Naturalistic Observation of Health-Relevant Social Processes: The Electronically Activated Recorder Methodology in Psychosomatics

Matthias R. Mehl, PhD, Megan L. Robbins, MA, and Fenne große Deters, Dipl Psych

This article introduces a novel observational ambulatory monitoring method called the electronically activated recorder (EAR). The EAR is a digital audio recorder that runs on a handheld computer and periodically and unobtrusively records snippets of ambient sounds from participants' momentary environments. In tracking moment-to-moment ambient sounds, it yields acoustic logs of people's days as they naturally unfold. In sampling only a fraction of the time, it protects participants' privacy and makes large observational studies feasible. As a naturalistic observation method, it provides an observer's account of daily life and is optimized for the objective assessment of audible aspects of social environments, behaviors, and interactions (e.g., habitual preferences for social settings, idiosyncratic interaction styles, subtle emotional expressions). This article discusses the EAR method conceptually and methodologically, reviews prior research with it, and identifies three concrete ways in which it can enrich psychosomatic research. Specifically, it can a) calibrate psychosocial effects on health against frequencies of real-world behavior; b) provide ecological observational measures of health-related social processes that are independent of self-report; and c) help with the assessment of subtle and habitual social behaviors that evade self-report but have important health implications. An important avenue for future research lies in merging traditional self-report–based ambulatory monitoring methods with observational approaches such as the EAR to allow for the simultaneous yet methodologically independent assessment of inner, experiential aspects (e.g., loneliness) and outer, observable aspects (e.g., social isolation) of real-world social processes to reveal their unique effects on health. **Key words:** ecological momentary assessment, experience sampling, ambulatory monitoring, ambulatory assessment, behavioral observation.

EAR = electronically activated recorder; **SECSI** = Social Environment Coding of Sound Inventory.

INTRODUCTION

Social environments play a pivotal role in the development of disease and maintenance of health (1). This scientific fact was recently reaffirmed by the National Institutes of Health (NIH) when it launched the Basic Behavioral and Social Science Opportunity Network (OppNet) in November 2009. OppNet is a trans-NIH initiative to "pursue opportunities for strengthening basic behavioral and social science research" and to promote "activities and initiatives that focus on basic mechanisms of behavior and social processes" (OppNet mission statement). The launch of OppNet was followed by a concerted release of unique funding opportunities specifically aimed at accelerating progress in the understanding of how social and behavioral processes influence health. The OppNet initiative showcases NIH's strong commitment to improving health by identifying, understanding, and addressing psychosocial risk factors.

This emphasis on psychosocial factors brings with it the need for their high-fidelity assessment. Because of their phenomenological complexity, this is usually best attained through systematic multimethod measurement (2). Interestingly, it is exactly in this regard, however, that the field of psychosomatic research is subject to a subtle but important imbalance. Whereas the measurement of the "somatic" side of the social health equation has developed rapidly with progress in noninvasive peripheral and central physiological, endocrinological, and immunological assessment (3,4), the measurement of the psychosocial context continues to rely largely on participants' self-reports.

Since the early 1980s, psychological ambulatory monitoring methods—also known as *ecological momentary assessment* (5), *experience sampling* (6), or *ambulatory assessment* (7)—have been developed to bypass methodological concerns around global and retrospective self-reports. By now, they have effectively become the criterion standard for assessing psychosocial factors in daily life, and their impact on psychosomatic research is hard to overstate. In essence, they have made it possible to study social context in vivo and in real time (8–10).

Despite this advance, the psychosomatic researcher's "tool kit" still lacks a method to directly observe "the social context in which [patients] live," which is central to the biopsychosocial model of medicine (11(p132)). Laboratory-based observational (as opposed to self-report based) methods such as videotaping couple (12) or family interactions (13) are available; however, methods for the direct observation of behavior in the natural environment are missing from the existing array of research methods. This is likely because it is not straightforward how one would go about collecting truly naturalistic behavioral data (14). Therefore, self-report–based momentary assessments are generally considered the best available proxy to behavioral observation in the field (15,16).

From a multimethod perspective, however, momentary and global or retrospective self-reports share important method variance because both derive their data from participants' reports of their introspections, perceptions, and cognitions. Thus, some concerns raised for global/retrospective self-reports potentially also apply to momentary self-reports (e.g., impression management, self-deceptive enhancement, limitations to what participants are aware) (16). Therefore, to complete the

From the Department of Psychology, University of Arizona, Tucson, Arizona.

Address correspondence and reprint requests to Matthias R. Mehl, PhD, Department of Psychology, 1503 E University Blvd, University of Arizona, Tucson, AZ 85721-0068. E-mail: mehl@email.arizona.edu

The preparation of this article was supported by the National Institutes of Health Grant R03CA137975.

The authors declare that they have no conflict of interest.

Received for publication December 27, 2010; revision received October 31, 2011.

DOI: 10.1097/PSY.0b013e3182545470

0033-3174/12/7404–0410

psychosomatic researcher's tool kit, it would be desirable to complement momentary self-report with momentary observational data. In this article, we review and discuss a novel ambulatory monitoring method for the unobtrusive sampling of naturalistic observations.

A METHOD FOR THE NATURALISTIC OBSERVATION OF SOCIAL PROCESSES IN DAILY LIFE: THE ELECTRONICALLY ACTIVATED RECORDER

Over the last 10 years, we have developed and validated the electronically activated recorder (EAR) (17), a method that unobtrusively samples acoustic observations of participants' momentary environments within the natural flow of their lives.

