

JOSHUA M. SMYTH and ARTHUR A. STONE

ECOLOGICAL MOMENTARY ASSESSMENT RESEARCH IN BEHAVIORAL MEDICINE

ABSTRACT. Behavioral medicine research often seeks to examine relatively complex theoretical models and clinical processes. Traditional designs, based on retrospective self-report in the natural environment or laboratory-based studies, cannot assess complex and temporally dynamic psychological, behavioral, and physiological processes in the natural environment. An alternate research approach, Ecological Momentary Assessment (EMA), provides a method to collect such data. This research approach and these data allow investigation of a number of research questions not amenable to other data collection methods. Examples of research questions uniquely suited for EMA data include, but are not limited to: (1) the comparison of retrospective data to momentary data; (2) evaluation of physiological and psychosocial processes in the natural environment; (3) the relation of symptoms and other disease processes to characteristics of the natural environment; and (4) examination of dynamic changes in psychosocial variables that may mediate improvement following treatment. Despite the strengths of EMA data, a number of challenges exist – including logistic, analytic, and interpretation problems – to the easy implementation of EMA approaches in behavioral medicine.

KEY WORDS: EMA, behavioral medicine, health psychology, research, methods, ambulatory.

Behavioral medicine is a discipline that is committed to understanding the relationship between social, psychological, and behavioral factors and health (i.e., the biopsychosocial model). It is closely related to other disciplines such as health psychology (which has a more psychological focus) and physiological psychology (which is more concerned with physiological mechanisms). Although behavioral medicine is a relatively young field (it was formally defined in 1978), it is especially important for the health of the nation given increasing emphasis of role of behavior in the genesis of disease and in the maintenance of health. The National Institutes of Health have recently emphasized the importance of behavioral medicine, broadly defined, through the establishment of a new office to stimulate additional research into biobehavioral influences on health and illness (Office of Behavioral and Social Science Research, established July 1, 1995).

The success of a scientific field such as behavioral medicine will depend, at least in part, on the study designs and assessments available to it. Our purpose in this paper is to discuss the rationale for the use of momentary data collection techniques in behavioral medicine and



to present several examples of research that has used these methods. We argue that momentary techniques will in many instances provide behavioral medicine researchers with powerful ways of addressing their research questions.

By its nature, behavioral medicine research is integrative and typically involves a wide array of outcome types (e.g., symptoms, behavior, mood, physiological function), a range of participant samples (e.g., individuals with chronic illness), and models of varying theoretical complexity. These attributes contribute to the need for sophisticated research approaches, ones that are able to simultaneously consider the complex patterns that are likely to exist among the outcomes. By way of example, a researcher might be interested in whether patients sharing a characteristic or trait (e.g., neuroticism) who are newly diagnosed with a serious medical condition experience increased day to day stress and/or altered coping responses that subsequently influence physiological processes and risk for the onset of upper respiratory infection. The example is notable in its complexity and the need to understand interactions over time.

We believe that the momentary techniques discussed in this special issue may be particularly useful for behavioral medicine researchers given the multi-disciplinary approach inherent in their work. Of course, one could make similar arguments for the utility of the momentary approach for other disciplines, and we do not in any way mean to limit the applicability of EMA approaches to this field.

RATIONALE FOR USING MOMENTARY DATA COLLECTION

We will discuss several reasons for the use of momentary data collection techniques and the designs that are associated with them in the following sections. As noted above, investigating complex phenomena require complex, time-sensitive designs. EMA approaches allow one to collect data that can answer many questions not amenable to data collected through more conventional approaches. We approach this by first discussing the rationale for using EMA for capturing self-report data, which is essentially a review of the definition of EMA. We then move to the study designs that are enabled by the use of this data capture technique and show how they may be helpful for the behavioral medicine researcher.

Characteristics of EMA

This data capture technique and its predecessors (the Experience Sampling Method or ESM; e.g., Csikszentmihalyi, 1994; DeVries 1992) were developed with multiple purposes in mind. Perhaps foremost was the concern that retrospective recall of self-reported experiences were fallible and that a variety of factors bias the recall of autobiographical memory (Schwartz and Sudman, 1994). Cognitive and memory limitations give rise to a variety of heuristics to conserve cognitive resources during recall that can systematically bias self-report. Although we have discussed these issues at length elsewhere (e.g., Gorin and Stone, 2001; Stone et al., 2000) several of the primary issues are noted below.

