Researchers have been incorporating ambulatory cortisol sampling into studies of everyday life for over a decade. Such work provides an important supplement to acute laboratory stress paradigms and provides a novel perspective on the interrelationships between stress, psychological resources, and health. However, the results of many field studies have been inconclusive and more studies have been undertaken than published. We describe some of the challenges facing naturalistic cortisol researchers, including lack of power, methodological and analytical problems, and patterns of confusing or conflictual results. We then summarize key findings of published naturalistic cortisol studies to date, grouped by type of cortisol outcome (morning awakening response, diurnal slope, area under the curve, and associations between momentary experiences and cortisol). We propose research questions relevant to everyday stress researchers and suggest next steps for researchers who are interested in incorporating naturalistic cortisol sampling into future studies.

**Keywords:** cortisol; stress; naturalistic research; diurnal cortisol slope; cortisol awakening response; area under the curve; research methodology

Researchers have long been intrigued by the influence of everyday stress and coping on mental and physical health. How might daily hassles, work overload, and even poor relationship quality translate into compromised immune functioning, increased disease risk, psychosocial disturbances, and shortened longevity? One plausible pathway is through alterations in the body’s stress responding systems. The allostatic load model (McEwen, 1998) posits that chronic strain leads to adjustments in the body’s regulatory “set points,” accelerating eventual wear and tear. One such system, the HPA (hypothalamic-pituitary-adrenal) axis, which secretes the hormone cortisol, has attracted particular research attention due to its sensitivity to perceived threat and its importance for multiple physiological systems.

When the brain detects a threat, the hypothalamus produces corticotropin releasing hormone, which stimulates the anterior pituitary to secrete adrenocorticotropic hormone, triggering the release of glucocorticoids (chiefly cortisol in humans) from the adrenal glands. These glucocorticoids then play an inhibitory role, signaling the system to shut down the stress response. At low levels, glucocorticoids help the body to maintain homeostasis and play a role in immune activity, growth, reproductive functioning, and energy metabolism. However, when HPA axis functioning is compromised within a high-stress environment – for example, when glucocorticoids become either chronically...
elevated or blunted – a slew of adverse outcomes can ensue, including hypertension and cardiovascular disease, dysregulations in inflammatory and immune processes, metabolic syndrome, and deficits in cognitive functioning (Sapolsky, Romero, & Munck, 2006). While the HPA axis is multifaceted, this review will focus specifically on studies of its end-product, cortisol. As a hormone that can be sampled in saliva, blood, or urine, is robust enough to store at room temperature, and has been associated with stressful experience (Dickerson & Kemeny, 2004), cortisol has appealed to researchers both as a physiological marker of subjective and objective stress, and a potential mediator between stressful experience and physical health.

Most published cortisol studies to date have been conducted within the laboratory, in the context of paradigms designed to elicit an acute stress response. However, many of the field’s primary models of the HPA axis’s role in physical health, such as the allostatic load model, focus on chronic, rather than acute, stress, with the notion that low levels of “wear and tear” can shape physiology over time. Laboratory inductions of acute stress can provide only a limited picture of allostatic processes; for example, they may capture reactivity to a stressor, but cannot provide a detailed picture of basal HPA axis activity, which may represent a separate physiological phenomena with its own relevance for health (Sapolsky et al., 2006). Additionally, many stressors encountered in the real world may differ in intensity and character from the types of paradigms used in research, making it difficult to extrapolate to everyday functioning.

It is important for researchers to study physiological phenomena not only within the lab but in everyday life. However, the field’s knowledge of day-to-day stress responding lags behind its understanding of acute stress reactivity. A meta-analysis of HPA axis responses to laboratory inductions of acute stress (Dickerson and Kemeny, 2004) reviewed 208 studies, and the acute stress literature has continued to grow since the review was published. A more recent meta-analysis of HPA axis responses to chronic stress (Miller, Chen, and Zhou, 2007) reviewed 107 studies, but a third of these studies focused on responses to pharmacological challenge within the lab, and a quarter of the studies only sampled cortisol at a single timepoint, making it difficult to capture patterns of change across the day. In all, only 38 studies reviewed by Miller and colleagues focused on outcomes, like total cortisol output or diurnal cortisol rhythm, that reflect cortisol fluctuations in daily life. Additionally, Miller, Chen, and Zhou (2007) emphasized populations reporting high levels of chronic stress, such as trauma survivors and the clinically depressed. While field studies have increased since the introduction of salivary cortisol assays in the late 1980s, the literature on naturalistic day-to-day and moment-to-moment cortisol fluctuations in normal adult populations has not yet been formally reviewed.

This paper aims to redress this imbalance by reviewing ambulatory studies of cortisol – that is, studies in which sampling of cortisol is incorporated into everyday settings, such as the workplace and the home. We focus primarily on salivary cortisol, in keeping with the vast majority of field studies currently being conducted. The review concentrates on studies of adults and adolescents, since research on children’s HPA axis functioning outside the lab (for example, at daycare) has already been reviewed and discussed extensively (e.g. Turner-Cobb, 2005; Tarullo & Gunnar, 2006; Vermeer & van IJzendoorn, 2006; Granger & Kivlighan, 2003; Brotman et al., 2003; Shircliff, Granger, Booth, and Johnson, 2005). Given that the naturalistic cortisol literature is highly variegated, reflecting the use of different sampling protocols, statistical approaches, and even outcome variables, we take a descriptive approach, organizing studies by outcome of interest and summarizing major research findings within each category. An additional goal of the review is to describe the
methodological pitfalls that ambulatory cortisol researchers often encounter and lay out guidelines for successfully incorporating cortisol measurement into studies of daily life.

**Limitations of laboratory stress paradigms**

Physiological reactivity to stress, including the HPA axis response, has been most often assessed by laboratory paradigms designed to tax mental or physical resources. Such paradigms range from cognitive tasks to exposure to noise or emotionally evocative material. The Trier Social Stress Task (TSST), the best-researched such paradigm, requires participants to give a speech and solve math problems in the presence of an unresponsive audience, and has been conceptualized as an acute psychosocial stress induction (Kirshbaum, Pirke, & Hellhammer, 1993). Over a wide variety of populations and laboratory settings, the TSST reliably triggers HPA axis and cardiovascular changes, typically a rise in cortisol followed by a decline to baseline or below-baseline levels.

