Daily Stress and Hassles

Aidan G.C. Wright
Elizabeth N. Aslinger
Blessy Bellamy
Elizabeth A. Edershile
William C. Woods

Department of Psychology, University of Pittsburgh


Aidan Wright’s effort on this project was supported by the National Institute of Mental Health (L30 MH101760). The opinions expressed are solely those of the authors and not those of the funding source. After AGCW, authorship was determined alphabetically, second authorship is shared.

Correspondence concerning this article should be addressed to Aidan Wright, Department of Psychology, University of Pittsburgh, 4119 Sennott Square, 210 S. Bouquet St., Pittsburgh, PA, 15260. E-mail: aidan@pitt.edu
Abstract

This chapter reviews historical and contemporary approaches to measuring and assessing daily stress and hassles. Conceptual and definitional issues are covered, followed by three generations of daily stress assessment: cross-sectional, ambulatory assessment, and passive sensing. A selective summary is provided of research on daily stress as it relates to mental health.

Keywords: Daily Stress, Daily Hassles, Assessment, Measurement, Ambulatory Assessment.
Daily Stress and Hassles

How was your day? Did you make it out of the house on time? Was traffic a breeze? Slow day at the office? Deadline looming? Boss acknowledge all your hard work? Did you get along with everyone, or did you have any disagreements? Did your kid get in trouble at school? Spouse complain about your habits? How’s your sex life? Now, are you feeling stressed? The study of daily stress and hassles occupies itself with the answers to these questions. In short, daily stress and hassles refers to quotidian adversity and the friction of moving through life. This emphasis on minor or micro-stressors was brought to the fore in the 1970’s by several investigators with concerns about the exclusive focus on major life events in the stress research of the time. Perhaps no other individual or group was as influential as Dr. Richard Lazarus and his colleagues working on the Berkeley Stress and Coping Project. This work led to a series of seminal papers in the early 1980’s that included the first effort to develop a catalogue and measure of daily hassles (Kanner, Coyne, Schaefer, & Lazarus, 1981). Since then, the study of daily stress, both separate from and as a supplement to major life events, has blossomed. Major findings in this area have relevance to both physical and mental health.

Here we consider how daily stress and hassles are conceptualized and operationalized across three generations of their measurement. The first generation borrowed its methodology from the life events literature, employing cross-sectional checklists and severity ratings of experienced stressors. The second generation involves the use of ambulatory assessment techniques, encompassing daily diaries, ecological momentary assessment, experience sampling, and ambulatory psychophysiology, to intensively and repeatedly assess stress in an individual’s natural environment. Now cresting over the horizon is the third generation, in which passive sensing devices are being used in conjunction with machine learning to facilitate the automated
capture of stress in daily life.

In much of the daily hassles and stress literature there is a parallel focus on coping. Depending on one’s perspective, coping may be viewed as a necessary component of the dynamic process of stress. Transactional views of stress notwithstanding, we will not cover the concept or measurement of coping in much detail here, and instead remain primarily focused on the definition and measure of daily stress and the hassles that drive it.

**Definitional and Conceptual Foundations**

Defining daily hassles, stressors, and stress would almost seem unnecessary for anyone likely to read this chapter. By virtue of being a human adult, the reader undoubtedly has more direct and informative experience with daily hassles and stress than any text can provide. Nevertheless, that with which we are most familiar often poses the greatest challenge for assessment and measurement in the behavioral sciences. Daily hassles and stress appear to be no different. As described by Lazarus and DeLongis (1983), hassles are the "irritating, frustrating, distressing demands and troubled relationships that plague us day in and day out" (p. 247). Examples of hassles (or stressors) include quotidian events such as inclement weather, traffic, work demands, arguments with one’s spouse, financial concerns, and disciplining a child, and many others. Beyond this basic description, articulating the essence and outlining the boundary conditions of the construct proves more difficult, especially in the context of studying their relationship to mental disorder (Dohrenwend, Dohrenwend, Dodson, & Shrout, 1984).

The motivation for measuring and studying daily hassles emerged from both practical and conceptual concerns with the study of stress up through the 1970’s (DeLongis et al., 1982; Kanner et al., 1981; Lazarus, 1984; Lazarus & DeLongis, 1983). In part, the modest association between major life events (e.g., divorce, job loss, heart attack, death of a loved one) and health outcomes,
a major early focus of the stress literature, was a catalyzing frustration (Kanner et al., 1981). More important was the fact that the assessment of major life events, often in cross-section, provided little access to the processes or mechanisms by which they conduced to ill health. This aligned with a growing emphasis on viewing stress as a contextualized and transactional process (Coyne & Lazarus, 1980; Lazarus & Launier, 1978). From this perspective, an event is stressful when it is appraised as having relevance for the individual, it is harmful or threatening, and it outstrips the individual’s capacity to cope or manage. The implication is that not all events will be equally stressful for all individuals under all circumstances. Embedded in this conceptualization of stress is a process that unfolds and exerts its effects across time.

Indeed, a consensual definition of stress has long eluded the field. It is not our goal to resolve these issues, and instead we adopt an ecumenical perspective which views stress, generically, as a process. As such, “stress” can be broken down into constituent components, which in turn have measurement implications. Specifically, a key distinction to be made is between external events or stimuli (i.e., stressors) and internal responses to those events (e.g., perceived stress). As Monroe (2008) summarizes, different perspectives exist on which components should be emphasized as a function of differing definitions of stress (see also Monroe & Slavich, this volume). Different theorists and researchers have adopted views of stress as stimulus driven, response-defined (either consensually or idiosyncratically), reflected in change or the potential thereof, in terms of biological or psychological processes, and so forth (Dougal & Baum, 2001). In our view, all of these approaches have relevance and validity, and it is incumbent on the researcher to clearly define and support their selected focus and measurement approach. These are likely to vary with study aims and feasibility.

As a general rule, however, when studying daily hassles and stress, we recommend making
attempts to assess both stimuli and responses, because these will allow the most rich and textured mapping of the stress landscape. In the context of daily hassles and stress, environmental events or stimuli may include such examples as social conflicts, insufficient time at work, or lack of money. In terms of the measurement of daily stress, Kamarck, Shiffman, and Wethington, (2011) further distinguished between affective- and appraisal-based approaches to capturing responses to stimuli. Common affective responses to stress might include anxiety, sadness, or anger, whereas appraisal-based measures might emphasize cognitive evaluation of demands and available resources to deal with stressors. A clear implication is that, to the degree possible, measurement approaches should attempt to avoid confounding events and responses (Harkness & Monroe, 2016). For instance, asking whether a traffic jam occurred as an initial prompt, and then asking how stressful that particular event was. Although this would seem to be straightforward, this distinction often requires thoughtful consideration. To illustrate, in a relevant early debate, Dohrenwend and Shrout (1985) noted that although Kanner and colleagues (1981) seminal measure of daily hassles included both a list of events and a rating of severity, the rating of severity began at *somewhat severe* and ranged to *extremely severe*. Thus, there was no way to endorse events as having occurred, but that the participant was unbothered by them. In effect, event and response were confounded, even if there was the ability to measure severity to some degree.