The recall of a past event tends to be in accordance with the individual's beliefs about behavior or the way the world functions ("effort after meaning"). The recall of an event may also be influenced by the outcome of the event. That is, recall can be biased by other events occurring after the event to be recalled ("retroactive reconstruction"). Similarly, recall is reliably influenced by one's current state, particularly mood. For example, emotionally valenced events are more easily recalled when the individual is experiencing a congruent current mood state. Longer time periods of recall, particularly recall in research protocols that require people to summarize events or experiential states, are particularly sensitive to recall heuristics. For instance, when asked to report on behaviors or mood states over the preceding week or month, recall will be heavily influenced by more salient events during that time interval. That is, those events or mood states that more readily come to mind, typically the peak or most recent experience, will bias (move) the retrospective self-report towards them. These processes inherent to self-report, particularly summary reports over longer periods of time, may thus give us a biased view of individuals' behavior and psychological states.

The second reason for developing EMA concerns the question of the ecological validity or generalizability of findings based on research conducted solely or in part in the laboratory. The laboratory environment is fundamentally dissimilar to peoples' natural environment. Although research conducted in the controlled laboratory setting offers the important benefit of experimental control, it is much less clear if processes (or the relationships between variables) observed in the laboratory environment are similar to what occurs in the "real" world.

As behavioral medicine researchers often investigate processes that are sensitive to transient or rapidly changing states (e.g., mood or blood pressure changes), issues of external validity become more central. It also may be the case that it is difficult or unethical to produce relevant experiential changes in the laboratory, although such changes can be captured in the natural environment of a carefully selected sample of participants (e.g., patients with a severe illness, individuals with serious marital discord, etc.).

One representation of ecological validity concerns in behavioral medicine is the “white-coat hypertension” (e.g., Pickering and Friedman, 1991). It is based on the observation that some individuals have higher blood pressure readings when taken in the laboratory or medical setting compared with those taken in a natural environment. It suggests that even when we are sufficiently sophisticated to replicate, for example, stressful interpersonal interactions in the laboratory, participants’ behaviors or psychophysiological responses may not be representative of their natural interpersonal behavior. In addition to differences in literal environmental cues, laboratory research rarely, if ever, includes aspects of the social environment. Furthermore, the research laboratory environment and all it entails (e.g., making a special trip to a medical center or university) may systematically alter or constrain affective and/or behavioral aspects of research participants.

The third reason for developing EMA was to enable a closer examination of the dynamic processes that surely are present in most biopsychosocial phenomena. One way to think about this is that cross-sectional studies provide a static snapshot of processes and that prospective studies provide a few snapshots. On the other hand, because EMA designs usually incorporate a large number of repeated measures, they provide a dynamic, movie-like view of processes, allowing examination of the temporal patterning of changes. Such a view is useful for imputing, though not confirming, causal associations among variables given the ability to know the ordering of changes. In other words, to know that variable *X* changed prior to changes in variable *Y* can be tremendously useful.

An EMA Approach

Standard research approaches struggle to encompass and represent this rich set of associations expected in many behavioral medicine study designs. Cross-sectional designs that relate a presumed antecedent

condition (e.g., diagnosis of disease) and its consequences (e.g., altered stress and coping responses) do not clearly demonstrate the temporal relationship between these variables. Such temporal sequencing is, of course, necessary to empirically test antecedent conditions of the variable in question. Many studies exclusively rely on the participants' retrospective recall of certain states, events or conditions. For example, a participant may be asked how much stress she/he experienced following an interpersonal conflict, or to recall affect before, during, and after such a conflict to evaluate temporal processes. One common solution to problems presented by retrospective recall is to use laboratory paradigms and/or direct observation in the laboratory. There are, of course, great strengths to laboratory designs, particularly experimental (randomized) designs. Of particular relevance to this discussion, these studies typically avoid problems associated with retrospective report and can add the rigor of an experimental design, helping to clarify causal relationships. In spite of these strengths, as mentioned earlier the external validity of laboratory-based findings may be questionable.

Exploring physiological processes in the laboratory environment as an analogue to physiological processes that occur in "real life" may also present problems. Although it initially may seem plausible to form inferential conclusions about naturalistic reactions based on laboratory studies, there is reason to suspect this is not always the case. For example, a surprisingly weak correspondence between laboratory and field (naturalistic) endocrine reactivity has been demonstrated in several studies. There exists relatively little concordance between laboratory and real-life findings of plasma catecholamine responses (laboratory stress of mental arithmetic and real-life social speaking), urine catecholamine response (laboratory stressors of reaction time and tracking task, cold pressor test real-life defense of doctoral research; Dimsdale, 1984), cortisol responses (contrasting actual and dramatized lecture situations; Houtman and Bakker, 1987), or salivary cortisol responses (a laboratory-based stressful speech task and stressful events in subjects' daily lives; van Eck et al., 1996). Findings such as these suggest that we must take observations of physiological responses or patterns observed in the laboratory as accurate reflections of that environment, but with perhaps less ecological validity than we might initially suppose. As such, examination and replication of these processes using EMA approaches may yield additional insight (a process we articulate further below, based on experiences from our own work).