The EAR System

The EAR is a portable audio recorder that periodically records brief snippets of ambient sounds. Participants wear it on their waistline while going about their daily lives. In tracking moment-to-moment ambient sounds around the participants, it yields acoustic logs of their days as they naturally unfold. In sampling only a fraction of the time instead of recording continuously, it makes large naturalistic observation studies feasible.

Although the EAR started out as a system based on analog microcassette recorders, recent versions of it have run on handheld computers. The handheld computer–based EAR system has critical advantages: a) it is software-based and runs on regular commercial devices. b) It is available at an affordable price (the cost of a handheld computer). c) It allows for freely programmed recording schedules and blackout periods with no recordings (e.g., overnight). Finally, because now, traditional self-report–based ambulatory monitoring methods and the EAR use the same platforms (18), it is possible to merge both methodologies.

Since 2005, we have used a version of the EAR that runs on Windows Mobile devices (e.g., HP iPAQ 110). It was successfully used in several research projects (19,20) and continues to be used in ongoing studies. Recently, Apple certified the iEAR app that runs on both the iPhone and the iPod touch and is now available for free download on iTunes. The development of the iEAR app constitutes a significant methodological advance because the iOS a) is currently among the most robust and refined mobile computing platforms, b) has an enormous user base and is expected to have long-term technology support, and c) is an operating system that maximizes compatibility across regions (with iPhones and iPod touches being sold in various countries).

A Comparison of Self-Report–Based Ambulatory Monitoring Methods and the EAR Method

As a real-time data capture method, the EAR method compares most directly to self-report–based ambulatory monitoring methods (6–8,21). Table 1 summarizes important similarities and differences between the two methods.

The most important difference between the two methods lies in the fact that traditional psychological ambulatory monitoring methods are based on momentary self-reports, whereas the EAR is based on momentary behavioral observation. The two types of methods hence adopt different assessment perspectives: the self, with the corresponding subjective account, versus the bystander, or observer, with the corresponding objective account (12,22,23).

The practical and conceptual differences between traditional self-report–based ambulatory monitoring and the EAR as outlined in Table 1 suggest that the two methodologies are best suited for different assessments. In capturing the agent's "insider" perspective, self-report–based ambulatory monitoring is optimized for the assessment of participants' subjective experiences and perceptions (e.g., thoughts, feelings). In contrast, in capturing the observer's "outsider" perspective, the EAR is optimized for assessing audible aspects of participants' objective social environments and interactions (e.g., social settings, communication behaviors, emotional expressions).

Psychosocial Information Contained in the EAR Recordings

Although other approaches to extracting information from the EAR recordings exist (24), we have primarily used a behavior

TABLE 1. A Comparison of Self-Report–Based Ambulatory Monitoring Methods and the EAR Method

	Self-Report–Based Ambulatory Monitoring	EAR Method
Approach	Naturalistic	Naturalistic
Medium	Paper and pencil, electronic (handheld computer or mobile phone)	Electronic (handheld computer or mobile phone)
Mode	Active (data provided through voluntary response)	Passive (data collected through automatic recording)
Method	Self-report	Behavioral observation
Perspective	Self (agent)	Other (observer)
Awareness of assessment	High	Low after habituation
Burden for participant	Practical (interruption of daily life)	Psychological (intrusion of privacy)
Burden for researcher	Preparing participants (instruction and training)	Preparing the sound data (coding and transcribing)
Data collection limited by	Response burden	Privacy considerations, laboratory capacity for data coding
Optimized for assessment of	Subjective experiences and perceptions	Objective social environments and interactions

EAR = electronically activated recorder.
Adapted with permission from Mehl and Conner (10).

coding or counting strategy. Specifically, in our research, trained coders listen to all of a participant's EAR recordings and code each sound file for the presence ("1") or absence ("0") of behaviors using a standardized coding system. Over the years, we have developed the Social Environment Coding of Sound Inventory (SECSI) (25,26) to capture acoustically detectible aspects of participants' social environments and interactions. In its basic form, the SECSI comprises four category clusters: a) the person's current *location* (e.g., at home, outdoors, inferred from ambient cues, such as sounds of television and housework at home or wind blowing outside), b) *activity* (e.g., watching television, eating), c) *interaction* (e.g., alone, talking, on the telephone), and d) *emotional expression* (e.g., laughing, crying, sighing).

Conceptually, it captures information about how individuals a) select themselves into social environments (e.g., displaying a preference for spending time in one-on-one versus group settings) and b) interact with their social environments (e.g., laughing or arguing a lot; Fig. 1). Adding to the basic SECSI system, we have then developed more specific coding systems such as one for capturing the aspects of patients' interactions with their support networks (e.g., disclosure, positive or negative support received) and other coping-relevant behaviors (e.g., complaining, expressing optimism, showing humor). In our laboratory, coders also routinely transcribe participants' utterances captured by the EAR. We then analyze the transcripts to obtain information about participants' linguistic styles (27).

Psychometric Properties of the EAR Data

The psychometric properties of EAR data can be determined through estimation of a) intercoder reliability, b) test-retest reliability, and c) parallel-test reliability. We routinely obtain estimates of intercoder agreement by having all coders code a standard set of training EAR recordings. Consistent with the specific, concrete, and behavioral nature of the codings (e.g., "talking" or "laughing"), intercoder reliabilities tend to be high (25,26,28). It is an advantage of the behavior counting approach that behavior codings of specific, narrowly defined behaviors (e.g., talking) are less susceptible to interpretational ambiguity (and thus coder unreliability) than ratings of broader psychological constructs (e.g., "relationship satisfaction").