A related issue that has long plagued the stress literature, is the importance and challenge of differentiating between *objective* and *subjective measures* of stressful events (See Monroe, 2008 for a review). At first glance, it may seem that measures of environmental stimuli would be objective, whereas measures of responses would be subjective. However, this is not necessarily the case. On the one hand, it is the case that the stress response, whether affective- or appraisal-based, will be subjective by definition. The interest is in how an individual perceives, evaluates,
and responds to any given event. On the other hand, in practice, environmental event measures vary in the degree of their objectivity. The issue is that there may be disagreement between subjects and researchers about which events qualify as stressors. For instance, does the flu count as a major medical event? What about an argument with a spouse, is that a major family conflict? Variability in the interpretation of life event categories like medical event, conflict, or even death of a close friend, can result in a highly heterogeneous, and therefore unreliable and potentially invalid assessment of events (Dohrenwend, 2006). Without exhaustive lists of life stressors, the ability to objectively evaluate experienced stress is eroded. Further, the conclusion reached is that certain measurement approaches, namely checklists of broad and ambiguous categories without a great deal of further instruction or probing result in poorer stress assessment as compared to detailed interviews (Harkness & Monroe, 2016; Dohrenwend, 2006; Monroe, 2008). Yet these limitations of checklist measures can be mitigated to some degree with additional modification (e.g., category descriptions/exemplars and severity ratings; Dohrenwend, 2006).

This issue has largely played out in the measure of major life events, not daily hassles and stress. Nevertheless, there has been some consideration of the issue in this context, as we review below (see Alemida, Stawski, & Cichy, 2011). Our intuition, although we have found little relevant discussion by other researchers, is that the measurement of categories of daily hassles and stress are less susceptible to impact of intracategory variability. One reason may be that many of the events that are the focus of daily hassles and stress measurement are so minor as to defy ambiguity or confusion. Harkness and Monroe (2016) raised this possibility as well. Another possibility is that, depending on assessment method, many daily hassles are events more proximal to the appraisal and response process than major life events thereby allowing clearer categorization. Or, perhaps it is just expected that categorization of daily hassles will be infused
DAILY STRESS AND HASSLES

with more subjectivity given the origins of the construct and their measurement. Regardless of why objectivity has received less attention in this context, the lessons learned in the assessment of life events likely have similar implications for assessing daily hassles and stress: researchers should strive to disambiguate event and response. We discuss these issues below, pointing out concerns and offering suggestions, across the three generations of daily stress measurement.

Building on this distinction between (a potential) stressor as an environmental event and stress as the subjective response, a further challenge is delineating the boundary of daily hassles and stress and symptoms of poor mental health. Indeed, almost from the outset, concerns were raised about measurement confounding of daily stress and psychiatric symptomatology (Dohrenwend et al., 1984; Kohn & Macdonald, 1992; Marziali & Pilkonis, 1986). For instance, concerning examples from the original Hassles Scale (Kanner et al., 1981) included: “Troubling thoughts about your future,” “thoughts about death,” “fear of rejection,” “not getting enough sleep,” “regrets over past decisions,” “not enough personal energy,” “nightmares,” “use of alcohol,” and “personal use of drugs.” All but the last two describe cardinal symptoms of the internalizing spectrum (e.g., depression, worry, social anxiety; see e.g., Kotov et al., 2017 for a review), whereas the final two might be considered hallmarks of the externalizing spectrum (e.g., substance use and antisocial behavior). Alternatively, alcohol or drug use might be understood in some contexts as maladaptive coping strategies, whereas nightmares are often used as markers for a severe traumatic stress reaction. If we take this to be confounding, then the concern is that associations with psychopathology would be artificially inflated, leading to overvaluing daily stress as a potential causal factor and possibly pathologizing normative experiences. This concern has been addressed somewhat on empirical grounds, in studies that have shown associations between stress inventories and key outcomes above and beyond some forms of symptomatology
(e.g., Cohen & Williamson, 1988). Others have sought to devise inventories “uncontaminated” by outcomes, including psychiatric symptomatology (e.g., Kohn & Macdonald, 1992).

At the same time, each of the above listed items from the Hassles Scale is likely to be normatively stressful to some degree. There are many others that could fit in there as well. In other words, all fall well within the bounds of normal adult experiences, we all experience them to some degree or another from time to time, and most might feel at least a twinge of passing stress, if not more, at their occurrence. That those who experience them more frequently, acutely, and for longer duration are considered to be of poorer mental health might only serve to further reinforce the point that the boundaries between normality and psychopathology are illusory. As Sullivan (1954) so eloquently put it, “We all show everything that any mental patient shows, except for the pattern, the accents, and so on (p. 183).” Indeed, some contemporary perspectives on the manifestation of classic psychiatric syndromes argue these types of mundane but stressful experiences operate as potential catalysts that can spread in activation to a network of other maladaptive behavior (Borsboom, 2017). However, who is vulnerable to these processes, when, and under what conditions is all a matter of future investigation. The key point is that the boundary between stressor, response, and feature of psychopathology is quite blurry. Although it remains a significant challenge in measurement, we do not view the overlap between daily hassles and stress as a fatal issue, but rather one that requires thoughtful attention from the investigator. The items one includes or excludes for a particular investigation or analysis may differ depending on the population or phenomena of interest, the context in which they are being studied, and the temporal resolution of sampling.

These boundary definitions highlight that the areas of mental health and daily hassles/stress is a natural, if not necessary, pairing. Many of the key concepts that have emerged in stress
research, such as reactivity or sensitivity to stress and maladaptive coping strategies (e.g., vicious cycles) suggest that stress processes, or more precisely maladaptation in these stress processes, are the core of many mental disorders. Take, for example, the first criterion of borderline personality disorder as articulated in the DSM, *frantic efforts to avoid real or imagined abandonment* (American Psychiatric Association, 2013). This describes, or at the very least implies, the outlines of a specific stress process, which includes *event* → *intermediary* → *response* (c.f., Wright, 2014). In this example, perceptions of abandonment (whether it is real or imagined is beside the point for this illustration) lead to catastrophic appraisals, ultimately triggering frantic efforts (e.g., threats, violence, suicide attempts). Many more such examples can be enlisted from throughout the DSM. As will be discussed below, a great deal of research has examined aspects of psychopathology that might be understood through the generic lens of daily stress processes, even if often framed within disorder specific terminology and assessments. Other concepts, such as stress generation, have emerged from the study of stress as it is relevant for psychiatric disorders (Hammen, 1991, 2006; Liu & Alloy, 2010). Although most often studied in the context of major life events, many of the specific manifestations of disorders (e.g., interpersonal dysfunction in depression) may lead to stress generation in daily hassles and stress, and possibly serve as the catalyst for certain types of major life events (e.g., “dependent” events like divorce or being fired for cause; Hammen, 2006).