Recall-based approaches, and often laboratory experiments as well, typically do not account for variations associated with time of day. We have previously demonstrated that there are diurnal patterns in mood; that is, fluctuations in mood predicted by time of the day (Stone et al., 1996). Without the intensive assessment of mood offered by EMA approaches, diurnal patterns will be obscured at best, and contribute to error or bias at worst (e.g., pre and post mood assessments conducted at different times of day). Diurnal patterns are also of interest because there are many physiological processes that have clear diurnal patterns, and it is plausible that mood and/or physiological patterns over the day are related to disease function or behavior. There is also evidence that social occurrences that are entrained to time of day are determinants of mood and/or physiological patterns (e.g., meals; Smyth et al., 1998; Stone et al., 1996). It is also the case that solitary activities, events, or settings that individuals experience can be potent predictors of mood, behavior, symptoms, and physiological parameters (e.g., driving in the car; Smyth et al., 1999a; Stone et al., 1996). Accordingly, it may be necessary in behavioral medicine research to examine not only naturally occurring fluctuations in activities, location, behavior, etc., but the influence of time of day as well. It again bears repeating that laboratory-based research will, by its very nature, lose this important information by removing participants from their natural environments. Naturalistic studies based on cross-sectional and/or retrospective methods, on the other hand, fail to capture the dynamic temporal flow of key variables, and may be biased through a reliance on retrospective report.

A variety of research methods and approaches that allow the assessment of individual's ongoing experiences in the natural environment have been developed. Such approaches have historically been used only infrequently, in large part due to the complexities of collecting and analyzing such data. A new data assessment protocol, typically referred to as Ecological Momentary Assessment (EMA; Stone and Shiffman, 1994) allows the study of behavior, psychological states, and physiological reactions in their natural context. EMA-like research approaches have been used to overcome the limitations of cross-sectional and panel designs in studies examining, although not limited to, stress and coping (Bolger and Zuckerman, 1995), personality traits and negative affect (Bolger, 1990), depression (Stader and Hokanson, 1998), substance abuse (Litt, 1998), cardiovascular disease (Kamarck, et al., 2002), mental illness (deVries, 1992), cigarette smoking (Shiffman et al., 1996), chronic pain (Stone et al., 1997), and asthma (Smyth

et al., 1999a). Such a methodology allows the delineation of antecedent stress conditions (e.g., interpersonal events), moderating or mediating variables (e.g., coping responses), and outcome measures (e.g., blood pressure) much more effectively than panel or retrospective designs, in which the time between assessments can be lengthy (i.e., weeks, months, or even years).

Implementation of EMA

EMA involves signaling participants, typically multiple times each day, for a period of days or weeks to report on current or recent psychological states, behaviors, and/or environmental conditions (see other papers in this issue for details). The signaling typically occurs through the use of a small electronic device such as paging devices, programable wrist-watches, and palmtop computers. Assessments can focus on immediate affective, cognitive, or behavioral experiences (e.g., events in the last 5 min; what is happening right now), or they can solicit recollections of events that occurred over longer time frames (e.g., in the last 30 min; since the previous assessment). This approach dramatically reduces the period of recall, from days/weeks/months to minutes/hours. As a result, biases associated with retrospective recall are reduced, and the need for participants to summarize events and/or experiences over extended periods of time is removed (Stone and Shiffman, 1994). Finally, EMA also allows observation of the natural range of behaviors and moods as they normally occur in participants' natural environments and in response to naturally occurring events.

EXAMPLES OF EMA IN BEHAVIORAL MEDICINE RESEARCH

For the remainder of the paper, we discuss several specific studies we have conducted that draw on or explicate the unique contributions of EMA to behavioral medicine research. More specifically, we will highlight (1) the relation of (retrospective) self-report of coping versus coping in the natural environment, (2) the evaluation of physiological processes in the natural environment, (3) the relation of symptoms (e.g., illness-specific symptoms) to characteristics of the natural environment, (4) evaluating compliance rates and accuracy from different reporting formats (using paper and pencil diaries and more technologically sophisticated EMA approaches in chronic pain patients), and (5) EMA approaches to studying mediation in behavioral medicine interventions.

Coping With Stressful Daily Events

A line of work by Stone and his colleagues has taken an EMA approach to the study of psychological coping with stressful events. The earliest quantitative assessments of coping, which are still in wide use today, depend on the accurate recall of the coping responses to difficult events. The Ways of Coping Questionnaire (WOC; Folkman and Lazarus, 1980) asks participants to recall the most stressful event of the past month and then to respond to checklist of thoughts and actions that were used to cope with the event. Recall problems with the WOC have been recognized for many years (Stone et al., 1991) and EOD diary coping assessments have been developed (Stone and Neale, 1984). Nevertheless, considering that many of the coping items, especially the thoughts (such as “Saw a silver lining”), could be mis-recalled even at the end of the day (e.g., as a result of effort after meaning processes), we studied coping with momentary assessment techniques.

