear: Detecting blood alcohol levels from wearables. In: IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (pp. 1–8). IEEE.

Miranda R, Jr., MacKillop J, Treloar H, Blanchard A, Tidey JW, Swift RM, Chun T, Rohsenow DJ, Monti PM (2016) Biobehavioral mechanisms of topiramate's effects on alcohol use: An investigation pairing laboratory and ecological momentary assessments. Addict Biol 21: 171–182. [PubMed: 25353306]

Miranda R, Jr., Monti PM, Ray LR, Treloar HR, Reynolds EK, Ramirez J, Chun T, Gwaltney CJ, Justus A, Tidey J, Blanchard A, Magill M (2014b) Characterizing subjective responses to alcohol among adolescent problem drinkers. J Abnorm Psychol 123: 117–129. [PubMed: 24661164]

Miranda R, Jr., Padovano HT, Gray JC, Wemm SE, Blanchard A (2018). Real-time assessment of alcohol craving and naltrexone treatment responsiveness in a randomized clinical trial. Addict Behav 83: 72–78. [PubMed: 29395188]

Miranda R, Ray L, Blanchard A, Reynolds EK, Monti PM, Chun T, Justus A, Swift RM, Tidey J, Gwaltney CJ, Ramirez J (2014a) Effects of naltrexone on adolescent alcohol cue reactivity and sensitivity: An initial randomized trial. Addict Biol 19: 941–954. [PubMed: 23489253]

Morean ME, Corbin WR, Treat TA (2013) The Subjective Effects of Alcohol Scale: Developing a psychometric evaluation of a novel assessment tool for measuring subjective response to alcohol. Psychol Assess 25: 780–795. [PubMed: 23647036]

Morgenstern J, Kuerbis A, Houser J, Muench FJ, Shao S, Treloar H (2016) Within-person associations between daily motivation and self-efficacy and drinking among problem drinkers in treatment. Psychol Addict Behav 30: 630–638. [PubMed: 27560995]

Muraven M, Collins RL, Morsheimer ET, Shiffman S, Paty JA (2005) One too many: Predicting future alcohol consumption following heavy drinking. Exp Clin Psychopharmacol 13: 127–136. [PubMed: 15943545]

Myin-Germeys I, Oorschot M, Collip D, Lataster J, Delsepaul P, van Os J (2009) Experience sampling research in psychopathology: Opening the black box of daily life. Psychol Med 39: 1533–1547. [PubMed: 19215626]

Neville FG, Williams DJ, Goodall CA, Murer JS, Donnelly PD (2013) An experimental trial exploring the impact of continuous transdermal alcohol monitoring upon alchol consumption in a cohort of male students. PLOS One 8: e67386. [PubMed: 23825656]

Paty J, Kassel J, Shiffman S. (1992). The importance of assessing base rates of clinical studies: An example of stimulus control of smoking, in The Experience of Psychopathology: Investigating Mental Disorders in Their Natural Settings (deVries MW ed), pp 347–352. New York: Cambridge University Press.

Piasecki TM, Alley KJ, Slutske WS, Wood PK, Sher KJ, Shiffman S, Heath AC (2012a) Low sensitivity to alcohol: Relations with hangover occurrence and susceptibility in an ecological momentary assessment investigation. J Stud Alc Drugs 73: 925–932.

Piasecki TM, Cooper ML, Wood PK, Sher KJ, Shiffman S, Heath AC (2014) Dispositional drinking motives: Associations with appraised alcohol effects and alcohol consumption in an ecological momentary assessment investigation. Psychol Assess 26: 363–369. [PubMed: 24274049]

Piasecki TM, Hufford MR, Solhan M, Trull TJ (2007) Assessing clients in their natural environments with electronic diaries: Rationale, benefits, limitations, and barriers. Psychol Assess 19: 25–43. [PubMed: 17371121]

Piasecki TM, Jahng S, Wood PK, Robertson BM, Epler AJ, Cronk NJ, Rohrbaugh JW, Heath AC, Shiffman S, Sher KJ. (2011). The subjective effects of alcohol-tobacco co-use: An ecological momentary assessment investigation. J Abnorm Psychol 120: 557–571. [PubMed: 21443289]

Piasecki TM, Wood PK, Shiffman S, Sher KJ, Heath AC (2012b) Responses to alcohol and cigarette use during ecologically assessed drinking episodes. Psychopharmacol 222: 331–344.