This brings us to a final relevant consideration related to daily hassles and stress, which is that although they may be the main factor of interest in their own right, they may also be the mechanism by which major life events and chronic stressors exert their effects. For instance, a loss of job, death of a spouse, or a major medical illness may be inherently stressful to some degree, but this may be partly or even wholly explained by the impact they have on daily routine and one’s ability to deal with life’s challenges. In the case of these three examples, it is not difficult to
imagine how they might result in a loss of financial, social, and physical resources necessary to meet everyday needs and challenges. In times of gainful employ, a trip to the grocer’s is a mundane occurrence. Unemployed, the same trip can serve as a poignant reminder of the struggle to make ends meet, and activate any number of concerns about finances, health, and family. Early research found that daily hassles and stress accounted for the association between major life events and health outcomes, as well as maintained an independent predictive effect (Eckenroade, 1984; Kanner et al., 1981; Wagner, Compas, & Howell, 1988). Interpreted in a mediational framework, these findings suggest that daily hassles and stress are how major life events lead to important outcomes.

To the extent that psychopathology often involves maladaptation in the normative management of daily stressors in various ways (e.g., sensitivity, reactivity, maladaptive coping strategies), then it is reasonable to conclude that those experiencing or vulnerable to psychopathology will be more likely to be impacted by a major life event. Events may even create sensitization effects that manifest via these pathways (e.g., Monroe & Harkness, 2005). These possible mechanisms are worth considering, in large part because daily hassles and stress processes are much easier to intervene upon than major life events, which may have no direct remedy (e.g., loss of a spouse). More research is needed in this area, especially as it relates to psychopathology.

In summary, daily hassles and stress refer to the often mundane and minor stressors most of us experience daily or with some regularity—the friction of moving through daily life. The study of daily hassles and stress emerged as somewhat of a counterpoint to the study of major life events. Despite this, there are many shared considerations in their conceptualization and measurement, including the focus on stress as a process, differentiating between events and responses, and understanding that there are likely to be individual and time/context-specific
variability in the links between events and responses. Daily hassles and stress have a very blurry boundary with symptoms of psychopathology, which is important to attend to when considering research linking the two. Because of this, though, the study of daily stress and psychopathology are a natural pairing, and stress processes may serve as a generic and flexible model for understanding maladaptation in mental and behavioral functioning. We now turn to the specifics of measuring daily hassles and stress.

Three Generations of Daily Hassles and Stress Assessment

We have parsed the assessment and measurement of daily hassles and stress into three broad “generations” of techniques. These are intended to organize the available methodology, not to suggest that the techniques contained in each are mutually exclusive or conceptually incompatible. They do, however, roughly correspond to early techniques, contemporary state-of-the-art, and emerging directions in the measurement of daily stress.

First Generation – Cross-sectional Inventories

As noted in the previous section, the early measure of daily hassles and stress followed from the assessment of life events. Accordingly, similar techniques were adopted and the first available measures were checklists. The major benefit of this approach is its facility and economy of administration. Cross-sectional self-report measures are cheap and easy to administer, require no special training of staff or participants, and can be quickly and easily scored by computer.

Measures in this category generally use a retrospective cross-sectional format, asking participants to report on their frequency, degree, or severity of stressors, responses, or appraisals in the recent past. Thus, by cross-sectional we do not mean to imply that the participant is only responding about their current state, but rather is usually being asked to aggregate over a period of time. One month (or 30 days) is a common time-frame (Kanner et al., 1981; Cohen, Kamarck, &
Mermelstein, 1983). Inventories of this type can and have frequently been administered in repeated waves to prospectively study the effect of stress over time (e.g., Delongis et al., 1982).

Quite a few checklist measures are available for assessing daily hassles. The original checklist measure, the *Daily Hassles Scale*\(^1\) (Kanner et al., 1981), was fashioned after similar inventories measuring life events. It includes a list of 117 statements that are “a number of ways in which a person can feel hassled” (p. 24), plus an “other” response option. Participants are asked to circle those hassles that have happened within the last month, and then rate those that have occurred on the level of perceived severity using a scale of *somewhat severe*, *moderately severe*, and *extremely severe*. Scoring has included both frequency (count of endorsed hassles) and intensity (mean severity score of all items checked). The Daily Hassles Scale enjoyed early widespread use after its publication over three decades ago, and a large number of scales have been based on it, have modified it to create neutral wording (Delongis, Folkman, & Lazarus, 1988), expanded it to include interpersonal events (Mayberry & Graham, 2001), or altered it for special populations like adolescents (e.g., Compas et al., 1987; Kohn & Milrose, 1993) and college students (e.g., Blankstein, Flett, & Koledin, 1991; Kohn, Lafreniere, & Gurevich, 1990). As such it serves as a good exemplar for this category of measures, although researchers interested in using this type of instrument may wish to audition several for match to their specific study (e.g., special population, age range, number of items).

The *Perceived Stress Scale* (Cohen, Kamarck, & Mermelstein, 1983) has taken a different tack to the measure of stress, focusing not on events, but largely on the appraisal. As such, it represents an unambiguous measure of the stress response, but not events. Similar to the Daily

\(^1\) Note that the Daily Hassles Scale was originally presented alongside a checklist of uplifts, which are quotidian positive events. However, the addition of uplifts to hassles/stress has rarely incremented the prediction of important outcomes, and therefore they have received much less attention in the literature.
Hassles Scale, the referenced time-frame is the past month, and participants are asked to endorse how frequently they have experienced various responses to stress (both appraisal and affective) on a scale ranging from *never* (0) to *very often* (4). Example items include, “In the past month, how often have you felt you could not cope with all the things that you had to do?” and “In the past month, how often have you been angered because things were outside of your control?” The original measure included 14 items, but 4- and 10-item short forms have been developed, and Cohen & Williamson (1988) recommend the 10-item version due to the preferred psychometrics. The Perceived Stress Scale has been very widely used, and the 10-item version has been included in the National Institutes of Health Toolbox (http://www.healthmeasures.net).

Despite their widespread use, major concerns have been raised about checklist measures because they often confound event and response. As discussed above, much more has been written about this in the life events literature, although others have noted the issue applies to daily hassles and stress as well (Harkness & Monroe, 2016). We agree that the available measures do little to disambiguate events from response, making it difficult to uncritically recommend their use. However, inventories do differ in their design features in ways that mitigate the concerns somewhat. For instance, in the original Daily Hassles Scale endorsing an event required also endorsing it as stressful to some degree (i.e., there was no “not stressful” response option), yet the revised form and other measures do allow endorsing that an event occurred, but that it was not experienced as stressful (Delongis et al., 1988). Additional concerns have been raised about overlap with symptomatology (e.g., fear of rejection, thoughts about death; Dohrenwend et al., 1984), and there have been efforts to make revised scales that remove items that might also be considered symptoms (e.g., Mayberry & Graham, 2001). Different scoring algorithms also raise concerns about capitalizing on chance. For instance, authors have suggested checklists might be
scored for “frequency” (count of items endorsed) or “severity” (e.g., Delongis et al., 1982). Unless theory is sufficiently precise about which aspect of daily stress is presumed to be in operation, then considering both predictors moves a study from confirmatory to exploratory, and should therefore be treated as such in the statistical inferences and conclusions. This is especially the case when there are otherwise concerns about confounding. A final concern, is that it is very challenging to retrospectively recall relatively minor daily events with much accuracy over long periods of time, further muddying the interpretation of these checklists.