We conducted two major studies with individuals who reported employment or marital stress and who said that they experienced stressful events associated with those problems several times week. Our general protocol was to study individuals for several days by randomly signaling them to report on stressful problems and by asking them to activate the electronic diary in response to the occurrence of stressful events (known as “event-driven” actions). In the more methodological of the two studies, we compared the reports of coping with a particular event that were made at the time of the event (recorded with EMA via the electronic diary) with coping reports made 1–2 days later. To summarize the findings, the comparison of the momentary and the recalled coping indicated that much of what was reported in real time was forgotten at recall and that coping efforts that were not reported in real time were reported at recall (Stone et al., 1998). These findings throw into question the validity of recalled coping inventories. Additional analyses of these data yielded findings concerning how individuals’ reports of their typical or usual coping relate to their real time views of coping efforts. Our findings were quite striking and, in our opinion, challenge our past understanding of coping as viewed from a recall perspective. That is, there was little correspondence between the recall of usual coping and aggregates of momentary coping efforts (Schwartz et al., 1999). Taken together, these findings have provided another view of psychological coping and may enhance our understanding of the dynamic processes at work here. Furthermore, these studies highlight the possibility that

attempts to replicate recall-based findings utilizing EMA approaches may result in dramatically different results.

Assessing Physiological Processes in the Natural Environment

As discussed above, EMA techniques and designs are also useful for studying non-self-report phenomena, such as physiological data. Given their interest in mind–body connections, behavioral medicine researchers have for many years have focused on the hypothalamic–pituitary–adrenal (HPA) axis. First, it is known to be stress-responsive and it appears that interpersonal stressors are the most potent at activating the axis. Second, the major end-product of the HPA axis in humans, cortisol, has a host of biological functions including the regulation of metabolism, as an anti-inflammatory agent, and as an immuno-suppressive agent. These functions make cortisol of special interest, because they suggest that cortisol may be associated with both the control of existing diseases (especially those requiring metabolic, immune, or inflammatory surveillance) and the onset of infectious diseases.

Until relatively recently, explorations of the cortisol were conducted either by blood draw (giving a measure of cortisol in the blood at that moment) or analysis of urine samples (representing cortisol accumulating in the urine over long periods; often, 12- or 24-h collections were used). Advances in salivary assay techniques has made the measurement of cortisol possible with more convenient, ambulatory methods. From the patient's perspective, saliva is sampled by chewing on a cotton wad for a couple of minutes, followed by placing the wad into a special test tube. New assays enable detection of free cortisol in saliva that correlates remarkably well with blood assays (Kirschbaum and Hellhammer, 1994). These advances have set the stage for EMA designs including the assessment of cortisol. In consideration of the strong diurnal cycle of cortisol present in most individuals (highest in the morning followed by a steep decline through the day), it is important to seriously consider the sampling of saliva in order to adequately capture diurnal variation. We recommend using a stratified, random sampling scheme that ensures sampling from all sections of the diurnal cycle. It is important to note that the diurnal cycle of cortisol is not discernable from urine (which reflects aggregate levels) or from a single blood draw. Rather, it is necessary to get continuous measures (e.g., indwelling intravenous collection of blood) or multiple assessments to

assess diurnal variability. As it is generally not feasible to have continuous or multiple blood draws in the natural environment (at least not without large reactive effects), salivary-based approaches to the assessment of cortisol appear better suited for EMA studies.

We have used EMA combining both self-report assessment and cortisol assessment, measured with salivary assays, in order to examine the association between naturally occurring chronic and acute psychological stress and HPA axis activation. In one study, we examined a sample of community dwelling individuals and classified them according to their general level of stress. Our results showed that stressed people had steeper diurnal slopes of cortisol than those with less stress (Ockenfels et al., 1995). We also showed that momentary instances of stress throughout the day were associated with higher levels of cortisol compared to times where stress was not reported (Smyth et al., 1998); these results replicated similar findings from other researchers (Nicolson et al., 1991). Again, the overarching issue is that it is feasible to study physiological processes in the natural environment. Furthermore, EMA approaches allow investigation of research questions that more traditional approaches could not address (e.g., examination of diurnal cycles, momentary stress reactivity, etc.).

Variability of Asthma Symptoms

We have also examined the relation of the natural environment to variability in chronic diseases, such as asthma. One study (Smyth et al., 1999a) examined diurnal cycles of both lung function and asthma symptoms in their natural environment, and additionally examined the contribution of the environment to the observed diurnal variability of asthma symptoms. The relationship between psychosocial factors such as activities, locations, social contact, stressors, and mood and lung function (peak expiratory flow rate, PEFR) and asthma symptoms was investigated. Several sets of psychosocial variables assessed via EMA (activities, location, social contacts, mood and stress, medication and drug use) significantly predicted PEFR and asthma symptoms. Examination of the explained variance of each predictor set, however, indicated that certain predictor sets explained more variance than others. Mood and stressors explained more variance than any other predictor set in PEFR, while social contacts explained more variance than the other predictor sets in asthma symptoms. These data suggest that care must be taken to select the theoretically-guided and appropriate psychosocial variables in EMA research.