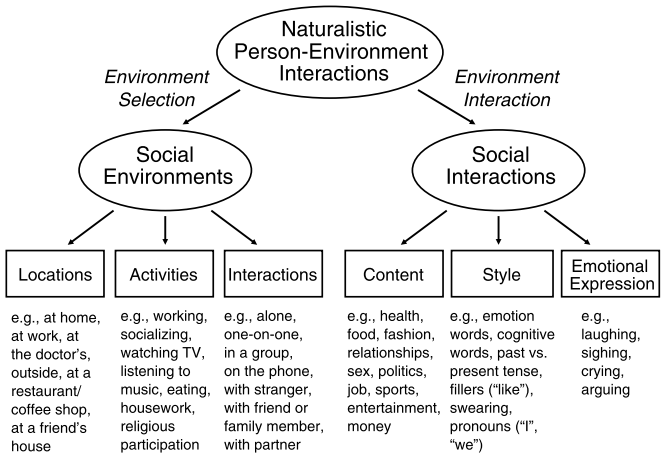

Figure 1. Electronically activated recorder (EAR) assessment of participants' daily social environments and interactions. Adapted with permission from Mehl and Conner (10).

Furthermore, Mehl and Pennebaker (26) analyzed the 4-week stability of different daily behaviors. Participants' activities and language use emerged as remarkably stable over this period, supporting the test-retest reliability of the EAR data. For example, the correlations between the two 2-day assessments (4 weeks apart) were 0.64 for how much time participants spent alone, 0.60 for how much they listened to music, and 0.77 for how much time they spent in transit. A similar picture emerged for aspects of their habitual language use. The degree to which participants used swearwords (r_{tt} = 0.86), nonfluencies (r_{tt} = 0.62), and filler words (r_{tt} = 0.59) was highly stable over the 4-week period.

Finally, estimates of parallel-test reliability can be obtained by simulating subsamples or "parallel tests" that are based on 50% of the collected data (i.e., aggregating across every other recording). Using this strategy on data from two published studies (based on the default sampling pattern of 30 seconds every 12.5 minutes) (25,29), we found that the "test halves" and the full data yielded essentially identical descriptive statistics (e.g., central tendency, dispersion). The measures based on the test halves and the full data also evidenced a very high degree of correspondence (test-half total correlation: median = 0.97, minimum = 0.66, maximum = 1.00). This was especially true for high base rate (being on the telephone and watching television) but also applied to low base rate (laughing and singing) behaviors.

More importantly, the measures based on the test halves also showed highly comparable criterion validity. Specifically, the analyses based on 50% of the data successfully "recovered" the published effects based on the full data. For example, Mehl and his colleagues (29) reported that small talk correlated −0.33 and substantive conversations correlated 0.28 with well-being. The correlations based on the test halves were −0.33/−0.32 and 0.29/0.26, respectively. Similarly, Mehl and colleagues (25) found that talking correlated 0.30 with extraversion and class attendance correlated 0.42 with conscientiousness. Again, the correlations based on the test halves were highly comparable (0.30/0.27 and 0.41/0.42).

Taken together, these analyses suggest that a) psychological information can be reliably decoded from the sampled ambient sounds, b) behavioral aggregates based on 2 days of monitoring evidence good temporal stability, and c) the default sampling pattern (30 seconds every 12.5 minutes) yields generalizable estimates of a person's daily behavior. We therefore suggest that—although longer monitoring periods are always preferable to shorter ones—a 48-hour assessment window represents a viable compromise between optimizing data quality and ensuring study feasibility and yields reliable estimates of habitual aspects of participants' daily social environments, behaviors, and interactions. Future research should test for potential day-of-the-week effects (30).

Ethical Considerations Around Conducting EAR Research

Recording ambient sounds raises ethical and legal questions. EAR studies conducted in our laboratory routinely implement a series of safeguards to protect participants' privacy and ensure the confidentiality of the data (31). We have found these safeguards to be highly effective at alleviating concerns that participants may have about the method. First, the EAR is programmed to record only a small fraction of a person's day. For example, sampling 30 seconds every 12.5 minutes leaves 95% of the time unrecorded. Second, the recordings are kept short: 30-second recordings are long enough to reliably extract basic behavioral information, yet they are short enough to capture only a small amount of contextualized personal information. Finally, and most importantly, before the investigators access the data, all participants can listen to their EAR recordings and delete the parts that they do not want on record. In one study, 19 (19.8%) of 96 participants reviewed their recordings but only erased a total of 10 sound files (25). In another study with 13 patients with rheumatoid arthritis (RA), only 1 of 2948 waking sound files was erased (19). This suggests that participants feel comfortable with sharing the sounds of their daily lives under our safeguards.

However, more serious concerns revolve not around the participants themselves but around bystanders whose behaviors are captured by the EAR. In the United States, there are few restrictions about recording people's utterances in public places. The situation concerning recording private conversations is more ambiguous. In most parts of the United States, recordings can be made legally if at least one of the interactants (e.g., the participant who is wearing the EAR) has a knowledge of the recording device. A small number of states only allow recordings if all interactants have a knowledge of the recording device. However, even in the most legally restrictive states, unauthorized recordings are only problematic if they are personally identifiable.

In our laboratory, we encourage participants to wear the microphone visibly and to readily mention the EAR in conversations with others. Irrespective of such notification, anonymity of other people's utterances is of critical importance because their behavior is collected without explicit informed consent. As mentioned previously, the brief recordings minimize the chance that personally identifying information about others is captured in the first place. Furthermore, in the coding process, any personally identifying information is omitted from the transcripts. It is thus highly unlikely that the EAR paradigm as we have established it violates privacy rights of people who are inadvertently recorded.