The insights yielded by research on the TSST and similar paradigms have been invaluable. However, such paradigms are designed to trigger a measurable short-term stress response, not necessarily to approximate stress in daily life, so their external validity and theoretical utility may be limited, especially for groups typically underrepresented in laboratory studies. The TSST requires participants to perform rapid math calculations and deliver a speech, stressors likely to resonate for college students, the group most often used in laboratory research, but with more limited generalizability to the types of stressors regularly encountered by most adults. Researchers are beginning to use the TSST with groups that vary by class, educational attainment, and age, which should help to shed light on the TSST's applicability to diverse populations (e.g. Fiocco, Joober, & Lupien, 2007).

Acute stress paradigms may also under-represent women's cortisol response to stress, which often appears attenuated in laboratory settings compared to men (Kudielka & Kirschbaum, 2005; Kajantie & Phillips, 2006). While this finding may reflect a biological difference, it has also been suggested that the competitive, achievement-oriented stress tasks often used in the lab elicit a stronger response from men, while women react more strongly to interpersonal or “communion” stressors like marital conflict or social exclusion (Kiecolt-Glaser, Glaser, Cacioppo, & Malarkey, 1998; Ewart et al., 1991; Malarkey et al., 1994; Stroud, Salovey, & Epel, 2002; Smith, Gallo, Goble, Ngu, & Stark, 1998). These more explicitly social stressors may be harder to simulate in a laboratory setting, making it harder for researchers to capture the full range of women's stress responding.

Some researchers have attempted to develop stress inductions that may be more ecologically valid than the TSST, such as driving simulation (Seeman, Singer, and Charpentier, 1995) or anticipation of upcoming exams (Weekes et al., 2006). Conflict or problem-solving discussions, in which couples discuss a pre-selected relationship issue, have also been used to assess physiological reactivity to dyadic interactions. Studies incorporating cortisol sampling into conflict discussion paradigms have found links between couples’ interactions, physiological responses, and relationship outcomes (e.g. Kiecolt-Glaser, Bane, Glaser, and Malarkey, 2003). However, given that they are often brief, highly structured, monitored by a researcher, and take place in unfamiliar settings, conflict tasks may not always generalize well to spontaneous behavior outside the lab, and, as with the TSST, have been most extensively tested (and may have greatest external validity) among fairly educated, affluent participants.

An implicit assumption of acute stress research is that participants’ responses to tasks like the TSST and the conflict discussion approximate their reactivity to the real-life stressors they might encounter at school, at work, and during interactions with a
relationship partner, capturing a slice of their coping and interactional style outside the lab. However, an early test of this assumption found no associations between laboratory reactivity to a speech task and cortisol responses to real-life stressors measured over five days (van Eck, Nicolson, Berkhof, & Sulon, 1996), and other studies have found inconsistencies between laboratory and ambulatory measures of cortisol (Harville et al., 2007; Powell et al., 2002). Additionally, the results of naturalistic stress studies have been more equivocal than laboratory stress studies might predict: for example, participants do not always show a strong cortisol response to momentary stressors, and their reports of high chronic stress levels or an aversive marital climate are not always associated with elevated baseline cortisol. The following sections will explore possible reasons for both the ambiguity and the relative paucity of published naturalistic cortisol studies, including methodological, analytical, and theoretical concerns.

Obstacles for naturalistic cortisol researchers

Methodological challenges. While acute stress studies may last a few hours, naturalistic studies typically persist over several days and can include a dozen or more sampling occasions. For that reason, naturalistic studies can be more costly, more complex, and more burdensome for participants than structured laboratory studies. Recruitment can be challenging given the time commitment required of participants combined with the need to screen for medication and oral contraceptive use. Ensuring participants’ compliance with study protocol, which can mean sampling saliva at specific intervals or times of day, is essential given the strong diurnal rhythm of cortisol. Noncompliance with sampling time instructions has been linked with significant differences in cortisol patterns (Kudielka, Broderick, & Kirschbaum, 2003). Participants’ behaviors, like eating, drinking, smoking, exercising, and consuming medication or caffeine, can also alter cortisol levels (Kirschbaum & Hellhammer, 1989, 1992), but unlike in the laboratory, where participants’ behavior can be more accurately monitored, participants in naturalistic studies often self-report their adherence to study protocol.

Given the many minor occurrences that can alter cortisol levels, and the pulsatile secretion of cortisol from the adrenal gland (Young, Abelson, & Lightman, 2004), it is reasonable to expect an individual’s cortisol to show many small spikes across the day. However, without sufficient sampling occasions over the course of the day, it is difficult to put the spikes in the context of a normal diurnal cortisol rhythm. Insufficient sampling occasions can also obscure individual differences in the diurnal slope of cortisol, which appears to vary from person to person; as discussed in the following sections, some people have fairly “flat” cortisol rhythms that change little over the course of the day, while others may have steep cortisol slopes, in which high morning levels show a strong diurnal decline (Smyth et al., 1997; Ice, Katz-Stein, Himes, & Kane, 2004). Additionally, laboratory studies have found difference in women’s cortisol responses that are tied to menstrual cycle phase (Kudielka & Kirschbaum, 2005; Kajantie & Phillips, 2006) but, despite a couple of exceptions (e.g. Davydov et al., 2005), most naturalistic studies have not taken women’s menstrual phase into account, given the additional recruitment challenge and the complexity of considering cycle phase in a study that may extend over several days or more.

Analytical challenges. Naturalistic cortisol studies that include multiple saliva sampling occasions are well-suited to statistical approaches such as latent state-trait analysis (e.g. Shirtcliff et al., 2005) and multilevel modeling (Hruschka, Kohrt, & Worthman, 2005), which can represent both state-level (within-person) and trait-level (between-person)
contributions to the variance in cortisol. However, before these statistical approaches became widely available, many early studies failed to adjust for the interdependency of multiple samples within individuals, or aggregated cortisol data collected across multiple timepoints, potentially distorting the effects of diurnal change.