Ramirez J, Miranda R, Jr. (2014) Alcohol craving in adolescents: Bridging the laboratory and natural environment. Psychopharmacol 231: 1841–1851.

Ray LA, Miranda R, Jr., Tidey J, McGeary JE, MacKillop J, Gwaltney C, Rohsenow DJ, Swift RM, Monti PM. (2010). Polymorphisms of the μ-opioid receptor and dopamine D4 receptor genes and subjective responses to alcohol in the natural environment. J Abnorm Psychol 119: 115–125. [PubMed: 20141248]

Riordan BC, Scarf D, Moradi S, Flett JAM, Carey KB, Conner TS (2017) The accuracy and promise of personal breathalysers for research: Steps toward a cost-effective reliable measure of alcohol intoxication? Digit Health 3: 1–5.

Roache JD, Karns TE, Hill-Kapturczak N, Mullen J, Liang Y, Lamb RJ, Dougherty DM (2015). Using transdermal alcohol monitoring to detect low-level drinking. Alcohol Clin Exp Res 39: 1120–1127. [PubMed: 25988708]

Roache JD, Karns-Wright TE, Goros M, Hill-Kapturczak N, Mathias CW, Dougherty DM (2018 Epub) Processing transdermal alcohol concentration (TAC) data to detect low-level drinking. Alcohol.

Rosen IG, Luczak SE, Weiss J (2014) Blind deconvolution for distributed parameter systems with unbounded input and output and determining blood alcohol concentration from transdermal biosensor data. Applied Math Comput 231: 357–376.

Sakai JT, Mikulich-Gilbertson SK, Long RJ, Crowley TJ (2006) Validity of transdermal alcohol monitoring: Fixed and self-regulated dosing. Alcohol Clin Exp Res 30: 26–33. [PubMed: 16433729]

Shiffman S (2009) Ecological momentary assessment (EMA) in studies of substance use. Psychol Assess 21:486–497. [PubMed: 19947783]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Shiffman S (2014).Conceptualizing analyses of ecological momentary assessment data. Nicotine Tob Res 16 (Suppl. 2): S76–S87. [PubMed: 24323571]

Shiffman S, Stone AA, Hufford MR (2008) Ecological momentary assessment. Annu Rev Clin Psychol 4:1–32. [PubMed: 18509902]

Shrout PE, Lane SP (2012) Psychometrics, in Handbook of Research Methods for Studying Daily Life. (Mehl MR, Connor TS eds), pp 302–320. Guilford Press, New York.

Simons JS, Wills TA, Emery NN, Marks RM (2015) Quantifying alcohol consumption: Self-report, transdermal assessment, and prediction of dependence symptoms. Addict Behav 50: 205–212. [PubMed: 26160523]

Sirlanci M, Rosen IG, Wall TL, Luczak SE (2018, Epub) Applying a novel population-based model approach to estimating breath alcohol concentration (BrAC) from transdermal alcohol concentration (TAC) biosensor data. Alcohol.

Suffoletto B, Kristan J, Callaway C, Kim KH, Chung T, Monti PM, Clark DB (2014) A text message alcohol intervention for young adult emergency department patients: A randomized clinical trial. Ann Emerg Med 64: 664–672. [PubMed: 25017822]

Suffoletto B, Gharani P, Chung T, Karimi H (2018) Using phone sensors and an artificial neural network to detect gait changes during drinking episodes in the natural environment. Gait Posture 60: 116–121. [PubMed: 29179052]

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

Swift R (2000) Transdermal alcohol measurement for estimation of blood alcohol concentration. Alcohol Clin Exp Res 24: 422–423. [PubMed: 10798575]

Swift R (2003) Direct measurement of alcohol and its metabolites. Addiction 98 (Suppl. 2) 73–80. [PubMed: 14984244]

Tidey JW, Monti PM, Rohsenow DJ, Gwaltney CJ, Miranda R, Jr., McGeary JE, MacKillop J, Swift RM, Abrams DB, Shiffman S, Paty JA (2008) Moderators of naltrexone's effects on drinking, urge, and alcohol effects in non-treatment seeking heavy drinkers in the natural environment. Alcohol Clin Exp Res 32: 58–66. [PubMed: 18028530]

Tomko RL, Saladin ME, McClure EA, Squeglia LM, Carpenter MJ, Tiffany ST, Baker NL, Gray KM (2017) Alcohol consumption as a predictor of reactivity to smoking and stress cues presented in the natural environments of smokers. Psychopharmacol 234: 427–435.