Obtrusiveness of and Participants' Compliance With Wearing the EAR Device

The EAR method requires participants to tolerate being intermittently recorded without exactly knowing when. This can create evaluation apprehension and result in reactance (i.e., censored or artificial behavior) or noncompliance (i.e., not wearing the EAR). Thus, it is critical to estimate how obtrusive the method is in daily life and how well participants comply with it.

Mehl and Holleran (32) addressed these questions by analyzing measures of both self-reported and behaviorally assessed EAR obtrusiveness and compliance in two samples: a short-term monitoring (2 days (25)) and a longer-term (10–11 days (33)) monitoring. Self-reported obtrusiveness was measured with items such as "To what degree were you generally aware of the EAR?" As a behavioral measure of obtrusiveness, the coders counted how many sound files the participants mentioned the EAR with others. As a self-report measure of compliance, participants reported what percentage of the day they were wearing the EAR. Finally, as a behavioral compliance measure, the coders counted the number of sound files that indicated that participants were not wearing the EAR. This was coded if, over the whole sound file, no ambient sounds at all were recorded—not even sounds of breathing or clothes rubbing against the microphone.

Results revealed that, soon after receiving the EAR, participants go through a brief period of heightened self-awareness in which conversations about the EAR are frequent. Within 2 hours, however, most participants habituate to the method and rarely mention it with others (Panel A of Fig. 2). This habituation effect was found for the short-term monitoring and longer-term monitoring.

The study further yielded the following findings about participants' compliance: in the short-term monitoring, participants' compliance was very high in the first hours after they had received the EAR. Noncompliance gradually increased over time and leveled off at approximately 10% to12% of the time on the second day of monitoring (Panel B of Fig. 2). Compliance in the longer-term monitoring was high for at least 6 days. After that, variability in noncompliance increased, suggesting that some participants' tolerance threshold may have been reached.

The compliance data reported in Mehl and Holleran (32) are based on two studies with student samples. We have since run a series of EAR studies with samples of older adults (e.g., individuals with RA (19), couples in which one member was receiving treatment for breast cancer (20), faculty members at a Tier One research institution (34)) and have obtained highly comparable results regarding EAR obtrusiveness and compliance.

Together, this suggests that a) EAR compliance and obtrusiveness can be reliably assessed, b) compliance is generally high and comparable with what has been reported for self-report–based ambulatory monitoring (35,36), and c) after an initial, short habituation period, the method operates fairly unobtrusively and does not interfere much with participants normal activities.

Practical Considerations for Running an EAR Study

Researchers can obtain a copy of the (old) Windows Mobile–based EAR software from our laboratory or download the new iEAR app directly and freely from iTunes. Also available from our laboratory is a researcher's guide with practical information on how to run an EAR study. It contains instructions for

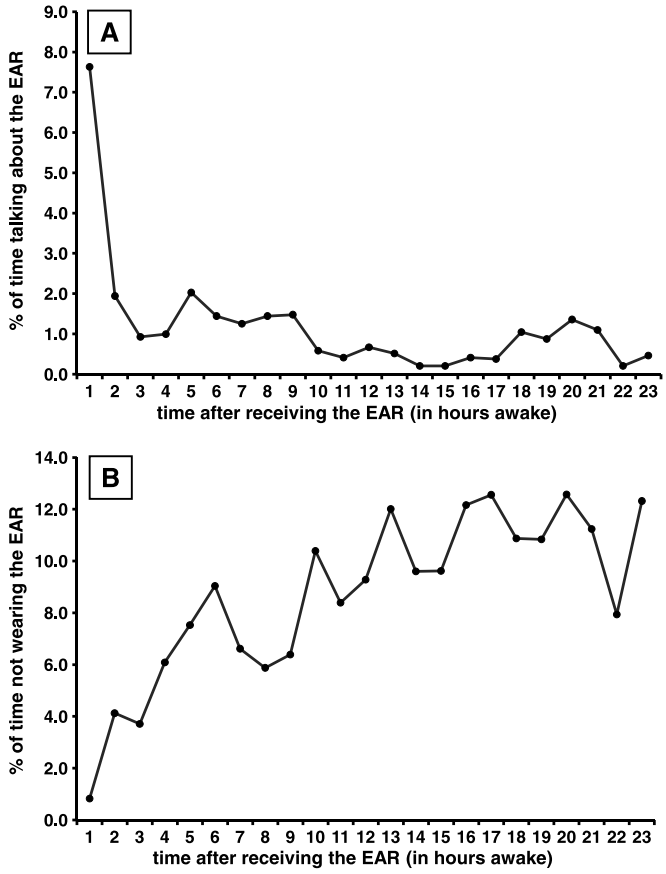

Figure 2. Behaviorally assessed obtrusiveness of and compliance with the electronically activated recorder (EAR) method. Reprinted with permission from Mehl and Holleran (32).

installation and use of the software, a sample consent form, a set of standard questionnaires (e.g., EAR compliance and obtrusiveness questionnaire), and a script on how to administer the EAR. Finally, we are always glad to help researchers start and troubleshoot EAR studies.

POTENTIAL OF THE EAR METHOD FOR PSYCHOSOMATIC RESEARCH

As a novel observational ambulatory monitoring method, the EAR needs to justify its use through bringing unique potential to the assessment of social processes. In this section, we discuss three concrete and important ways in which naturalistic observation of social processes with a method such as the EAR can enrich psychosomatic research.