On a more conceptual level, researchers have sometimes failed to distinguish appropriately between cortisol indices reflecting reactivity – that is, state-level fluctuations in cortisol across situations – and basal, or trait-level, cortisol patterns or output. Conflation of these levels can lead to misleading results. For example, an overworked person may excrete lower basal cortisol than a person with a lighter workload, but that same individual may show higher-than-average cortisol during busy times at work. A within-person approach is typically employed in studies of acute laboratory stress and in naturalistic studies of momentary stress (e.g., van Eck et al., 1996; Smyth et al., 1998; Adam, 2005). In contrast, studies of basal cortisol patterns, such as diurnal cortisol slope and total cortisol output, have often been more likely to focus on between-person comparisons (e.g. Giese-Davis, Sephton, Abercrombie, Durán, & Spiegel, 2004; Miller, Cohen, & Ritchey, 2006; Steptoe et al., 2003), although there are exceptions to this rule (e.g. Adam et al., 2006). Lack of clarity about the appropriate level of analysis may account for some of the difficulties researchers have experienced in translating between acute and chronic stress research findings. For example, as discussed below, the finding that chronic stress may be associated with low cortisol output appears to conflict with evidence that cortisol increases in response to momentary stress, but these findings concern two aspects of HPA axis activity that may not only not parallel each other, but may serve different physiological functions altogether (Sapolsky et al., 2000).

Null or conflictual results. The challenges of collecting and analyzing naturalistic cortisol data help to partially explain the shortage of work in this area, but it is also likely that naturalistic cortisol research has suffered from publication bias, or a “file drawer effect.” Given that cortisol responses to everyday stressors tend to be subtle, and easily obscured by sampling error or statistical noise, many studies, especially those with small samples or limited sampling occasions, may have failed to find consistent patterns of results. Additionally, some legitimate findings may have been neglected because they contradicted prevailing expectations. For example, as Miller, Chen, and Zhou (2007) report, early physiological researchers found chronic stress to be associated with reduced overall cortisol output, but work in this area languished due to this finding’s incompatibility with animal models suggesting that stress should heighten glucocorticoid excretion. Only when studies emerged linking certain types of chronic stress with higher average cortisol output did cortisol research become reinvigorated. However, mixed findings have continued to emerge, creating confusion about whether high or low cortisol output is “desirable,” given that both hypercortisolism (high cortisol across the day) and hypocortisolism (chronically low cortisol) have been associated with chronic stress and psychosocial dysfunction (Burke, Davis, Otte, & Mohr, 2005; Gunnar & Vazquez, 2001; Fries, Hesse, Hellhammer & Hellhammer, 2005). These inconsistencies may be due to a number of factors, including participant characteristics and study methodologies. In terms of participant characteristics, specifically clinical sequelae, hypocortisolism has been linked with post-traumatic stress disorder (Yehuda, Giller, Southwick, Lowy, & Mason, 1991), and other disorders (e.g., chronic pelvic pain, fibromyalgia) that have been associated with adverse early life experiences (Gunnar & Vazquez, 2001), while hypercortisolism has been most consistently found in clinical depression (e.g., Burke et al., 2005). The timing of a life stressor might be important, and may be the key factor in differentiating the “signature” cortisol patterns of
different sequelae. Miller and colleagues (2007) suggest that diurnal cortisol output may increase in the presence of ongoing or recently experienced trauma or loss, but decrease, even below “normal” levels, as a stressor becomes more distal. So, for example, a young child in a high-conflict home might show heightened daily cortisol excretion, but that same child might exhibit chronically low cortisol in adolescence or adulthood.

Methodological considerations may also play a role, particularly the timing of sample collection. As discussed further in the following sections, diurnal cortisol slope has increasingly emerged as an outcome with predictive value for health (Sephton, Sapolsky, Kraemer, & Spiegel, 2000; Bower et al., 2005). The daily pattern of cortisol (a post-awakening peak followed by a steep decline throughout the morning hours, then a plateau before reaching an evening nadir) suggests that high and low cortisol may have different implications at different times of the day; a steep diurnal rhythm of cortisol, typically seen as “healthy” (Miller, Chen, & Zhou, 2007) would be marked by high cortisol levels in the morning and low levels in the afternoon and evening. Therefore, a study that collected more morning cortisol samples might find psychosocial difficulties linked with lower-than-average cortisol, while a study that emphasized afternoon or evening collection timepoints could discover the opposite pattern. Insights like these are necessary in order to contextualize apparently conflicting patterns of results, but within the HPA axis literature, research has often outpaced theoretical development.

Given the ambiguous and even contradictory nature of cortisol researchers’ pronouncements about maladaptive or adaptive cortisol patterns, it is perhaps no wonder that many cortisol studies have yielded difficult-to-interpret results. A meta-analysis of 14 studies of occupational stress and cortisol (Hjortskov, Garde, Orbæk, & Hansen, 2004) found equivocal results, with some studies predicting a positive and some a negative association between work stress and cortisol levels; the authors concluded that there is insufficient evidence for an association between self-reported psychological stress in the workplace and cortisol output.

In addition to confusion about the interpretation of results, lack of statistical power may be another source of null findings. In their meta-analysis of HPA axis functioning in populations experiencing high chronic stress, Miller, Chen, & Zhou (2007) argue that, with an average of 80 participants per study, most of the between-subject studies they cited were underpowered; with effect sizes in the small to medium range (.20-.50), twice as many participants would be needed to show reliable effects. By this standard, many between-subject studies focusing on “normal” everyday stress—for example, stress related to the workplace or the family—are insufficiently powered, especially given that normative work- or family-related stress may affect the HPA axis less than the more aversive stressors (like war, trauma, and bereavement) cited by Miller and colleagues.