Trela CJ, Piasecki TM, Bartholow BD, Heath AC, & Sher KJ (2016). The natural expression of individual differences in self-reported level of response to alcohol during ecologically-assessed drinking episodes. Psychopharmacol 233: 2185–2195.

Treloar H, Miranda R, Jr. (2017). Craving and acute effects of alcohol in youths' daily lives: Associations with alcohol use disorder severity. Exp Clin Psychopharmacol 25: 303–313. [PubMed: 28627928]

Trull TJ, Ebner-Priemer U (2014) The role of ambulatory assessment in psychological science. Curr Dir Psychol Sci 23: 466–470. [PubMed: 25530686]

Trull TJ, Wycoff AM, Lane SP, Carpenter RW, Brown WC (2016) Cannabis and alcohol use, affect, and impulsivity in psychiatric out-patients' daily lives. Addiction 111: 2052–2059. [PubMed: 27270874]

Voogt C, Kuntsche E, Kleinjan M, Poelen E, Engels R (2014). Using ecological momentary assessment to test the effectiveness of a web-based brief alcohol intervention over time among heavy-drinking students: Randomized controlled trial. JMIR 16: e5. [PubMed: 24401555]

Walls TA, Schafer JL (2006) Models for intensive longitudinal data. New York: Oxford University Press.

Wang Y, Fridberg DJ, Leeman RF, Cook RL, Porges EC (2018 Epub) Wrist-worn alcohol biosensors: Strengths, limitations, and future directions. Alcohol.

Weaver ER, Horyniak DR, Jenkinson R, Dietze P, Lim MSC (2013) "Let's get Wasted!" and other apps: Characteristics, acceptability, and use of alcohol-related smartphone apps. JMIR Mhealth Uhealth 1: e9. [PubMed: 25100681]

White AM, Kraus CL, Flom JD, Kestenbaum LA, Mitchell JR, Shah K, Swartzwelder HS (2005) College students lack knowledge of standard drink volumes: Implications for definitions of risk drinking based on survey data. Alcohol Clin Exp Res 29: 631–638. [PubMed: 15834229]