The study of life events has dealt with many of the problems with self-report checklists through the use of detailed interviews (see Harkness & Monroe, 2016; Monroe, 2008 for reviews). Although interviews serve to address many of the concerns about intracategory variability and confounding described above, it is not clear that a similar approach would be successful at mitigating this same concern in the assessment of daily hassles and stress. This is in part because the nature and number of the events differs substantially. In the case of life events, these are generally major and rarely occurring events that are likely to stand out with some prominence. In the case of daily hassles, the mundane nature of the targets is likely to undermine the utility of a long-term (e.g., 1-month) retrospective interview about myriad daily events. As such, the added time/cost of an interview is unlikely to be worth it, because it is not clear that it would lead to improved data quality given the nature of what is being recalled. For instance, other research areas have adopted interviews for assessing daily events, namely the alcohol and substance abuse literature, with limited success. Termed the time-line follow back approach (Sobell & Sobell, 1992), an interviewer reviews a recent period asking participants to recall, for example, whether and how much a participant drank. Standard techniques, such as anchoring to holidays and the participant’s schedule are used to increase the fidelity of the responses. However, this approach
has performed unevenly when compared to daily diaries, suggesting it does poorly at capturing the exact patterning of use and may even underestimate use (Carney, Tennen, Affleck, Del Boca, & Kranzler, 1998). This raises questions about its ultimate utility and incremental validity over retrospective generalizations captured by the much less cumbersome self-report questionnaires. Further, adding multiple, possibly many, categories of events would likely make the approach cumbersome. Therefore, the differences between daily hassles and stress and life events motivate construct specific solutions. However, as we discuss in the next section, an interview approach may prove beneficial in daily diary approaches to stress assessment in a manner similar to life events assessment.

As discussed throughout, thoughtful selection of items that take care not to confound event and response (e.g., neutral wording) and separate assessment of severity that gives an option for non-stressful response are important. Nevertheless, for some, this may not adequately address concerns about confounding when using checklist measures. An alternative approach is to forego attempts to measure the events, and instead just focus on subjective or perceived stress. If this is acceptable given the study aims (i.e., the emphasis is on perceived stress or response to stressors, not the events themselves), then this would be the most economical and “clean” approach. The 10-item Perceived Stress Scale would provide a well validated measure for this purpose, and we would recommend this approach.

Second Generation – Ambulatory Assessment

Up to now we have considered cross-sectional measures of daily hassles and stress. But many of the most interesting questions about daily hassles and stress as they relate to mental health require a shift to dynamic assessments (Wright & Hopwood, 2016). As Chassan (1959) put it:

“The clinician knows he [sic] is dealing with process. He [sic] cannot help but
remain unimpressed with statistical procedures and results which are applied to observations made at comparatively isolated points in time, and which do not tell him [sic] something of what has been happening along the way (pg. 397)’

In other words, if we are interested in understanding stress as a process, then it should be measured as such. Above we emphasized the importance of measuring both events and response, here we discuss approaches that are designed to measure the components of the stress process on a relevant time-scale. By definition daily hassles are intended to be those events whose effects are acute in nature, and the peak impact is likely to be transient. But how frequently an event occurs, how delayed is the peak of the response, how long lasting is the response, and what then does the response predict (e.g., maladaptive coping) are all questions of stress dynamics. Moreover, daily stress is likely contextualized, with its triggers and responses varying according to place and circumstance. To answer these questions, dynamic data is needed.

Traditionally, dynamic assessment of daily hassles and stress was difficult. However, over the past 25 years there has been revolution in the availability and affordability of the tools needed for capturing and analyzing ambulatory assessment data. Ambulatory assessment encompasses a range of techniques (e.g., ecological momentary assessment, experience sampling methodology, ambulatory psychophysiology, daily diaries) to assess behavior, physiology, and settings in an individual’s lived environment, and has experienced a dramatic rise in use in recent years (Hamaker & Wichers, 2017). Several excellent reviews have extensively covered ambulatory assessment in psychological research (e.g., Moskowitz, Russell, Sadikaj, & Sutton, 2009; Shiffman, Stone, & Hufford, 2008; Trull & Ebner-Priemer, 2013; Wrzus & Mehl, 2015), and so we will not provide an in-depth review here. We will, however, highlight several notable features as they relate to the assessment of daily stress.
Regardless of the specific instantiation (e.g., daily diary, ecological momentary assessment), ambulatory assessment is designed to sample people’s behavior intensively and repeatedly in their natural environment. As such, ambulatory assessment offers two major benefits to the daily hassles and stress researcher. First, as we alluded to above, by sampling individuals many times, the goal is to capture different points along a dynamic process. For instance, if one is interested in the affective response to stressors, by giving participants surveys about the occurrence of stressful events and affect several times a day, one could determine the effect of a stressful event on one’s affect in the moment relative to times when there were no stressors. Individual differences in this effect could be evaluated. More complex questions, such as lead-lag relationships, duration of response, and mitigating factors (e.g., interpersonal support) could be examined. Second, by asking about an individual’s current or very recent (e.g., past two hours) experiences, biases of retrospection are greatly reduced (Ebner-Priemer & Trull, 2009). Retrospective biases have been shown to affect ratings of events, and by asking immediately (or shortly; Himmenlstein, Woods, & Wright, 2018) after, these biases can be mitigated (see Shiffman et al., 2008 for a review). In the current context, this feature of ambulatory assessment is likely to be critical, because many of the targets of assessment are likely to be mundane and easily forgotten or selectively remembered upon later recall. For instance, it may be that individuals relatively frequently experience events that have the potential to be stressful (e.g., traffic jams, work demands), but only retrospectively recollect those instances that are highlighted in memory by actually being stressful. This would be consistent with literature that has found overestimates of symptom intensity and frequency (see e.g., Van Den Brink et al., 2001 for a review). By assessing potentially stressful events in situ shortly after they happened, both those that lead to a strong response and those that do not will be captured, therefore allowing for more valid estimates of the
effect of events per se as opposed to responses.

A major consideration for ambulatory assessment of stress is the sampling frame. For some questions, total amount of events or stress over a period of time may be the focus, but other designs will want to target specific events to understand their effects. Common sense dictates that assessment schedules should match the occurrence and timing of the phenomena of interest. In practice this can be challenging. In part this is because there is infrequently detailed information on the timing and frequency of relevant events. For instance, one might be interested in studying whether self-mutilation is used as a strategy to reduce negative affect or in response to stress. The question that immediately follows is does the mutilation occur immediately when stress is experienced, after some sustained period of time, or threshold of severity (Nock & Prinstein, 2004)? Even if these types of processes were well-articulated by theory, real world considerations, such as burden, fatigue, and rarity of events impinge upon the assessment enterprise. For one, events and behavior of interest may not occur on the same time-scale. Nightly sleep, often studied as both an outcome and predictor of daily stress, occurs once per day, whereas stressful events and their evoked response can happen many times within a day. Some high-value events may be quite rare (e.g., interpersonal conflict, self-mutilation), such that even with frequent sampling an assessment is unlikely to capture many instances. One might be tempted to reduce the frequency of sampling (e.g., once daily), but sample for a longer period of time (e.g., 100 days; Wright & Simms, 2016). But this may blur the specific processes leading up to the target behavior if they occur on a briefer time-scale. Selecting a sampling frame often involves trade-offs between frequency and fidelity. Being able to make reliable inferences requires some minimum of observations of an event, whereas making inferences at the right level of granularity requires the correct timing. Both are important, but may be at odds with each other in many scenarios. Thus,
rules that require sampling more or less frequently than an event can be expected to occur are generally misguided, and instead the exact timing schedule should be selected based on substantive and methodological concerns (both practical and quantitative).