Selection of appropriate outcome variable

One challenge of naturalistic cortisol research is the difficulty of choosing an appropriate outcome variable, given that there are multiple ways to represent cortisol, and a lack of consensus within the literature about what outcomes are most meaningful to health. However, focusing on an inappropriate outcome, or improperly modeling the outcome (e.g., using insufficient sampling occasions) can lead to misleading results. For example, 25% of the chronic stress studies reviewed in the Miller, Chen, & Zhou (2007) meta-analysis only examined morning cortisol, most of them relying on a single sampling occasion. Given the variability of cortisol across the day, and the many phenomena that can influence cortisol, a single measurement of cortisol, especially during the volatile
morning hours, does not accurately capture HPA axis functioning. In fact, different cortisol outcomes may be only weakly associated with each other; for example, in one study (Edwards, Clow, Evans, & Hucklebridge, 2001), the size of the cortisol awakening response was uncorrelated with total secretion of cortisol across the day. Additionally, there is considerable variability in the measurement of cortisol and the computation of key cortisol parameters, as described in the relevant sections below.

Major findings to date

Most published studies that sample cortisol outside the laboratory focus on at least one of five general types of outcomes: 1) the size of the cortisol awakening response (CAR) (a spike, followed by a rapid decline, which typically occurs within the first hour after waking); 2) the overall shape and pattern of cortisol across the day, as indexed by measures of diurnal slope (that is, the flatness and steepness of the drop in cortisol from morning to evening); 3) area under the curve, a calculation of the total cortisol excreted during the day, and other measures of average or total cortisol output; 4) the momentary (state-level) reactivity of cortisol, including within-person variability in cortisol levels; and 5) a focus on specific values or time periods (not including the CAR), such as evening cortisol levels or the drop in cortisol from afternoon to evening. Here, and in Table 1, we briefly summarize major research findings and methodological considerations within each of these broad categories.

1) Cortisol Awakening Response

Salivary cortisol typically shows a rapid rise upon awakening, peaking within the first 30-45 minutes after waking before beginning its diurnal decline. This rise, or Cortisol Awakening Response (CAR; also sometimes described as the ACR, or Awakening Cortisol Response), was first identified as a useful index of HPA activity by Pruessner and colleagues (Pruessner, Kirschbaum, & Hellhammer, 1995). The CAR appears in about 75% of healthy participants and its size and timing appear moderately intra-individually stable over several days or weeks (Pruessner et al., 1997, 1999; Wust et al., 2000), although there is recent evidence that day-to-day changes in experience or affect may affect the CAR (e.g., Adam et al., 2006). A twin study also found evidence for heritability of the CAR (Wust, Federenko, Hellhammer, & Kirschbaum, 2000). However, despite apparent consistency within individuals, a review of 12 published studies focusing on the CAR found wide variation across studies in absolute values of cortisol, even when sampled at comparable post-awakening times, perhaps because of differences in hormone assay procedures or sampling instructions (Clow, Thorn, Evans, & Hucklebridge, 2004). The CAR appears to be highly sensitive to sampling conditions and participant compliance, and poor adherence to study protocol has been associated with a significantly attenuated CAR (Kudielka et al., 2003; Broderick et al., 2004). In a study that used sleep actigraphy to measure objective waking time, participants who sampled “waking” saliva after more than a 15 minute delay had significantly higher values, which would lead to underestimation of the CAR (Dockray, Bhattacharyya, Molloy, & Steptoe, 2008). The CAR may also be larger among earlier risers (Edwards, Evans, Hucklebridge, & Clow, 2001; Kudielka, Federenko, Hellhammer, & Wüst, 2006; Kudielka & Kirschbaum, 2003; Federenko et al., 2004) although some studies have reported no associations between the size of the CAR and awakening time (Pruessner et al., 1997; Wust et al., 2000; Brooke-Wavell et al., 2002; Kunz-Ebrecht, Kirschbaum, Marmot, & Steptoe, 2004). A study in which women were forcibly awakened during the night and morning found little cortisol response to waking during the first half of the night but a significant waking response in the morning before getting out of
<table>
<thead>
<tr>
<th>Outcome</th>
<th>Major findings</th>
<th>Methodological issues</th>
<th>Sample citations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cortisol Awakening Response (CAR)</td>
<td>↑ CAR associated with loneliness, social stress, lower employment grade and SES, overweight, depression; ↓ CAR associated with burnout, health problems, early loss, material hardship, noise exposure, depression. May be larger on weekdays and in advance of exams. Appears moderately heritable. Women may have larger/more sustained CAR.</td>
<td>Moderate demands on participants; compliance important; sleep quality, duration, time of waking may be meaningful</td>
<td>Steptoe, et al. 2004; Schultz et al. 1998; Wust et al. 2000, Wright &amp; Steptoe, 1995, Wallerius et al., 2003; Pruessner et al., 1999; Kudielka &amp; Kirschbaum, 2003; Nalini et al., 2005; Waye et al., 2003, Pruessner et al., 2003, Stetler &amp; Miller, 2006, Kunz-Ebrecht et al., 2004, Schlotz et al., 2004l</td>
</tr>
<tr>
<td>Diurnal cortisol slope</td>
<td>“Flat” slope observed among parents of children with cancer, veterans with PTSD, the unemployed, women reporting relationship distress or lower social support; steeper slopes associated with marital satisfaction, posttraumatic growth, non-Hispanic whites. Flatter slope linked with more adverse health outcomes and earlier mortality</td>
<td>High demands on participants; compliance important, particularly with first sample(s). Possible gender effects and predictors of within-person changes in slope both under-explored</td>
<td>Miller et al., 2006; Lauc et al., 2004, Giese-Davis et al., 2004, Ockenfels et al., 1995, Szösgren, et al., 2006, Abercrombie et al., 2004, Adam &amp; Gunnar, 2001, Vedhara et al., 2006; Gallager-Thompson et al., 2006, McCallum et al., 2006, Sephton et al., 2000, Bower et al., 2005, Matthews et al., 2006</td>
</tr>
<tr>
<td>Area under the curve/total cortisol output</td>
<td>↑ average cortisol associated with early loss (men), high perceived chronic stress, mothers with children at home, anxiety symptoms (men); higher job status (women); ↑ AUCg associated with mastery (women with breast cancer); ↓ AUCI correlated with distress (women with breast cancer); ↓ average cortisol associated with work-related stress (women), higher job status (men)</td>
<td>Moderate-to-high demands on participants; approach does not consider diurnal rhythm of cortisol, has not been conclusively linked with health outcomes</td>
<td>Nicolson, 2004; Luecken et al., 1997; van Eck &amp; Nicolson, 1994; Kurina et al., 2004; Edwards et al., 2003; Pruessner et al., 2003; Vedhara et al., 2006; Matthews et al., 2006; Sephton et al., 2000; Turner-Cobb et al., 2000; Steptoe et al., 2003</td>
</tr>
<tr>
<td>Outcome</td>
<td>Major findings</td>
<td>Methodological issues</td>
<td>Sample citations</td>
</tr>
<tr>
<td>----------------------------------------------</td>
<td>-------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------</td>
</tr>
<tr>
<td>Cortisol response to momentary experience</td>
<td>↑ momentary cortisol associated with experience of concurrent stressor, anticipation of future stressor, negative mood; negative mood may mediate stress-related rise in cortisol, ↓ momentary cortisol associated with positive moods. rise with; ↑ within-subjects variability in cortisol associated with depression, PTSD</td>
<td>High demands on participants; “reactivity” appears variable among participants, may be stronger for men, blunted in depression; temporal proximity of experience to saliva sampling important</td>
<td>van Eck et al., 1996; Smyth et al., 1998; Adam, 2005; Hanson et al., 2000; Polk et al., 2005; Nicolson, 1992; Peeters et al., 2003; Peeters et al., 2004; Yehuda et al., 1996</td>
</tr>
<tr>
<td>Cortisol at specific time periods</td>
<td>↑ evening cortisol among divorcing women, women experiencing financial strain; ↓ evening cortisol associated with higher income and education level, and after a more stressful workday</td>
<td>Low demands on participants; few significant results due to signal-to-noise ratio of cortisol</td>
<td>Powell et al., 2002, Grossi et al., 2001, Cohen et al., 2006, Saxbe et al., in press</td>
</tr>
</tbody>
</table>
bed (Dettenborn, Rosenloecher, & Kirschbaum, 2007). While it has been assumed by most researchers that the CAR represents a response to awakening, rather than the circadian increase in HPA axis activity during the early morning hours, this assumption was only recently tested by a laboratory study that sampled serum cortisol during and after sleep, finding that the CAR does indeed seem to be triggered specifically by awakening (Wilhelm, Born, Kudielka, Schlottz, & Wüst, 2007).