Calibration of Psychosocial Effects on Health Against Frequencies of Real-World Behavior

A persistent problem in research on psychosocial influences on health is the calibration and interpretation of effects. In a seminal article, Sechrest et al. (37(p1065)) pointed out that "very few psychological measures of any kind are expressed in a metric that is intuitively or immediately meaningful." Instead, most standardized scales use arbitrary metrics, such as 5- or 7-point rating scales. Such measures have no clear referents that inform about what scoring at a certain level (e.g., a "5" on optimism) means for how a person fares in important domains of life. For example, how much less time does a person with a "4" on a social integration scale spend alone compared with a person with a "2"? In addition, how does a person's daily life change if an intervention reduces his depression score by 7 points? Finding answers to questions like these is crucial for understanding the implications of psychosocial effects on health. Yet, the field has struggled greatly with accomplishing this (38).

One advantage of the EAR is that its sound file–based behavioral codings can be readily converted into a metric that is nonarbitrary, intuitively meaningful, and inherently real-world relevant. If the EAR captures a person talking in 40 of 120 recordings, one can estimate that the person spent approximately a third of her time awake (approximately 5 hours) talking. Or, if television sounds are present in 20% of the recordings, one can estimate that the person was (actively or passively) watching television 20% of the waking day (approximately 3½ hours). By linking EAR-derived frequencies of daily behavior to the metrics of measures, a better understanding of the real-world implications of psychosocial influences on health can be obtained.

In a recent study, for example, Mehl and colleagues (29) found that greater subjective well-being is related to a) spending less time alone and more time with others and b) having less small talk and more substantive conversations. The magnitudes of these effects were vividly illustrated by the fact that, compared with the unhappiest participants (−2.0 SD), the happiest ones (1.5 SD) spent 25% less time alone (58.6% versus 76.8%) and 70% more time talking (39.7% versus 23.2%). They also had roughly one third as much small talk (10.2% versus 28.3%) and twice as many substantive conversations (45.9% versus 21.8%).

Similarly, testing the myth that women are by a factor more verbose than men, Mehl and colleagues (39) estimated based on six EAR studies that both men and women use approximately 16,000 words per day. A sex difference of 546 words compared with a range of more than 46,000 words between the least and the most talkative individuals (695 versus 47,016) rendered significance testing close to meaningless and vividly illustrated the large magnitude of individual differences in social engagement.

Finally, and directly relevant to the health context, we recently began a collaboration with Charles Raison where we use the EAR to study how proinflammatory cytokines regulate prosocial and affiliative behavior. Prior research has found inflammatory and socioemotional processes to be tightly linked, but how and how much specific, real-world social behaviors are affected by chronic inflammation is largely unknown (40–42). In using the EAR in patients with hepatitis C before and after interferon alfa (or no) treatment, subtle changes in daily social life (e.g., amount of arguing or swearing) can be objectively traced and quantified in an intuitive, nonarbitrary metric.

Thus, in facilitating an absolute metric in the measurement of daily social processes, the EAR can help "benchmark" and interpret psychosocial influences on health.

Ecological Measurement of Health-Related Social Processes Independent of Self-Report

Understanding both "the quantity and quality of social relationships and support in society" is central to unpacking how social relationships affect health (43(p544)). However, how can the effects of objective versus subjective aspects of social processes be studied separately if the two components cannot be assessed independently without shared method variance? Clearly, it is possible to ask participants to report on both aspects, and often, their answers are meaningful and interpretable (e.g., feelings of loneliness (44), size of social network (45)). Yet, ultimately, in both cases, the data are fundamentally based on participants' accounts of their social worlds and thus susceptible to biases that are inherent to self-reports (e.g., social desirability, impression management, lack of self-knowledge, mood effects, cultural relativity). For example, a recent study of the differential impact of loneliness and social isolation on health (46) brought to light the need for independent measures of subjective (loneliness) and objective (social isolation) aspects of the social environment. Specifically, the authors concluded that the complex nature of social environments necessitates a more complex and multimethod measure of social isolation.

In addition, researchers studying the impact of psychosocial factors on health are often interested in constructs that also influence the *reporting* of health factors that renders self-report as a measure of both predictors and the outcomes problematic. For example, self-focus and rumination are characteristic depressive symptoms that, at the same time, significantly affect symptom reporting (47). Using a behavioral observation approach to replace self-reports of either audible depressive behaviors or illness symptoms (e.g., sighing, groaning) could eliminate critical shared method variance. Thus, at a basic level, the EAR provides researchers with an opportunity to assess social processes with a method that is independent of participants' self-reports.

An example of how this independent assessment can matter comes from a recent cross-cultural EAR study. Ramírez-Esparza and her colleagues (48) collected self-report and observational data on how sociable Americans and Mexicans are. Consistent with prior self-report research, they found that American participants reported being more sociable than their Mexican counterparts. At the same time, however, they spent significantly *less* time with others and had *fewer* social conversations. Furthermore, whereas Americans rated themselves significantly *higher* than Mexicans on the item "I see myself as a person who is talkative," they spent, in fact, almost 10% less time talking.

Mehl and colleagues (39) revealed a similar self-report versus observation dissociation in their study of male and female talkativeness. In the absence of any sex difference in observed daily word use, women reliably rated themselves higher than men on the item "I see myself as a person who is talkative." Together, the two studies provide strong evidence for the need to assess culturally embedded social constructs from both the agent's and the observer's viewpoints.

More relevant to the health context, in collaboration with Charles Raison, we are currently using the EAR to study the potential of compassion meditation to change practitioners' real-world social lives. There is consensus among meditation researchers that a clear demonstration of the real-world, prosocial effects of meditation is crucial for the field. Importantly, however, retrospective and even momentary self-reports cannot unambiguously distinguish change at the level of the self-concept from change at the objective, behavioral level. If participants report having emerged as calmer and more compassionate after meditation training, it is not clear whether this is a result of meditation transforming self-perceptions or due to practitioners, in fact, engaging in more calm and compassionate acts in their daily lives. Counting objectively (i.e., acoustically) documented prosocial and compassionate acts can circumvent this methodological issue and help provide the direct behavioral evidence that the field has been seeking. Interestingly, pilot data based on 24 participants suggest that compassion meditation, compared with a control intervention, increased the amount of time that the participants spent a) interacting with others, b) having substantive conversations, c) laughing, and d) expressing empathy, suggesting that meditation may in fact have the potential to transform practitioners' outer social lives independent of (or in addition to) their inner worlds.