We also examined the well-established diurnal cycle of asthma symptoms and PEFr. The EMA results obtained in the natural environment replicated the diurnal cycles of PEFr and asthma symptoms typically found in laboratory research. PEFr was strongly related to time of day, with highest values (best lung function) in the afternoon. Asthma symptoms were also related to time of day, with fewer symptoms in the afternoon. These diurnal cycles are typically attributed only to various physiological factors; we examined whether psychosocial variables may also contribute to this pattern. After controlling for psychosocial factors in the diurnal cycle of PEFr and asthma symptoms, the effects of time of day were no longer significant. This finding suggests that psychosocial factors significantly impact diurnal variations in asthma symptomatology and provide more predictive utility than time of day alone. More generally, this study suggests that a careful EMA assessment of psychosocial factors (including activities, locations, social contacts, and mood and stressors) can provide important information about fluctuations in the severity of chronic physical illness.

Momentary Assessments of Pain in Chronic Pain Patients

Our personal experience (and a great deal of research) tells us that pain can be a rapidly fluctuating phenomenon. Similar to some of the other constructs discussed above, recall questionnaires have been the predominant way that the pain experience has been assessed. Of course, there are very real practical reasons for this (time, expense). Nevertheless, some pain researchers have employed diaries for many years in an effort to capture the variability in the pain experience and in an effort to avoid the recall bias that has been associated with retrospective pain reports. To hone our ability to completely measure the pain experience, several research teams have gone beyond EOD diaries to momentary reports of immediate pain taken several times a day. We discuss one of more recent studies below.

The study was focused on a methodological issue that has been of concern to diary researchers for many years, one that has probably undermined the use of diaries in research and the information they yield. Many researchers and clinicians who have had their patients use diaries are familiar with the poor compliance associated with paper-and-pencil diaries. In fact, there is a name for the hoarding of diaries and their last minute completion before turning them in – “parking lot compliance,” so named, because diaries are often completed in the

parking lot prior to an appointment. Of course, deviation from a diary study protocol can have pernicious effects on the resulting data given that the memory biases the diary technique was supposed to reduce may be as present as in typical recall reports.

One of the issues tackled in this study was that there had been no way to know how compliant individuals were with the paper diaries. We developed an instrumented paper diary (IPD) that had all of the features of a typical paper diary, but was different in one critical respect: a very thin computer was embedded in the diary binder and a set of unobtrusive photo cells were attached to the binder's internal spine. Software developed for the computer enabled us to record when the diary was open (meaning that a person could be completing it) and closed (not completing it). The second issue addressed in the study concerned the level of compliance that could be achieved with an electronic diary that was enhanced with features to make it especially user-friendly (such as a alarm clock and an auditory reminder to complete the diary).

Eighty chronic pain patients participated in the study for 3 consecutive weeks. Each day of the study they were instructed to record their pain levels and the quality of their pain in a very brief question (with paper and pencil for the first group and on the electronic diary's small, touch-sensitive screen for the second group). In summary, we found that the paper diary group had a compliance rate of 11–19% and that there was evidence that the patients hoarded diaries (completed on days other than the one they were supposed to) about a third of the time. The electronic diary (which automatically date and time stamped all entries) had a compliance rate of 94% (Stone et al., 2002). Of course, we are not sure which of features of the electronic diary are responsible for the difference in compliance rates; for example, it might be the auditory reminder or the knowledge that patients were being monitored by the electronic diary, and we are currently investigating several possibilities.

The important message from this study is that compliance issues are very real for diary studies and that, at the very least, it is a major advantage, to be able to actually know the compliance being achieved in a study. The majority of new EMA approaches, such as the use of palm-top computers or electronic diaries, provide the benefit of each data entry being date and time stamped. This feature eliminates many of the compliance "faking" problems present in paper and pencil diaries, and further allows the examination of a variety of compliance issues (e.g., punctuality of entries, length of time required to complete them, and so forth).

EMA as a Tool to Understand How a Behavioral Medicine Intervention Works

The primary purpose of the study to be described was to determine if a relatively new psychological manipulation – structured writing about past traumatic events – could impact objective disease indicators in two chronic diseases, asthma and rheumatoid arthritis. Its experimental design was classical in that patients within each disease were randomly assigned to a control condition (writing about emotionally neutral topics) or an the experimental condition (writing about past traumas in an emotional way). Primary outcomes for the arthritis patients were blind ratings of disease activity by the treating physicians and for the asthma patients it was a standard measure of lung function (FEV_1) evaluated using spirometry. Results of the intervention 4 months later were statistically significant as well as clinically meaningful, demonstrating improvement in patients completing the structured emotional writing intervention over those completing emotionally neutral writing tasks (Smyth et al., 1999b).