In an illustration of the inconsistency of CAR research to date, chronic psychological stress and adverse health outcomes have been linked both negatively and positively with the size of the CAR. For example, loneliness, work overload, social stress, life regrets, type-D (depressive) personality, and a lack of social recognition have all been associated with a larger-than-usual CAR (Steptoe, Owen, Kuntz-Ebrecht, & Brydon, 2004; Schultz et al. 1998; Wrosch, Bauer, Miller, & Lupien, 2007; Whitehead, Perkins-Porras, Strike, Magid, & Steptoe, 2007; Wust et al. 2000) as has a lower employment grade among government employees (Kunz-Ebrecht et al., 2004) and lower socioeconomic status (Wright & Steptoe, 1995). A smaller CAR has been linked with reductions in financial strain (Steptoe, Brydon, & Kunz-Ebrecht, 2005), and with greater positive affect (Evans et al., 2007; Steptoe, Gibson, Hamer, Wardle, 2007). The CAR also appears to be larger among overweight men, although results have been equivocal for women (Wallerius et al., 2003; Steptoe, Kunz-Ebrecht, Brydon & Wardle, 2004). Among normal-weight young men (but not women), the CAR was positively associated with fathers’ adiposity, suggesting that abnormalities in cortisol may proceed weight problems rather than vice versa (Steptoe, Wright, O’Donnell, Brydon, & Wardle, 2006).

In the other direction, burnout, self-reported health problems, early loss experiences (such as the death of a parent), high material hardship, depressive rumination, amnesia, and nighttime exposure to low-frequency noise have all been linked with an attenuated CAR (Pruessner et al., 1999; Kudielka & Kirschbaum, 2003; Meinschmidt & Heim, 2005; Ranjit, Young, & Kaplan, 2005; Kuehner, Holzhauer, & Huffziger, 2007; Waye et al., 2003; Wolf, Fujiwara, Luwinski, Kirschbaum, & Markowitsch, 2005), while some positive indicators of health (such as bone density and a measure of cardiovascular fitness) have been associated with a larger CAR (Brooke-Wavell et al., 2002; Eller, Nettetstrom, & Hansen, 2001). Clinical depression has been associated with a larger CAR (Pruessner, Hellhammer, Pruessner, and Lupien, 2003), although a blunted CAR characterized depressed women in an all-female community sample (Stetler & Miller, 2006), and a study that compared depressed and non-depressed psychiatric inpatients found the depressed patients to have a comparatively blunted CAR (Huber, Issa, Schik, & Wolf, 2006). Similarly, clinical burnout and burnout symptom severity have been associated both with a smaller CAR (e.g., Sonnenschien et al., 2007), a larger CAR (de Vente et al., 2003; Grossi Perski, Ekstedt, & Johansson, 2004), and no difference in CAR (Mommersteeg, Heijnen, Verbraak, & Van Doornen, 2006), suggesting that assessment, measurement, and compliance issues may lead to discrepant results.

A cognitive-behavioral stress management intervention was associated with a larger CAR on the morning of an important exam (Gaab, Sonderegger, Scherrer, & Ehlert, 2006), supporting the “boost hypothesis” (Schultz et al., 1998; Adam, Hawkley, Kudielka, and Cacioppo, 2006), which suggests that the post-awakening rise may reflect an adaptive effort to cope with anticipated challenges. The weekday cortisol response to awakening appears to be larger than the weekend response, and this difference is exaggerated when participants report greater work overload and worry, supporting the contention that anticipation of future challenge may affect the size of the CAR (Kunz-Ebretch et al., 2004; Schlottz, Hellhammer, Schulz, & Stone, 2004). However, many early studies of the CAR did...
not distinguish between weekday and weekend sampling occasions, clouding past research results (Clow et al., 2004). Nevertheless, given this preliminary evidence that the CAR may reflect a mustering of physiological arousal to cope with anticipated challenges, one theoretical possibility deserving further study is that the CAR reflects perceptions of controllability regarding the source of chronic stress. Participants who show signs of learned helplessness or burnout may therefore show a smaller CAR, a contention supported by some of the studies reviewed above.