Taken together, in providing ecological behavioral assessments that are independent of self-reports, the EAR can help disentangle subjective aspects from objective aspects of participants' social environments and thereby facilitate progress in the scientific understanding of how social processes are implicated in health and disease.

Assessment of Subtle and Habitual Social Behaviors With Important Health Implications

Asking participants to accurately report on subtle or habitual social behaviors is a task that often goes beyond what self-report can accomplish. For example, mundane behaviors, such as sighing or laughing, tend to be particularly difficult to self-report because occurrences become indistinguishable and irretrievable (49). Momentary or end-of-day event diaries can evade memory problems associated with retrospective self-reports, but even in event diaries, participants can only report what is noticed (16). In daily life, many subtle and habitual behaviors do not pass this threshold of conscious recognition. Therefore, the study of how these more elusive—yet psychologically nontrivial—aspects of our social lives influence health necessitates observational approaches.

In a recent study that exemplifies how the EAR can capture subtle social behaviors, we tested the degree to which spontaneous sighing is an objective indicator of depression among patients with RA (19). Thirteen patients with RA wore the EAR for two weekends separated by 1 month. Depression and physical symptoms were assessed via questionnaires. As an "an obvious exaggerated exhalation of breath" (50(p366)), incidents of sighing could be reliably coded from the sampled ambient sounds. Interestingly, and consistent with the notion that sighing is a depression behavior more so than a pain behavior,

the number of sighs per hour was significantly and strongly related to patients' depressive symptoms and nonsignificantly and less strongly to their experienced pain and flare days. These findings suggest that sighing is a subtle yet observationally traceable social behavior that has clear implications for mental health and therefore warrants further investigation in psychosomatic research.

Other behaviors are less subtle but highly automatic and thus difficult to report (49). For example, swearing's habitual nature in everyday conversations makes it much less salient than the interaction as a whole (51). In a recent study of the intrapersonal and interpersonal consequences of swearing, we combined data from two pilot studies (20). Thirteen women with RA and 21 women with breast cancer wore the EAR on weekends to track their daily conversations. All sound files were transcribed and submitted to Linguistic Inquiry and Word Count (52), a computerized text analysis program, to calculate the frequency of spontaneous swearing. In addition, participants completed self-reported measures of depressive symptoms and emotional support at the start of the EAR weekend and several months later at the follow-up. Consistent with the idea that swearing can repel support at the expense of psychological adjustment, swearing in the presence of others, but not swearing alone, was related to decreases in emotional support and increases in depressive symptoms over the study period. Furthermore, decreases in emotional support mediated the effect of swearing on disease severity–adjusted changes in depressive symptoms. Similar to the findings regarding the sighing-depression link, they highlight the importance of investigating social behaviors that have potentially important health implications but are often too subtle or habitual for participants to report.

CONCLUSIONS AND FUTURE DIRECTIONS

As a naturalistic observation method, the EAR eavesdrops on people's daily lives and provides highly naturalistic, experientially vivid, and psychologically rich information about their moment-to-moment (acoustic) social worlds. This information can be useful for psychosomatic research to determine how objectively traceable aspects of people's social worlds influence outcomes such as disease progression and psychological adjustment. Within the array of ambulatory monitoring methods (7,8,10), the EAR certainly occupies a methodological niche: it is not for everyone and everything. As an observational method, it is labor intensive and thus requires careful deliberation as to when it should be used instead of more economic self-report–based ecological momentary assessment methods (53). However, in providing ecological behavioral measures that are independent of self-report and often beyond what self-report can capture, it can yield insights into findings that are difficult to obtain otherwise and thereby support the field in the mission to identify "basic mechanisms of behavior and social processes" (OppNet mission statement).

With the existing possibility of assessing objective and subjective aspects of these behavioral and social processes truly independently—that is with no shared method variance (2)—the most important future direction will be to begin to merge the EAR method with traditional self-report–based ambulatory monitoring methods. This would create a uniquely powerful way of studying social influences on health in their natural habitat and with the fine-grained multimethod approach that is afforded by the complexity of the phenomenon.

Specifically, combining self-report–based and observational ambulatory monitoring methods can yield important progress in three ways. First, having time-synchronized momentary self-reports along with the moment-to-moment EAR codings can help contextualize and disambiguate behavioral measures. For example, when studying the health implications of laughing (54), it could help clarify the degree to which an observed spontaneous laughter is an emotional expression (and thus an indicator of positive affect) or a response to the social context (and thus an indicator of sociability)—implicating two distinct psychosomatic pathways (55).

Second, it can help anchor momentary experiences in or calibrate them against measures of social reality. For example, among coping researchers, the question of "just how much cancer intrudes upon and organizes the lives of couples confronted with the disease" (56(p24)) continues to be a contentious issue. In a recent meta-analysis, Hagedoorn and his colleagues (56(p24)) recommended that "direct sampling of their interactions and daily experiences that does not presuppose an answer to that question could prove illuminating in this regard," and clearly, it is the *joint* sampling of daily interactions and experiences that will provide the most informative answer.