Although the overall outcome of the study was promising and suggested that additional exploration of this technique is needed, the basic design did not provide any insight into the possible mechanisms underlying the effect. Of course, at the outset of the study we were interested in mechanism and therefore overlaid an EMA protocol on the basic trial design. The basic idea during the planning of the study was that if the experimental group evidenced better outcomes, that this might come about by alterations in mood or health behaviors. By intensively examining patients both prior to the manipulation and after the manipulation (although for a limited time), we effectively created the opportunity to examine both between-persons associations (post-manipulation levels of outcomes by group) and within-person associations (changes over time in outcomes within groups).

One week prior the three experimental writing sessions, which were conducted at our laboratory, patients completed brief diaries assessing affect, sleep, substance use (alcohol, tobacco, and caffeine) medication use, social setting (with other people), and stress three times a day. Daily monitoring continued through the 3 days of the experimental manipulation, and then for another 14 consecutive days.

Results from the EMA portion of the study indicated no reliable pattern of group differences or changes over time that could explain the beneficial effects of the intervention (Stone et al., 2000). There are several possible reasons for this negative result, including the design's

limitation of examining only 2 weeks post-manipulation (especially in light of the primary outcome variables showing significant effects 4 months later) and selection of potential mediators that were studied (e.g., relatively few cognitive change variables were included). Nevertheless, we believe that a strong case was made that those variables that were studied were not affected by the manipulation and that was, after all, the primary purpose of the EMA component of the design. In other words, it is difficult to conjure up scenarios where affect and the health behaviors that we studied were actually affected by the structured writing, but that this effect was actually missed (apart, of course, from power considerations: it is always possible to argue that greater power might have uncovered small effects). More generally, this study reveals that EMA approaches can be used assess potential mediators of intervention effects in a dynamic and intensive fashion.

OTHER CONSIDERATIONS FOR EMA DESIGNS

Although we may have waxed poetic about the virtues of EMA, we recognize several limitations to and challenges for the field (for other, see Scollon et al., 2003). The first challenge concerns the implementation of EMA designs. The study described above comparing compliance with paper versus electronic diaries certainly leads us to recommend the use of the latter, especially if the device is structured so that it enhances compliance by being user-friendly and minimally intrusive. However, field implementation of electronic diaries poses considerable challenges for the researcher. There are a large number of issues ranging from the design of the sampling scheme, the thoughtful consideration of questionnaire design, training of participants, monitoring participants in the field and providing help when problems arise, motivating participants to follow the protocol, and dealing with the technological difficulties inherent in the use of palmtop computers. Christensen et al. (2003) go a long way to enabling researchers to conduct momentary data collection, and we applaud this activity. Nevertheless, the user needs to be aware that EMA is complex and can be a frustrating research strategy, though there are large rewards to be reaped from successful implementations. It is also an expensive proposition given all of the considerations that go into conducting a successful study and researchers interested in large-scale studies with EMA need to plan accordingly. The interested reader may wish to review a recent paper

detailing many of the considerations that need to be reported about a momentary study (Stone and Shiffman, 2002).

Second, momentary data collection techniques yield masses of complex, time-dependent data. At first glance, this may appear to be a strength of the technique and, indeed, it is. The downside to the quantity and quality of these data sets is that they require considerable statistical and data management acumen. For example, EMA data sets are usually oriented with moment as the unit of analysis, yielding very “tall” data files. Often sophisticated data manipulation is required to re-orient these files for particular analyses, because aggregation may be required over persons, days, and/or situational characteristics. These manipulations often go well beyond the programming skill level of the average social scientist. Statistical analyses of the data can also be challenging for the same reasons: by their nature, EMA data are multi-level, meaning that the analysis techniques most appropriate for the data are specifically designed for multi-level data. Recent articles have discussed the pros and cons of different analytic approaches (Affleck et al., 1999; Reis and Gable, 2000; Schwartz and Stone, 1998). Despite these helpful guides, there is still a considerable degree of technical expertise required for effective analyses of momentary data. Finally, missing data presents a common, yet vexing, problem for EMA research. An assessment protocol requiring multiple assessments each day over a period of days or weeks will produce missing data in the majority of participants. Such missing data may not be balanced either within or between participants, may reflect characteristics of an individual, an assessment or study design, or even represent a response to a treatment. As such, great care must be taken in the handling of missing data, both in terms of the statistical/analytic approach taken (e.g., the imputation or representation of missing data in models) as well as in the final interpretation of the study data (even when utilizing statistical approaches, such as multi-level random effects modeling, that relatively seamlessly account for missing data in model testing).