Several studies have reported a larger and more sustained CAR among women than men (e.g. Pruessner et al., 1997, 1999; Wright & Steptoe, 2005), but other studies have failed to find gender differences (Edwards et al., 2001; Kudielka & Kirschbaum, 2003). Oral contraceptive use has been associated with an attenuated CAR (Pruessner et al., 1997, 1999), but the CAR does not appear to be linked with menstrual phase (Kudielka & Kirschbaum, 2003). Among pregnant women (de Weerth & Buitelaar, 2005), the absolute value of the mean increase in cortisol after awakening was greater, but the relative mean increase over the waking value was the same – about 40% in both cases. Age has not been consistently associated with the CAR (Pruessner et al., 1997; Wüst et al., 2000; Edwards et al., 2001), but the CAR does appear to be moderately heritable (Wüst et al., 2000; Clow et al., 2004).

In summary, the CAR appears to be a cortisol pattern with decent intra-individual stability and documented linkages to mental and physical health functioning. It appears that both a larger-than-average and a smaller-than-average CAR may signal poor adaptation to chronic stress. These inconsistencies may be due to several of the factors mentioned above: for example, study compliance, waking time and anticipation of waking, and the anticipation and controllability of stress. For example, a study that found a blunted CAR among amnesiac patients (Wolf, et al., 2005) mentioned several important differences between patients and controls: the controls woke up earlier than the patients, and the controls were responsible for their own awakening, while the patients were woken by their caregivers. Since, as described above, both an earlier time of awakening and sleep conditions before waking have been linked with a larger CAR, this study provides a good example of one in which methodological considerations may have shaped an apparent link between patient status and a blunted CAR. It is likely that sleep quality and duration influence the CAR, as does waking time, but more research is needed to establish clear effects. The apparently wide range in absolute values of cortisol across studies of the CAR (Clow et al., 2004) provides a clue to the mixed results that have reported in the literature, suggesting that the size of a “normal” CAR might depend on the comparison group. A robust CAR might be generally adaptive, but stressful circumstances may elevate the CAR above the ideal threshold; among participants experiencing typically high chronic stress, such as nurses, the unemployed, or even undergraduates, a lower-than-“average” CAR might be a marker of better functioning relative to the group, while, in a community sample of participants with low average stress, a below-average CAR might signal a blunted response to awakening. Many of the CAR studies cited above explored between-subject effects; there is still little understanding of what might cause the CAR to fluctuate on the within-person level (for example, day-to-day changes in workload or social experiences). One intriguing study found that feelings of loneliness, sadness, and threat were not associated with the size of the CAR that day, but predicted a larger CAR the following day (Adam et al., 2006). More exploration of possible causal associations between stressful experiences, mood states, and the CAR is needed.

The time demand placed on participants by a study focusing on the CAR is moderate: multiple sampling occasions are required, over at least two to six days (Hellhammer, et al.,
2007), but only during the first hour or two following awakening. However, participants’
compliance with study protocol is especially important; whether a participant samples
saliva immediately upon awakening or fifteen minutes later may affect the overall shape
and size of the CAR estimate, so researchers must educate participants carefully about the
time-sensitive nature of the hormone. Many studies measure the CAR at just two
timepoints – upon waking and 30 minutes after waking – but more samples are advisable,
especially with female participants, who may peak later than males and show a more
delayed decrease in cortisol (Wust et al., 2000). Using actigraphy to improve monitoring of
participants’ sleep quality and waking time can be costly, but allows for more accurate
assessment of the CAR. Saliva sampling kits that include some kind of time stamp (such as
electronic monitoring caps on collection vials) may also be helpful, although these tools
cannot detect actual participant waking time.

2) Diurnal cortisol slope

Cortisol shows a strong diurnal rhythm, peaking within the first hour after awakening,
declining rapidly over the morning hours, and then tapering off over the rest of the day
before reaching a nighttime nadir. Inter-individual variability in cortisol slope has been
reported; in two recent studies, 30-50% of participants produced flattened or inconsistent
slopes (Smyth et al., 1997; Ice, Katz-Stein, Himes, & Kane, 2004), while another study
found flat slopes in over 10% of participants (Stone et al, 2001). The flat slope, which may
be characterized by low morning levels, an attenuated drop in cortisol levels over the
afternoon and evening, or both, appears to be associated with a high chronic stress burden
and poor psychosocial functioning. A recent meta-analysis focusing on the HPA axis
found that chronic stress appears to be consistently associated with a flatter diurnal rhythm
of cortisol and with its correlates, lower morning cortisol values and higher afternoon and
evening values (Miller, Chen, & Zhou, 2007). Interestingly, this review also found that the
low morning levels associated with a flattened diurnal rhythm were linked with the
perceived uncontrollability of the stressor; when it was seen as controllable, morning levels
were likely to be higher, as if the body were mustering its resources to confront the source
of stress. Therefore, a flat rhythm that includes low morning levels may indicate not only
high stress but low perceived control, such as might be seen among the parents of children
with cancer (Miller, Cohen, & Ritchey, 2006) or workers using highly automated
production systems (Garde et al., 2003), both of whom had flatter-than-usual slope than
a comparison group. Flattened slopes may also be predictive of eventual health outcomes:
for example, among women with breast cancer, flat cortisol cycles were linked with earlier
mortality (Sephton et al., 2000), and among breast cancer survivors, flattened slope was
associated with persistent fatigue (Bower et al., 2005). Flatter slopes have also been linked
with coronary calcification (Matthews, Schwartz, Cohen, & Seeman, 2006) and body mass
index, although the latter link may be mediated by education and positive affect (Daniel
et al., 2006).