Finally, having parallel access to both the perspective of self and the observer can help with theoretically clarifying those social phenomena where systematic discrepancies between the two accounts are an inherent and important part of the phenomenon. For example, because of the operation of self-protective defense mechanisms, persons with narcissism may rate social interactions (momentarily and retrospectively) as engaging, pleasant, or inspiring, whereas observers may judge the same interaction to be rejecting, aversive, and offensive (57). Similarly, the consequences that provided social support that is or is not explicitly perceived as such (i.e., invisible support) has for psychological adjustment also necessitate the simultaneous, event-centered assessment of both the objective support delivery and the subjective support appraisal (58,59). In short, because every social process has an "inner" part and an "outer" part, it is often desirable, and at times essential, to capture the two components independently and with equally high measurement fidelity. Often, traditional self-report–based and observational ambulatory monitoring will yield comparable findings; at times, they will reveal important discrepancies—both outcomes are scientifically informative.

We thank James Pennebaker for his mentoring over the years.

REFERENCES

1. Karelina K, DeVries AC. Modeling social influences on human health. Psychosom Med 2011;73:67–74.
2. Eid M, Diener E. Handbook of Multimethod Measurement in Psychology. Washington, DC: American Psychological Association; 2006.

3. Cacioppo J, Tassinary L, Berntson G. Handbook of Psychophysiology. New York, NY: Cambridge University Press; 2007.
4. Harmon-Jones E, Beer J. Methods in Social Neuroscience. New York, NY: Guilford Press; 2009.
5. Stone AA, Shiffman S. Ecological momentary assessment (EMA) in behavioral medicine. Ann Behav Med 1994;16:199–202.
6. Hektner J, Schmidt J, Csikszentmihalyi M. Experience Sampling Method: Measuring the Quality of Everyday Life. Thousand Oaks, CA: Sage; 2007.
7. Fahrenberg J, Myrtek M, Pawlik K, Perrez M. Ambulatory assessment—monitoring behavior in daily life settings: a behavioral-scientific challenge for psychology. Eur J Psychol Assess 2007;23:206–13.
8. Stone A, Shiffman S, Atienza A, Nebeling L, editors. The Science of Real-time Data Capture: Self-reports in Health Research. New York, NY: Oxford University Press; 2007.
9. Tennen H, Affleck G, Armeli S. Daily processes in health and illness. In: Suls J, Wallston KA, editors. Social Psychological Foundations of Health and Illness. Malden, MA: Blackwell Publishing; 2003:495–529.
10. Mehl M, Conner T, editors. Handbook of Research Methods for Studying Daily Life. New York, NY: Guilford Press; 2012.
11. Engel GL. The need for a new medical model: a challenge for biomedicine. Science 1977;196:129–36.
12. Heyman R. Observation of couple conflicts: clinical assessment applications, stubborn truths, and shaky foundations. Psychol Assess 2001;13:5–35.
13. Margolin G, Oliver P, Gordis E, O'Hearn H, Medina A, Ghosh C, Morland L. The nuts and bolts of behavioral observation of marital and family interaction. Clin Child Fam Psychol Rev 1998;1:195–213.
14. Campos B, Graesch A, Repetti R, Bradbury T, Ochs E. Opportunity for interaction? A naturalistic observation study of dual-earner families after work and school. J Fam Psychol 2009;23:798–807.
15. Conner T, Tennen H, Fleeson W, Barrett L. Experience sampling methods: a modern idiographic approach to personality research. Soc Personal Psychol Compass 2009;3:292–313.
16. Piasecki T, Hufford M, Solhan M, Trull T. Assessing clients in their natural environments with electronic diaries: rationale, benefits, limitations, and barriers. Psychol Assess 2007;19:25–43.
17. Mehl MR, Pennebaker JW, Crow DM, Dabbs J, Price JH. The electronically activated recorded (EAR): a device for sampling naturalistic daily activities and conversations. Behav Res Methods Instrum Comput 2001; 33:517–23.
18. Ebner-Priemer UW, Kubiak T. Psychological and psychophysiological ambulatory monitoring: a review of hardware and software solutions. Eur J Psychol Assess 2007;23:214–26.
19. Robbins ML, Mehl MR, Holleran SE, Kasle S. Naturalistically observed sighing and depression in rheumatoid arthritis patients: a preliminary study. Health Psychol 2011;30:129–33.
20. Robbins ML, Focella ES, Kasle S, Weihs KL, Lopez AM, Mehl MR. Naturalistically observed swearing, emotional support and depressive symptoms in women coping with illness. Health Psychol 2011;30:789–92.
21. Bolger N, Davis A, Rafaeli E. Diary methods: capturing life as it is lived. Annu Rev Psychol 2003;54:579–616.
22. Weiss RL, Heyman RE. Observation of marital interaction. In: Fincham FD, Bradbury TN, editors. The Psychology of Marriage: Basic Issues and Applications. New York, NY: Guilford Press; 1990:87–117.
23. Sillars A. Behavioral observation. In: Montgomery B, Duck S, editors. Studying Interpersonal Interaction. New York, NY: Guilford Press; 1991:197–218.
24. Mehl MR. Eavesdropping on health: a naturalistic observation approach for social health research. Soc Personal Psychol Compass 2007;1:359–80.
25. Mehl MR, Gosling SD, Pennebaker JW. Personality in its natural habitat: manifestations and implicit folk theories of personality in daily life. J Pers Soc Psychol 2006;90:862–77.