Three general categories of sampling frames have been developed: event-contingent, fixed, and random interval. Event-contingent recording is tied to the occurrence of some event, a situation, or internal state. This approach is often used for rare or unpredictable events, as might be the case in meaningful interpersonal interactions (Moskowitz & Zuroff, 2004) or binge-eating episodes (Smyth et al., 2007). The target of the assessment is instructed to complete a survey or begin recording at the start of or soon after an event has occurred. A longstanding challenge associated with this approach is the difficulty determining whether all relevant events have been reported. Assuming they all were captured would be overly optimistic, and realistically some degree of missing data is expected. It is difficult to get a sense of the amount and nature of the missing data, and one must hope that the observed cases are representative of the total set of cases (i.e., missing is completely at random). A recent study that experimentally varied the sampling frame suggests that event-contingent recordings result in the same basic descriptive features (means and variances) and associations (between- and within-person) as random sampling (Himmelstein et al., 2018), although more work in the stress domain would be valuable. Further, passive sensing of behavior or ambient context could be leveraged to safeguard against missed events by detecting them and prompting the participant to respond. Alternatively, one sensor could be used to trigger other sensors (e.g., Bluetooth contact between smartphones in proximity could engage the microphone to capture an interpersonal interaction). Assessing daily hassles or stressors using an event-contingent design may prove challenging if the goal is to broadly sample hassles, because there may be too many “events” for a subject to keep track of. Yet, event-
contingent recording has been frequently used for specific types of events. For instance, theoretical models of personality disorders emphasize interpersonal situations as being triggering events for symptoms, which can be understood in a stress-response model (Hopwood et al., 2013; Miskiewitz et al., 2015). Focusing assessments on circumscribed domain, such as interpersonal situations specifically, works well in an event-contingent design.

An alternative approach is to adopt a fixed or random interval of assessment. Daily diaries commonly use a fixed interval approach. The assessment target is asked to complete a survey each evening. Less common, although also feasible, is an hourly or sub-hourly interval. For example, Kamarck and colleagues (1998) examined the associations among stressors (social conflict and task strain), emotional activation, and cardiovascular activity in daily life by sampling individuals every 45min over several days. Naturally, the narrower the interval, the more participant burden, therefore brief fixed intervals generally require stronger justification. Fixed intervals need not be symmetrical, and building in some asymmetry to more densely sample important parts of a process is conceivable (e.g., few samples during the day but frequent assessments on nights when substance use occurs; Piasecki et al., 2011). A noted problem with fixed-interval recording is that the assessment target may learn to anticipate the prompts and as a result change their behavior in anticipation (i.e., reactivity). Continuous sampling, as is the case with most passive sensors, would be considered on a fixed schedule, albeit with very high sampling rate (e.g., every minute or second).

Random or pseudo-random prompts have been used to get around these concerns. A typical approach might be to sample 6-8 times per day, but at random times or at random within a time-block (e.g., two-hours) to avoid large periods without sampling. A limitation of truly random prompts arises when one wants to use statistical models that examine auto-regressive effects (i.e.,
time-series models), which have traditionally relied on an assumption of equal sampling interval. New methods are being developed that circumvent or account for unequal intervals. Alternatively, random assessments within defined blocks might be assumed to be equidistant on average, and estimated effects interpreted accordingly. Pseudo-random sampling schedules appear to be the most frequently used in contemporary mental health research, although fixed-time and event-contingent sampling are well represented.

It should also be noted that these various sampling approaches are not mutually exclusive, such that they can be combined with good effect. For instance, one might combine event-contingent sampling with random prompt sampling in order to decouple assessments of events and affect (e.g., Greeno, Wing, & Shiffman, 2000). This approach, although more burdensome, allows for examining antecedents and consequents of events.

Some of the measurement considerations associated with intensive longitudinal daily hassles and stress assessment are similar to those discussed in the preceding section on cross-sectional methods (i.e., First Generation). For instance, accurately categorizing events and differentiating event from response remain a formidable challenge. However, the structural aspects of ambulatory assessment raise specific considerations that must be addressed. First, participant burden is a much larger concern. The most obvious and direct way this manifests is in the need for briefer assessments at each time-point. Long lists of very specific events are inadvisable, and shorter but more general assessments are preferred. For instance, using something like the revised Daily Hassles Checklist (DeLongis et al., 1988), which contains over 50 items, would likely overburden participants if given daily or more than once per day, and further it would be grossly inefficient given that any particular event is likely to be rare. Furthermore, long inventories of specific events would come with the opportunity cost of not being able to ask
detailed questions about antecedents, context, and outcomes. Brevity is at a premium.

Among the available ways to assess stressors, events, or stimuli, Kamarck et al. (2011) have distinguished between close-ended and open-ended approaches. Close-ended approaches directly ask about stressors, either generally (e.g., Has a problem occurred in the last 30 min? Marco & Suls, 1993) or specifically (e.g., In the past 10 minutes, did your activity require working hard? Kamarck et al., 1998) and participants are asked to respond with either a binary (yes vs. no) or ordinal (e.g., YES|yes|no|NO) scale. In contrast, an open-ended approach would provide a prompt (e.g., Did anything go wrong today in the house, with the children, or others in the household, at work, or elsewhere?) and solicit a free-response report of the event that could then be coded later (e.g., Caspi, Bolger, & Eckenroade, 1987). Kamarck and colleagues (2011) concluded that the “closed-ended data appear to be sufficient, in the sense that these are associated with concurrent measures of negative affect and salivary cortisol” (pp. 601-602). However, these different approaches do have potentially important implications. Close-ended approaches are more vulnerable to subjectivity, but are less burdensome for the participant and researcher. High burden for participants may introduce concerns about data quality and bias. It is worth noting that in a typical ambulatory assessment study, each participant might complete 50-100 records. If even a minority of those (e.g., 30%; Kamarck et al., 2011) report on problems, this will mean 15-30 codings per participant. Closed- and open-ended approaches to measurement are useful distinctions, but they should not be considered mutually exclusive, and combinations have been employed as well (e.g., Almeida et al., 2002).

We next consider several specific implementations of these approaches. In line with the cross-sectional assessments, the first measures created for intensive repeated measurement were checklists, largely intended for once daily assessments (i.e., daily diary designs). Examples
include the *Daily Stress Inventory* (Brantley, Waggoner, Jones, & Rappaport, 1987) and the *Assessment of Daily Events Checklist* (Stone & Neale, 1982). These measures are very similar in structure and content to traditional cross-sectional checklists, although the items are geared more to a daily assessment routine. For instance, the *Daily Stress Inventory* includes items like “unable to complete a task” and “had car trouble,” which are then checked if they occurred, and rated on a scale from “not stressful” to “caused me to panic.” As we noted above, issues related to objectivity in the assessment of the stimuli have been less commonly debated in this context (cf. Harkness & Monroe, 2016), but even a cursory review of this class of measures raises this concern. The same can be said of construct creep, wherein a subset of items clearly straddle the boundary between traditional conceptualizations of stress and symptoms of mental disorders. Some inventories (e.g., *Daily Experiences Survey*; Hokanson, Stader, Flynn, & Tate, 1992) even borrow directly from common measures of depression in their item construction. Thus, our admonitions from the introductory section remain relevant for some measures in this space. Investigators using this approach should consider whether responses are best interpreted from an environmental or perceived stress perspective.