Flattened cortisol slopes have been observed in male veterans with PTSD (Lauc,
Zvonar, Vuksic-Mihaljevic, & Flogel, 2004), women high in repressed or anxious coping
(Giese-Davis et al., 2004), unemployed men and women (Ockenfels, Porter, Smyth, &
Kirschbaum, 1995), men and women with high ratings of cynicism, depression, and
exhaustion (Sjögren, Leanderson, & Kristenson, 2006), and women reporting high stress
and low perceived social support (Abercrombie et al., 2004). Among middle-class mothers
of toddlers, women with more insecure attachment styles and less rewarding marital
relationships had flatter cortisol slopes (Adam & Gunnar, 2001). At least two other studies
have found the steepness of the diurnal slope to be positively correlated with marital
satisfaction among women (Vedhara, Miles, Sanderman, & Ranchor, 2006; Saxbe, Repetti,
A study that sampled both men and women (Barnett, Steptoe, & Gareis, 2005) found that participants who reported higher marital-role concerns had a significantly flatter cortisol slope over the day.

Some racial or ethnic differences in cortisol slope have been reported. For example, flatter slopes have been reported among Hispanic women compared to non-Hispanic women matched on age and education level (Gallager-Thompson et al., 2006) and African-American women compared to European-American women (McCallum, Sorocco, and Fritsch, 2006). Another study of African-American, Latino, and Caucasian men and women found a tendency towards steeper slopes among non-Hispanic whites and women (Adam et al., 2006). In a sample of older adolescents, both African-American and Latino youths had flatter cortisol slopes, with African-American males having flatter slopes than females (DeSantis et al., 2007), even when socioenvironmental factors were controlled. Few other examinations of gender differences in cortisol slope have been reported, perhaps because, oddly, a majority of the published studies focusing on slope have sampled only women (e.g. Abercrombie et al., 2004; Bower et al., 2005; Giese-Davis et al., 2004; Adam & Gunnar, 2001; Vedhara et al., 2006; Moskowitz & Epel, 2006; Sephton et al., 2000; Gallagher-Thompson et al., 2006; Garde et al., 2003; Daniel et al., 2006). This gender discrepancy, especially striking because acute stress studies have tended to oversample males, may reflect the interests of researchers who have focused on diurnal slope (for example, Sephton, Giese-Davis, Bower, and Abercrombie all report on slope in the context of breast cancer, an illness which disproportionately affects women). Another possibility, albeit an untested one, is that women’s diurnal rhythms may also be especially sensitive to psychosocial factors, and all-male or mixed-gender studies of cortisol slope may have languished in the file drawer. To that end, a recent study of couples found marital satisfaction and diurnal slope may have languished in the file drawer. To that end, a recent study of couples found marital satisfaction and diurnal slope were significantly associated for women but not men (Saxbe et al, 2008), and a follow-up study found that couples’ descriptions of the home environment predicted wives’ cortisol slope but not husbands’ (Saxbe & Repetti, in preparation).

The link between diurnal slope and positive emotions has received relatively little research attention. One study of maternal caregivers (Moskowitz & Epel, 2006) found that women with higher levels of posttraumatic growth (that is, positive adaptation to the stress of caregiving) had steeper cortisol slopes, but only if they also reported higher levels of daily positive emotion. Another study found steeper slope to be associated with high ratings of social support and coping, general health, and well-being (Sjogren et al, 2006). But other studies have failed to find associations between positive psychological states and cortisol slope (e.g. Adam et al., 2006). A study of women found steeper slopes on days that included more regular activities and more social contact, supporting the notion that “social zeitgebers” may help to entrain biological circadian rhythms (Stetler, Dickerson, & Miller, 2004).

Flattened slopes can be due to alterations of “typical” HPA axis profiles at several different occasions during the day, for example, low morning levels that remain low throughout the day (a flat, “low” cycle), or high morning levels that fail to show a normal diurnal decrease (a flat, “high” cycle, often characterized by higher-than-expected values in the afternoon or evening). It is likely that these cycle types reflect different physiological underpinnings and have different health consequences. For example, the link between mortality and cortisol slope among women with breast cancer (Sephton et al., 2000) appeared to be due to women with a high, flat rhythm who did not show the expected drop in cortisol in the evening. However, other studies (e.g., Adam & Gunnar, 2001; Saxbe et al., 2008) found associations between impairments in psychosocial functioning and flattened...
slope that appeared to be driven by low morning values rather than high evening values. The majority of studies referenced above did not differentiate explicitly between “flat low” and “flat high” cycles, focusing on the trajectory of change across the day rather than absolute values of cortisol. It is important for researchers to attend not only to daily change in cortisol over the day but also to cortisol levels at the peak and the nadir of the daily cycle. Failure to do so may lead to discrepant results. For example, Polk, Cohen, Doyle, Skoner & Kirschbaum (2005) found that men low in trait positive affect had an attenuated afternoon decrease and a high, flat diurnal rhythm of cortisol, while women high in trait positive affect had low morning cortisol and a low, flat rhythm. Miller, Chen and Zhou (2007) suggest that, as with the morning rise (discussed in the previous section), the controllability of a stressor may be important: the higher morning values consistent with a high, flat rhythm might suggest a challenge that is potentially surmountable, while the low, flat rhythm could indicate stress that is not or no longer controllable, as in the case of post-traumatic stress or burnout.

In summary, diurnal slope appears to be the rare cortisol parameter that has been consistently associated with psychosocial functioning, with flattened slope indicative of higher overall chronic stress, lower perceived control, and poorer relationship functioning. Many of these findings are based on between-subject level predictors, such as those listed above; only a few studies (e.g. Adam et al., 2006; Stetler et al., 2004) have explored within-person changes in diurnal slope (for example, whether a flatter slope occurs on or after more stressful days). Additionally, many published slope studies use only female participants, although it is unclear whether men tend to show weaker associations between diurnal slope and other variables.