Third, it is important to keep in mind that momentary data collection is often used in the context of observational or non-experimental research designs. Given the rich detail and the temporal discrimination afforded by EMA designs, it is sometimes tempting to covertly reconceptualize an observational trial as experimental, at least in terms of the discussion of causality among studied variables. We should take care to remember that most momentary studies (there are exceptions, such as their use in randomized controlled trials) will not have randomly

assigned the independent variable to either occasions or to individuals, and, thus, we need to remain circumspect in our conclusions. This is not meant as a discouraging coda to the article, but only as a sharp caution that we note the limitations and challenges of EMA approaches.

CONCLUSION

EMA and other strategies for capturing momentary data provide researchers with a new tool to examine daily and within-day phenomena. Within the domain of self-reported phenomena, momentary data capture has the ability to reduce retrospective recall biases, which may provide an alternative view of these constructs. More generally, momentary data capture provides the user with a very rich, dynamic picture of peoples' daily lives. Such data provide both insight into temporal (e.g., diurnal) processes, as well as have the potential to uncover mediational pathways linking concepts. The ability to move out of the laboratory and into real life is especially important for research areas where the psychological and/or environmental milieu could affect the individual. We have argued that behavioral medicine research could profit by incorporating EMA into its research and have attempted to show by example how this could occur. There are limitations and challenges to momentary research paradigms, but progress has been made to overcome these issues yielding an approach that should continue to prove fruitful in behavioral medicine research.

REFERENCES

- Affleck, G., A. Zautra, H. Tennen and S. Armeli: 1999, 'Multilevel daily process designs for consulting and clinical psychology: a preface for the perplexed', *Journal of Consulting and Clinical Psychology* 67, pp. 746–754.
- Bolger, N.: 1990, 'Coping as a personality process: A prospective study', *Journal of Personality and Social Psychology* 59, pp. 525–537.
- Bolger, N. and A. Zuckerman: 1995, 'A framework for studying personality in the stress process', *Journal of Personality and Social Psychology* 69, pp. 890–902.
- Christensen, T.C., L. Feldman Barrett, E. Bliss-Moreau, K. Lebo and C. Kaschub: 2003, 'A practical guide to experience-sampling procedures', *Journal of Happiness Studies*, this issue.
- Csikszentmihalyi, M.: 1994, *Flow: The Psychology of Optimal Experience* (Harper Collins, New York).
- DeVries, M.: 1992, *The Experience of Psychopathology* (Cambridge University Press, Cambridge, England).

- Dimsdale, J.: 1984, 'Generalizing from laboratory studies to field studies of human stress physiology', *Psychosomatic Medicine* 46, pp. 463–469.
- Folkman, S. and R.S. Lazarus: 1980, 'An analysis of coping in a middle-aged community sample', *Journal of Health and Social Behavior* 21, pp. 219–239.
- Gorin, A.A. and A.A. Stone: 2001, 'Recall biases and cognitive errors in retrospective self-reports: A call for momentary assessments', in A. Baum, T. Revenson and J. Singer (eds), *Handbook of Health Psychology* (Erlbaum, New Jersey), pp. 405–413.
- Houtman, I. and F. Bakker: 1987, 'Stress in student teachers during real and simulated standardized lectures', *Journal of Human Stress* 13, pp. 180–187.
- Kamarck, T.W., D.L. Janicki, S. Shiffman, D.E. Polk, M.F. Muldoon, L.L. Liebenauer, J.E. Schwartz: 2002, 'Psychosocial demands and ambulatory blood pressure: A field assessment approach', *Physiology & Behavior* 77, 699–704.
- Kirschbaum, C. and D.H. Hellhammer: 1994, 'Salivary cortisol in psychoneuroendocrine research: Recent developments and applications', *Psychoneuroendocrinology* 19, pp. 313–333.
- Nicolson, N.: 1991, 'Stress, Coping and Cortisol Dynamics in Daily Life', in M. deVries (ed.), *The Experience of Psychopathology* (Cambridge University Press, Cambridge, UK), pp. 21–42.
- Ockenfels, M., L. Porter, J. Smyth, C. Kirschbaum, D. Hellhammer and A. Stone: 1995, 'The effect of chronic stress associated with unemployment on salivary cortisol: Overall cortisol levels, diurnal rhythm and acute stress reactivity', *Psychosomatic Medicine*, 57(5), 460–467.
- Pickering, T. and R. Friedman: 1991, 'The white coat effect: A neglected role for behavioral factors in hypertension', in P. McCabe, N. Schneiderman, et al. (eds), *Stress, Coping, and Disease* (Lawrence Erlbaum Associates Inc., Hillsdale, NJ), pp. 35–49.
- Reis, H.T. and S.L. Gable: 2000, 'Event-sampling and other methods for studying everyday experience', in H.T. Reis and C.M. Judd (eds), *Handbook of Research Methods in Social and Personality Psychology* (Cambridge University Press, New York), pp. 190–222.
- Schwartz, J.E., J.M. Neale, C.A. Marco, S. Shiffman and A.A. Stone: 1999, 'Are there really trait-like ways of coping?' *Journal of Personality and Social Psychology* 77, pp. 360–369.
- Schwartz, J.E. and A.A. Stone: 1998, 'Data analysis for EMA studies', *Health Psychology* 17, pp. 6–16.
- Schwartz, N. and S. Sudman: 1994, *Autobiographical Memory and the Validity of Retrospective Reports* (Springer-Verlag, New York).
- Scollon, C.N., C. Kim-Prieto and E. Diener: 2003, 'Experience sampling: Promises and pitfalls, strengths and weaknesses', *Journal of Happiness Studies*, this issue.
- Shiffman, Paty, Gnys, Cassel and Hicocks: 1996, 'First lapses to smoking: within subject analysis of real time reports', *Journal of Consulting and Clinical Psychology* 64, pp. 366–379.
- Smyth, J., M. Ockenfels, L. Porter, C. Kirschbaum, D. Hellhammer and A. Stone: 1998, 'The association between daily stressors, mood and salivary cortisol secretion', *Psychoneuroendocrinology* 23, pp. 353–370.