Others have adopted a less detailed approach, more in line with the needs of studies that employ multiple assessments per day. For instance, from Suls and his colleagues, a single item was used to measure stress (Has a problem occurred in the last 30 min?), often paired with a second question related to whether it is a new or ongoing stressor (e.g., Suls & Martin, 2005). Economy is prioritized over comprehensiveness and detail. This approach might work very well if the assessment of daily hassles and stress are part of a larger bouquet of constructs that need to be assessed at each time point, but it is probably best interpreted as perceived stress, not a strict objective measure of environmental events. Others have taken a similarly concise approach, but
leverage open-ended responding to elicit detail that can be coded and used to distinguish between events and non-events, or exclude symptoms of mental disorders to avoid confounding (e.g., Bolger et al., 1995; Caspi et al., 1987; Eckenrode, 1984). The open-ended approach always affords the investigator with option of treating the data as a binary close-ended marker, or subjecting the responses to the more detailed and effortful coding.

Readers familiar with the assessment of major life events (Monroe & Slavich, this volume; Monroe, 2008) may note the lack of interview measures in the assessment of daily life stress. In large part this is because the demands of collecting intensive longitudinal data preclude repeated interviewing. There are some exceptions, such as the Daily Diary Interview (Walker, Garber, Smith, Van Slyke, & Claar, 2001). However, this measure largely used the interview format to ultimately administer a checklist. The measure that stands apart from others is the Daily Inventory of Stressful Events (DISE; Almeida, Wethington, & Kessler, 2002), which was designed to include the benefits of both the interview and the checklists approaches. Designed to be administered as a semi-structured end-of-day phone interview, the DISE uses seven stem questions to probe daily stressor occurrence, including questions about arguments or disagreements, avoided arguments or disagreements, events at work or school, events at home, experiences with discrimination, events that occurred to a relative or close friend, and finally a catch-all “other” question. In response to these, the interviewer solicits an open-ended narrative description of the event, which is recorded. Follow-up probes are used to ensure enough detail about the objective features of the event to later rate several components. Participants are also asked for a severity rating of each event using a 4-point scale (1 = not at all stressful; 2 = a little, 3 = somewhat, 4 = very stressful). For those events scored as somewhat or very stressful on this scale, further questions probe primary appraisal of what was perceived to have been at risk in the situation (e.g., finances, health and safety; Lazarus...
& Folkman, 1984). Almeida and colleagues (2002; see also Alemida, Stawski, & Cichy, 2011) provide additional detail about measurement development.

What distinguishes the DISE from similar inventories is the use of an investigator rating method, wherein the recorded answers are transcribed and rated by trained research staff using standardized criteria. In this way, the DISE gets around the thorny issue of subjective and objective ratings of events. Alemida and colleagues (2002; 2011) report high interrater reliability of the DISE features, which can be used to generate three broad classifications of events: interpersonal tensions, overloads, and network (e.g., something happens to close friend or relative) events. A full accounting of the criteria used and the resulting scores are beyond the scope of this chapter, but some notable implications bear mentioning. First, one might wonder whether objective ratings result in many fewer participant reported events being recognized as true events according to the DISE’s criteria. It appears that this is not a major issue, as only 5% of participant reported events in the National Study of Daily Events did not meet this threshold. However, it should be noted that the DISE appears to estimate a lower prevalence of daily events than checklist measures (Almeida et al., 2011). A second question is whether participant endorsement of stem questions results in a different classification of events than do staff ratings. There appears to be a larger discrepancy about the content, such that “nearly 14%” of events’ content differed across self and expert ratings. A final issue is whether and how much this impacts substantive questions that might be posed in the data. It would seem that these discrepancies do not have a major impact, as event level associations with key outcomes like physical health symptoms are similar to checklist style inventories, and the DISE is often administered in self-report form (e.g., Neupert, Almeida, Mroczek, & Spiro, 2006; Wright, Hopwood, & Simms, 2015) or only the self-report responses from the interview are used in analyses (e.g., Chiang, Turiano, Mroczek, & Miller, 2018). Thus,
it would seem reasonable to question whether the intensive and costly nature of the interview-based administration of the DISE is, in fact, necessary and worth any incremental gain in objectivity. Nevertheless, the DISE is one of the few measures that offers this potential added value.

This section has largely focused on traditional general assessments of daily hassles and stress that seek to broadly if not comprehensively cover daily stressors. At the same time, as it pertains to mental disorders, there may be specific hypothesized pathways or triggers of negative affect (response) or symptoms (e.g., maldaptive coping). When, due to theory or some other rationale, the focus is on a particular environmental event(s), alternative approaches are available. For instance, Smyth and colleagues (2007) used a greatly reduced subset of the Daily Stressors Inventory items rated as relevant for bulimia nervosa patients to investigate stress as a trigger for binge and purge episodes. This type of design selectively uses items from existing stress inventories. An alternative approach is to leverage general assessments of the environment, behavior, or cognition as specific antecedents in an unfolding stress process. Basic measures of affect could then be used to measure the response (e.g., Kamarck et al., 2011). The momentary assessment of interpersonal behavior in situations (Moskowitz, 1994; Moskowitz & Zuroff, 2005) has been profitably used in this way to study interpersonal triggers of negative affect in social phobia (Sadikaj et al., 2015), borderline personality disorder (Sadikaj et al., 2013), and narcissistic personality pathology (Wright et al., 2017). In each of these examples a general model of interpersonal behavior (the interpersonal circumplex) was used to examine the types of behaviors from interaction partners that lead to amplifications of negative affect among those with certain forms of psychopathology (e.g., those with social anxiety are sensitive to withdrawal from others, whereas those high in pathological narcissism respond negatively when others are dominant or
assertive). By using a general model that covers all relevant behavior, the specificity of the pattern can be established (e.g., negative affect is linked with withdrawal, but not submissiveness). In principle, this approach could be expanded to other domains of interest for specific forms of psychopathology.

A final consideration for the intensive repeated measurement of daily hassles and stress are statistical modeling issues. The resulting data can be used to answer straightforward questions using only basic analytic approaches (e.g., mean number of days people diagnosed with major depression experience interpersonal stress). However, some of the more exciting questions related to dynamic processes go beyond average levels, by definition. These might include questions related to instability in affect over time, establishing lead lag relationships (e.g., stress generation), within-person coupling of stress and affect (i.e., stress sensitivity/reactivity), individual differences in these dynamic features, or even personalized (i.e., idiographic) models of stress processes (e.g., Wright et al., in press). A full discussion of these issues extends well beyond the current chapter; therefore, we recommend several accessible resources such as Walls and Schafer (2006) and Bolger and Laurenceau (2013).