Wray TB, Merrill JE, Monti PM (2014) Using ecological momentary assessment (EMA) to assess situation-level predictors of alcohol use and alcohol-related consequences. Alcohol Res 36: 19–27. [PubMed: 26258997]
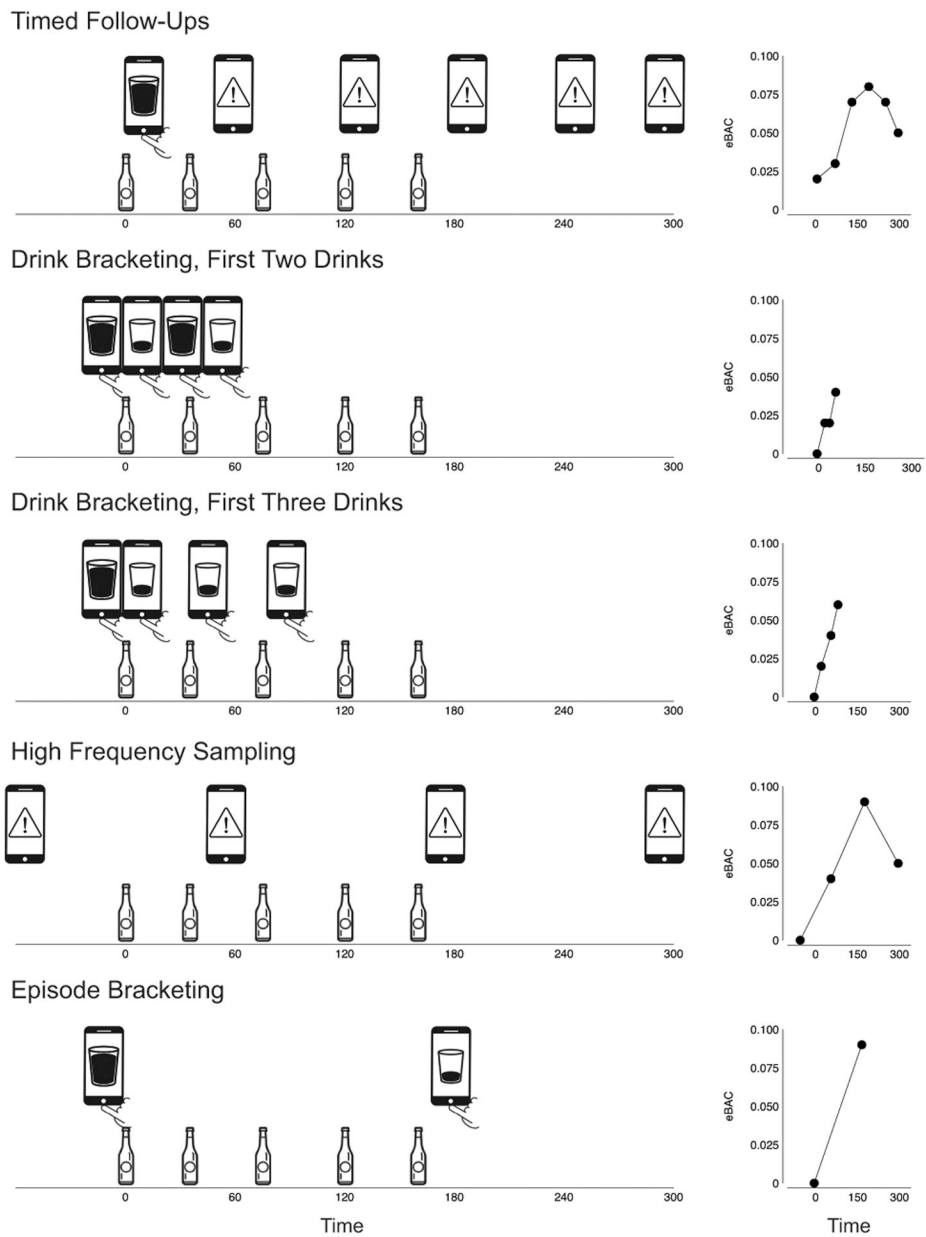
Figure 1.

Schematics (left) depicting how various high-resolution EMA designs from the existing literature might apply to a single hypothetical drinking episode involving consumption of five drinks. In this *Timed Follow-Ups* example, a user-initiated first drink report (depicted as phone with a drink icon) is logged shortly after completing the first drink, triggering prompted follow-up assessments (alert phone icon) at 60 minute intervals. The sequence ceases after no new drinks are reported in two consecutive follow-ups. In the *Drink Bracketing, First Two Drinks* example, the user logs begin-drink reports (depicted as full drink phone icon) just before each of the first two drinks in the episode and end-drink reports (empty drink phone icon) after each the first two drinks. In the *Drink Bracketing, First Three Drinks* example, the user logs a begin-drink report prior to the first drink and

end-drink reports after each of the first three drinks. In the *High Frequency Sampling* example, prompted assessments are collected at two-hour intervals and each assessment asks the user to report the number of drinks consumed since the prior report. In the *Episode Bracketing* example, the user logs a begin-drink report prior to consuming the first drink and an end-drink report after completing the drinking episode. Estimated BAC points (right) are plotted for each EMA assessment example. The eBAC data points are plotted for each time point where drinking information and other self-report data are collected by the corresponding assessment scheme. Estimates were generated using the formula of Matthews and Miller (1979) and assuming the drinker was a 150 lb. female. Note that, when applied to a single drinking episode, the different EMA protocols produce varying numbers of data points, cover different time intervals and the portions of the eBAC range, and vary in the extent to which they capture moments associated with falling eBAC.