- Smyth, J., M. Soefer, A. Hurewitz, A. Kliment and A. Stone: 1999a, 'Daily psychosocial factors predict levels and diurnal cycles of asthma symptomatology and peak flow', *Journal of Behavioral Medicine* 22, pp. 179–193.
- Smyth, J.M., A.A. Stone, A. Hurewitz and A. Kaell: 1999b, 'Effects of writing about stressful experiences on symptom reduction in patients with asthma or rheumatoid arthritis: A randomized trial', *Journal of the American Medical Association* 281, pp. 1304–1309.
- Stader, S. and J. Hokanson: 1998, 'Psychosocial antecedents of depressive symptoms: An evaluation using daily experiences methodology', *Journal of Abnormal Psychology* 107, pp. 17–26.
- Stone, A.A. and J.M. Neale: 1984, 'A new measure of daily coping: development and preliminary results', *Journal of Personality and Social Psychology* 46(4), pp. 892–906.
- Stone, A. and S. Shiffman: 1994, 'Ecological Momentary Assessment (EMA) in behavioral medicine', *Annals of Behavioral Medicine* 16, pp. 199–202.
- Stone, A.A. and S. Shiffman: 2002, 'Capturing momentary, self-report data: A proposal for reporting guidelines', *Annals of Behavioral Medicine* 24, pp. 236–243.
- Stone, A., J. Broderick, L. Porter and A. Kaell: 1997, 'The experience of rheumatoid arthritis pain and fatigue: Examining momentary reports and correlates over one week', *Arthritis Care Research* 10, pp. 185–193.
- Stone, A.A., M. Greenberg, E. Kennedy-Moore and M.G. Newman: 1991, 'Self-report, situation-specific coping questionnaires: What are they measuring?', *Journal of Personality and Social Psychology* 61, pp. 648–658.
- Stone, A., J. Smyth, T. Pickering and J. Schwartz: 1996, 'Daily mood variability: Shapes of diurnal patterns and determinants of diurnal patterns', *Journal of Applied Social Psychology* 26(14), pp. 1286–1305.
- Stone, A.A., S.S. Shiffman, J.E. Schwartz, M. Hufford and J.B. Broderick: 2002, 'Patient non-compliance with paper diaries', *British Medical Journal*, 324, pp. 1193–1194.
- Stone, A., J. Turkkan, C. Bachrach, J. Jobe, H. Kurtzman and V. Cain: 2000, *The Science of Self-Report* (Lawrence Erlbaum Associates, Mahwah, New Jersey).
- Stone, A.A., J.E. Schwartz, J.M. Neale, S. Shiffman, C.A. Marco, M. Hickcox, J. Paty, L.S. Porter and L.J. Cruise: 1998, 'How accurate are current coping assessments? A comparison of momentary versus end-of-day reports of coping efforts', *Journal of Personality and Social Psychology* 74, pp. 1670–1680.
- van Eck, M., H. Berkhof, N. Nicolson and J. Sulon: 1996, 'The effects of perceived stress, traits, mood states, and stressful daily events on salivary cortisol', *Psychosomatic Medicine* 58, pp. 447–458.

Address for Correspondence:

JOSHUA M. SMYTH

Department of Psychology

402 Huntington Hall

Syracuse University, Syracuse

NY 13244-2340, USA

E-mail: jmsmyth@psych.syr.edu