**Examples from Clinical Literature**

In the prior two sections, we reviewed what we have termed the first and second generations of daily hassle and stress measurement. Prior to pointing to future directions in the next section (i.e., the third generation), we pause to review some of the ways these approaches have been implemented in clinical research, primarily focusing on depression supplemented by some additional examples from other forms of psychopathology.

Early research of the association between daily stress and major depressive disorder (MDD) relied typically on cross-sectional designs consisting of retrospective reports gathered at
one or, rarely, two data collection sessions. Despite the advent of more sophisticated ambulatory assessment techniques which can capture dynamic processes, first generation designs are still used in contemporary research. Measures of daily stress in this literature often list a number of potential stressors from a broad host of domains such as home life, work/school, finances, social life, romantic life, and general hassles (D’Angelo & Wierzbicki, 2003; Frison & Eggermont, 2015; Jung & Khalsa, 1989). Specialized measures of stress have been used with special populations, such as minority groups (Safdar & Lay, 2003) and expecting mothers (Field et al., 2006). Cross-sectional studies of daily stress and depression have varied in their approaches assessing extent of experience (Anderson, Goddard, & Powell, 2010; Barker, 2007; McIntosh, Gillanders, & Rodgers, 2010; Sim, 2000), intensity (Bouteyre, Maurel, & Bernaud, 2007; Jung & Khalsa, 1989), or stressfulness (Frison & Eggermont, 2015). Cross-sectional research of this kind has routinely implicated daily hassles in the maintenance of depressive symptoms (Barker, 2007; Bouteyre et al., 2007; D’Angelo & Wierzbicki, 2003; Jung & Khalsa, 1989; McIntosh et al., 2010; Safdar & Lay, 2003; Sim, 2000). Studies assessing daily stress across multiple domains have found that some domains pull for more depressive symptoms than others (D’Angelo & Wierzbicki, 2003; Safdar & Lay, 2003; Sim, 2000), although the exact domains vary across studies. Some research with two collection sessions have also linked depression at baseline to greater daily hassles at follow-up (Barker, 2007) although this finding has been contradicted by in-the-moment studies of daily stress and depressive symptoms (Peeters, Nicolson, Berkhof, Delespaup, & DeVries, 2003). More conclusive research is needed on daily or momentary stress generation processes.

Researchers interested in the longitudinal association between daily stress and MDD have often used daily diary designs to assess participants over time. As we noted above, one benefit of adopting this research strategy is the ability to assess within-person, dynamic processes in stress
and depression (Almeida, Neupert, Banks, & Serido, 2005; Cohen, Gunthert, Butler, O’Neill, & Tolpin, 2005). As with cross-sectional daily stress research, researchers using daily diary designs often rely on measures that assess multiple domains, such as home, work, social network, and finances (Butler, Hokanson, & Flynn, 1994; Sliwinski, Almeida, Smyth, & Stawski, 2010). In one such study, the authors compared the affective reactivity of depressed patients to interpersonal and non-interpersonal stressors, finding that the former was more strongly linked with depressive symptoms (Gunthert, Cohen, Butler, & Beck, 2007). As with cross-sectional studies, measures of daily stress have varied in how they capture stress, assessing undesirability (Gunthert, Cohen, Butler, & Beck, 2005; Gunthert et al., 2007; O’Neill, Cohen, Tolpin, & Gunthert, 2004), unpleasantness (Dunkley et al., 2017), severity (Sliwinski et al., 2010) or stressfulness (Bell & D’Zurilla, 2009; O’Neill et al., 2004). Measures used in daily diary studies have been in both checklist (Butler et al., 1994; Cohen et al., 2008) and Likert (Bell & D’Zurilla, 2009; Dunkley et al., 2017; Gunthert et al., 2005, 2007; O’Neill et al., 2004; Sliwinski et al., 2010) formats. Daily diary studies have replicated the association of hassles and depression reported in cross-sectional studies (Bell & D’Zurilla, 2009; Dunkley et al., 2017; Gunthert et al., 2007; O’Neill et al., 2004). Some studies have measured the role of daily stress in the treatment of depression, finding that patients with more negative perceptions of stressors typically have poorer treatment response (Cohen et al., 2008; Gunthert et al., 2005). A common interest for daily diary researchers is affective reactivity to daily stressors, defined as the association between stress and later affect (Cohen et al., 2008; Dunkley et al., 2017; O’Neill et al., 2004; Sliwinski et al., 2010). Often, researchers calculate participants’ unique associations between momentary stress and later affect as a measure of affective reactivity (e.g., Wichers et al., 2010). A recent review of ambulatory assessment studies of MDD found that much of this literature focused on understanding how the
disorder influences momentary links between stress and affect (Aan het Rot, Hogenelst, & Schoevers, 2012). Cumulatively, these studies have found that daily stress and affective reactivity to stress both exacerbate MDD symptoms (Booij et al., 2018; Bylsma et al., 2011; Myin-Germeys et al., 2003; Peeters et al., 2010, 2003; van Winkel et al., 2015; Wichers et al., 2010). The ability of ambulatory assessment designs to study the dynamic relationship between daily stress, affective reactivity, and MDD has led to both psychological (Wichers, 2014) and biopsychosocial (Sher, 2004) theories of the disorder as well as complex quantitative models (Dunkley et al., 2017).

Although daily stress has most often been studied in terms of its associations with depression, it is likely to have implications to psychopathology writ large. Although there are notable studies in other domains of psychopathology, such as thought disorders (Collip et al., 2013; Compton et al., 2008; Myin-Germeys et al., 2013; Tessner et al., 2009), personality disorders (Coifman et al., 2012; Glaser et al., 2008; Lariviere et al., 2016; Tolpin et al., 2004; Wright, et al., 2015), and general maladaptive behaviors like aggression (Sprague, Verona, Kalkhoff, & Kilmer, 2011), to name a few, this remains an area ripe for more programmatic inquiry. We encourage researchers focused on other domains of psychopathology to consider adapting the lens of stress processes in their programs of research as has been fruitfully done in depression. It would be important to understand which of these processes are unique to depression, general to psychopathology, or differ in their manifestation across domains.

Third Generation – Future Directions with Passive Sensing

We end this chapter with a discussion of passive-sensing, which represents an exciting and promising future direction for stress research. Due to the highly technical nature of this area, we provide a conceptual overview and point to extant and emerging issues. Although ambulatory assessment measures of daily hassles and stress represent the state of the science, and we believe
they should be the method to turn to for those interested in understanding daily stress processes, they are not without limitations. For one, they are burdensome, and the burden places an upper limit on the frequency and amount of assessment. This could erode data quality by leading to participant fatigue, and it leaves many potentially interesting events or responses not sampled. As we noted above, in most cases the sampled events can be assumed to be representative of the full population of events, and this should mitigate major concerns. Nevertheless, in some scenarios this might be a major problem. For instance, when seeking to statistically model the antecedents and consequences of events and responses, assuming exact temporal ordering might be a problem. Second, most ambulatory assessment studies rely on self-reported events and responses. An individual’s subjective experience may very well be the most important psychological feature of stress, but it is unarguably an incomplete picture. The longstanding issue of objective measurement of stress resurfaces here, and presumably non-self-report measurement would be advantageous for addressing lingering concerns about biases. Largely due to these reasons, a number of researchers have turned to passive sensing in order to detect daily stress in the moment.