26. Mehl MR, Pennebaker JW. The sounds of social life: a psychometric analysis of students' daily social environments and natural conversations. J Pers Soc Psychol 2003;84:857–70.
27. Pennebaker JW, Mehl MR, Niederhoffer KG. Psychological aspects of natural language use: our words, our selves. Annu Rev Psychol 2003;54:547–77.
28. Vazire S, Mehl MR. Knowing me, knowing you: the accuracy and unique predictive validity of self-ratings and other-ratings of daily behavior. J Pers Soc Psychol 2008;95:1202–16.
29. Mehl MR, Vazire S, Holleran SE, Clark CS. Eavesdropping on happiness: well-being is related to having less small talk and more substantive conversations. Psychol Sci 2010;21:539–41.
30. Larsen RJ, Kasimatis M. Individual differences in entrainment of mood to the weekly calendar. J Pers Soc Psychol 1990;58:164–71.
31. Conner T, Lehman B. Getting started: launching a study in daily life. In: Mehl M, Conner T, editors. Handbook of Research Methods for Studying Daily Life. New York, NY: Guilford Press; In press.
32. Mehl MR, Holleran SE. An empirical analysis of the obtrusiveness of and participants' compliance with the electronically activated recorder (EAR). Eur J Psychol Assess 2007;23:248–57.
33. Mehl MR, Pennebaker JW. The social dynamics of a cultural upheaval: social interactions surrounding September 11, 2001. Psychol Sci 2003;14:579–85.
34. Holleran S, Whitehead J, Schmader T, Mehl M. Talking shop and shooting the breeze: a study of workplace conversation and job disengagement among STEM faculty. Soc Psychol Personal Sci 2011;2:65–71.
35. Green A, Rafaeli E, Bolger N, Shrout P, Reis H. Paper or plastic? Data equivalence in paper and electronic diaries. Psychol Methods 2006;11:87–105.
36. Stone A, Shiffman S, Schwartz J, Broderick J, Hufford M. Patient non-compliance with paper diaries. BMJ 2002;324:1193–4.
37. Sechrest L, McKnight P, McKnight K. Calibration of measures for psychotherapy outcome studies. Am Psychol 1996;51:1065–71.
38. Kazdin A. Arbitrary metrics: implications for identifying evidence-based treatments. Am Psychol 2006;61:42–9.
39. Mehl MR, Vazire S, Ramírez-Esparza N, Slatcher RB, Pennebaker JW. Are women really more talkative than men? Science 2007;317:82.
40. Miller AH, Maletic V, Raison CL. Inflammation and its discontents: the role of cytokines in the pathophysiology of major depression. Biol Psychiatry 2009;65:732–41.
41. Raison C, Capuron L, Miller A. Cytokines sing the blues: inflammation and the pathogenesis of depression. Trends Immunol 2006;27:24–31.
42. Eisenberger NI, Inagaki TK, Mashal NM, Irwin MR. Inflammation and social experience: an inflammatory challenge induces feelings of social disconnection in addition to depressed mood. Brain Behav Immun 2010;24:558–63.
43. House J, Landis K, Umberson D. Social relationships and health. Science 1988;241:540–5.
44. Hawkley LC, Preacher KJ, Cacioppo JT. Loneliness impairs daytime functioning but not sleep duration. Health Psychol 2010;29:124–9.
45. Bickart KC, Wright CI, Dautoff RJ, Dickerson BC, Barrett LF. Amygdala volume and social network size in humans. Nat Neurosci 2011;14:163–4.
46. Shankar A, McMunn A, Banks J, Steptoe A. Loneliness, social isolation and behavioral and biological health indicators in older adults. Health Psychol 2011;30:377–85.
47. Howren MB, Suls J, Martin R. Depressive symptomatology, rather than neuroticism, predicts inflated physical symptom reports in community-residing women. Psychosom Med 2009;71:951–57.
48. Ramírez-Esparza N, Mehl M, Álvarez-Bermúdez J, Pennebaker J. Are Mexicans more or less sociable than Americans? Insights from a naturalistic observation study. J Res Pers 2009;43:1–7.
49. Schwarz N. Retrospective and concurrent self-reports: the rationale for real-time data capture. In: Stone A, Shiffman S, Atienza A, Nebeling L, editors. The Science of Real-time Data Capture: Self-reports in Health Research. New York, NY: Oxford University Press; 2007:11–26.
50. Keefe F, Block A. Development of an observation method for assessing pain behavior in chronic low back pain patients. Behav Ther 1982:363–75.
51. Jay T. The utility and ubiquity of taboo words. Perspect Psychol Sci 2009;4:153–61.
52. Pennebaker J, Booth R, Francis M. Linguistic Inquiry and Word Count: LIWC 2007. Austin, TX: LIWC; 2007.
53. Slatcher RB, Robles TF, Repetti RL, Fellows MD. Momentary work worries, marital disclosure, and salivary cortisol among parents of young children. Psychosom Med 2010;72:887–96.
54. Vlachopoulos C, Xaplanteris P, Alexopoulos N, Aznaouridis K, Vasiliadou C, Baou K, Stefanadi E, Stefanadis C. Divergent effects of laughter and mental stress on arterial stiffness and central hemodynamics. Psychosom Med 2009; 71:446–53.
55. Provine R. Laughter: A Scientific Investigation. New York, NY: Viking; 2000.
56. Hagedoorn M, Sanderman R, Bolks HN, Tuinstra J, Coyne JC. Distress in couples coping with cancer: a meta-analysis and critical review of role and gender effects. Psychol Bull 2008;134:1–30.
57. Clifton A, Turkheimer E, Oltmanns T. Contrasting perspectives on personality problems: descriptions from the self and others. Pers Individ Dif 2004;36:1499–514.
58. Bolger N, Zuckerman A, Kessler R. Invisible support and adjustment to stress. J Pers Soc Psychol 2000;79:953–61.
59. Howland M, Simpson JA. Getting in under the radar: a dyadic view of invisible support. Psychol Sci 2010;21:1878–85.