Passive sensing encompasses a large number of specific techniques and procedures, but all generally involve the continuous collection of data from sensors that are either worn by an individual (e.g., respiration coil) or embedded in his or her mobile devices (e.g., smartphone). The use of standard medical sensors (e.g., blood pressure cuffs) in ambulatory assessment has been around for quite some time (see e.g., Fahrenberg, 1996 for a review), but leveraging the impressive sensor array embedded in our phones (e.g., ubiquitous computing) is more novel. Indeed, the modern smartphone really only gained prominence in the past decade following the release of Apple’s iPhone in 2007. Today’s smartphone is likely to include (there are some difference across models) sensors for movement (e.g., accelerometer, gyroscope), location (GPS, Wifi),
temperature, moisture, altitude, noise (microphone), light, and proximity with other devices (e.g., Bluetooth). As Mohr and colleagues (2017) recently pointed out in a review of passive sensing as it relates to mental health, these individual sensors directly generate low level data. Which is to say, they likely are not immediately useful for measuring complex constructs like stress, but rather require several levels of aggregation or transformation. We will return to this issue shortly, but for now the key point to take away is that devices that seem as mundane as those we carry everywhere in our pockets generate highly diverse and continuous streams of data. These can be augmented with sophisticated psychophysiological sensors in specialized arrays or as part of other common devices (e.g., smartwatches and step counters) that also continuously record across several channels of sensors. Additional data on device usage can be recorded, such as screen on/off, call logs, texting, application usage, and even keystroke recording.

Some of these streams of data have obvious utility for extracting higher-level features that are easily interpretable. For instance, GPS coordinates can be used to inform types of location, such as work, home, gym, bar, and so forth. Others, such as light sensors might seem less clearly relevant, unless one considers that it might be a strong marker for something like a sleep-wake cycle. A relatively complex behavior like sleep-wake cycle will not rest entirely on a sensor like ambient light, but instead would be informed by several others (e.g., ambient noise captured from the microphone, screen activation, accelerometer). This brings us to the challenge of fruitfully using the data from passive sensing. Moving up along the conceptual hierarchy, from basic sensor-level data to mid-level feature and on to higher-level features (Mohr et al., 2017), is unlikely to be a linear affair. Many of the informative streams of data for identifying complex states are likely to be unintuitive; one would not be able to divine the important variables a priori. To make these abstractions, one is likely to need some form of machine learning. Machine learning may sound
highly technical, and in some forms it can be, but on its surface it is just some form of supervised or unsupervised learning algorithm. Arguably, step-wise regression, which most readers will remember as being discouraged in basic graduate statistics, is a basic form of machine learning. Predictors are auditioned for a regression model, and only those that are statistically significant at some criterion (e.g., $p < .05$) are retained in the final model. The basic logic is that the data drives the decisions of which variables to retain as informative. When coupled with something like logistic regression, machine learning can be used to develop a probability model for classifying some outcome (e.g., stressful event).

So, how are passive sensing and machine learning leading to the future of daily stress assessment? The basic logic is that the array of potential sensors alluded to above can be used to predict when an individual is experiencing “stress.” In order to develop a model for predicting stress, it would require instances of both stress and non-stress to differentiate between. With sufficient instances to work with, the model would be able to “learn” which features from the passive sensors predict stressful experiences with high fidelity. Once the algorithm has “learned” when an individual is experiencing stress, the participant would not be required to respond to report on stress. Therein lies the very exciting aspect of this methodology. It takes the recording of stress out of the participant’s hands, making it truly continuous, putatively objective, and non-burdensome. Beyond just assessing daily stress in research, this sort of approach could also be used to facilitate just in time interventions to disrupt stress processes.

Work in this area is still in its infancy, but progressing rapidly. One aspect of this area that is simultaneously exciting and challenging is the potential to develop personalized models of stress. That is, algorithms would be employed to develop a model for each individual that would be tailored to their own stress profile. Similar to prior sections, in passive sensing the challenge
remains of what to use as “ground truth” in defining stress, and arguably it is an even larger issue. If one uses self-reports as the target, then the model will be a model of perceived stress. We highlight one study by Plare and colleagues (2011) that illustrates how some of these issues are being dealt with in passively sensing stress. Participants ($N=21$) underwent several standard stress induction procedures in the laboratory while undergoing psychophysiological recording (Plare et al., 2011). Stressors included public speaking, mental arithmetic, and cold-pressor. Throughout the procedure participants rated their stress as well. The psychophysiological data was used to classify stressful states from the in-lab self-reports (i.e., “physiological classifier”). It achieved >90% accuracy. These algorithms were then tested in real-life using an ambulatory assessment procedure, by comparing the algorithm’s stress predictions based on worn ambulatory sensors to self-reported stress in the moment (i.e., a “perceived stress model”). The model achieved a median correlation of .72, although the value ranged across participants from ~.2 – 1.0 for individual participants.

This is just one example, and other investigators have adopted alternative procedures, for instance sampling vocal frequency from smartphone microphones (Adams et al., 2014). There are pros and cons to different systems (e.g., psychophysiological sensors can be obtrusive, microphones may not be reliable when in a purse), and significant challenges remain (e.g., accounting for other potential influences on sensors). Much work remains to be done in the passive sensing of stress. In particular, studies require much larger samples before they can be considered confidence inspiring. Indeed, it is not uncommon for studies to report single digit or low double-digit sample sizes. Thus, the field really remains at the “proof of concept” stage and remains to be scaled to the point where results might be considered reliable. Nevertheless, the potential benefits of taking the measurement of stress out of the hands of participants and patients and
offloading that to a perpetually vigilant mobile device is very appealing.

**Conclusion**

The measurement of daily hassles and stress would seem to be a mundane affair. Indeed, if the interest is in a basic assessment of perceived stress, it can be straightforward and easily implemented. But there are also many compelling research questions about the links between daily hassles and stress and mental health that require considerably greater thought and care if they are to be convincingly addressed. Major issues that plague the entire stress measurement enterprise, such as distinguishing between event and response, between subjective and objective assessment of events, and between stressors and symptoms are just as important for daily hassles and stress as they are for any other area. This has not always received detailed attention. Some aspects of the phenomenon (i.e., events are quotidian by definition) make differentiating these issues challenging, although not impossible. There is also much to be gained by contemporary approaches that assess individuals intensively and repeatedly in their natural environment in order to faithfully represent dynamic stress processes. We reiterate that one’s research goals may not be best served by using stress measures per se, and instead assessing specific contextual features that might serve as specific triggers for certain forms of psychopathology. Finally, we look forward to advances in the areas of passive sensing that are likely to lead to new insights as methods improve and sample sizes increase.
References


psychophysiological methods in monitoring and field studies. Seattle, WA: Hogrefe and Huber.


handbook of stress science: Biology, Psychology, and Health (pp. 597-617). New York: Springer.


Myin-Germeys, I., Krabbendam, L., Delespaul, P. A. E. G., & Van Os, J. (2003). Do life events have their effect on psychosis by influencing the emotional reactivity to daily life stress? *Psychological Medicine, 33*(2), 327–333.