Studies focusing on diurnal cortisol slope place fairly high demands on participants, usually requiring at least four to six saliva samples to be collected across each day for at least two or three days (one day’s worth of data is usually insufficient for calculating diurnal slope; Kraemer et al., 2006). The statistical computation of cortisol slope varies, but most often includes regressing cortisol level by time since waking, sometimes in order to use the resulting coefficient as a dependent variable (e.g. Kurina, Schneider, and Waite, 2004), but more often as part of a hierarchical linear model with between-subjects predictors at Level 2 or day-level predictors at Level 2 and between-subjects predictors at Level 3 (e.g. Adam et al., 2006). Since the shape of a typical diurnal slope is actually curved (with cortisol levels decreasing sharply over the morning hours and then tapering off over the rest of the day), a more accurate regression equation may include coefficients for time and time$^2$ (Adam, 2005; Vedhara et al., 2003).

3) Total cortisol output/area under the curve

Many early cortisol studies focused on aggregate or average measures of cortisol excretion over the course of the day. Perhaps unsurprisingly, given that this approach does not take the diurnal rhythm of cortisol into account, these studies have yielded few consistent results, particularly in non-pathological populations (see discussion of hypercortisolism and hypocortisolism above). A few exceptions are mentioned here. One study found that men who had suffered the early loss of a parent had higher daily average cortisol (Nicolson, 2004) than men who had not. Another study of working women found higher 24-hour excretion of urinary cortisol among mothers with children living at home (Luecken et al., 1997). A study comparing individuals experiencing high or low levels of chronic stress found higher average cortisol excretion among the high-stress group (van Eck & Nicolson, 1994). Among working parents (Kurina, Schneider, & Waite, 2004), average cortisol levels were lower among women who felt stressed at work and higher among men who felt anxious, but another study of workers found a conflicting pattern of
gender differences, with higher job status associated with lower average workday cortisol levels in men and higher workday levels in women (Steptoe et al., 2003). An alternative to summing or averaging cortisol levels is to compute “area under the curve” (AUC), an estimation of total cortisol output performed by trapezoidal calculation. It appears that, overall, a higher total AUC is associated with chronic stress (Miller, Chen, & Zhou, 2007), but studies using the AUC have reported inconsistent results (e.g. Edwards, Hucklebridge, Clow, & Evans, 2003), and there has not been full consensus among researchers about how to estimate the AUC. AUC can be calculated in several different ways, and some researchers have advocated computing a basal AUC (“AUC with respect to ground,” or AUCg) or, less commonly, an AUC that captures the reactivity in cortisol across the day (“AUC with respect to increase,” or AUCi) (Fekedulegn et al., 2007; Pruessner, Kirschbaum, Meinlschmid, and Hellhammer, 2003). AUCi is less often used by researchers, in part because it may be confounded with the CAR. AUCg was linked to childhood history of Upper Respiratory Illnesses in one study, while AUCi was not (Vedhara et al., 2007). A study that examined both indices (Vedhara et al., 2006) found AUCg to be positively associated with personal mastery among women with breast cancer, while AUCi was negatively correlated with distress. The study also found surprisingly few consistencies between AUCg, AUCi, diurnal slope, and CAR, suggesting that these outcome measures may be tap into different aspects of HPA axis physiology. Another study found diurnal cortisol slope to be linked with coronary calcification and psychosocial adjustment in a large, diverse sample, while AUC was not (Matthews et al., 2006). Similarly, while diurnal cortisol slope has been linked with breast cancer mortality, AUC does not appear to be associated with mortality (Sephton et al., 2000; Turner-Cobb et al., 2000).

4) Relationship between momentary events and cortisol

In additional to overall patterns across the day, researchers have looked at within-person covariation between moods, experiences, and cortisol levels, borrowing from approaches like Experience Sampling Methodology (Larson & Csikszentmihalyi, 1983) and Ecological Momentary Assessment (Stone & Shiffman, 1994; Smyth & Stone, 2003). Within such paradigms, multiple observations are collected from the same participant over several hours, days, or weeks. Unlike the studies discussed above, which focused on indices like CAR, cortisol slope, and AUC mostly in conjunction with between-person or “trait” factors, these studies emphasize “state,” the changes in cortisol and psychological well-being across the day within individuals. A comparison of trait and state level variance in a sample of adolescents suggested that, after time of day effects were controlled, trait factors accounted for about 30% and state factors for about 68% of the variance in cortisol, with the remaining variance attributable to error (Shirtcliff et al., 2005). However, the relative contributions of trait and state have not been explored extensively in adults. Interestingly, higher variability of repeated cortisol samples has been found in depression (Peeters, Nicolson, & Berkhof, 2004) and PTSD (Yehuda et al., 1996), suggesting that erratic patterns of within-person cortisol release may themselves indicate risk.

Several studies have reported that stressful experiences are associated with higher-than-expected cortisol when time of day effects are controlled, and that negative mood states may also contribute to elevated cortisol (e.g. van Eck et al., 1996; Smyth et al., 1998; Adam, 2005; Hanson, E.K., Maas, C.J., Meijman, T.F., Godeart, G.L., 2000). This finding parallel the results of laboratory studies such as the TSST, which typically find a stress-related increase in cortisol (Dickerson & Kemeny, 2004). For example, one early study found that cortisol levels increased among college students anticipating stressful examinations, and that the cortisol increases were associated with the students’ negative...
mood states (Nicolson, 1992). A study of white-collar males (van Eck et al., 1996) found both recent stressful events and negative mood states to be associated with higher cortisol levels, with the mood states apparently mediating the association between stressful experiences and cortisol. Similarly, an ESM study found that daily stressors predicted increased cortisol, with effects apparently mediated by negative affect (Jacobs et al., 2007), while another study found momentary rises in cortisol to be linked with feelings of performance pressure during daily tasks, with trait anxiety (but not state negative affect) playing a mediating role (Schlotz, Schulz, Hellhammer, Stone, & Hellhammer, 2006). A follow-up to the van Eck study (Smyth et al., 1998) found that both the experience of a current stressor and the anticipation of a future stressor were associated with higher-than-expected cortisol sampled 20 minutes later, but that having experienced a stressful event within the last two hours was not associated with any change in cortisol. As with the van Eck and Jacobs studies, affect appeared to mediate the link between stressful experiences and cortisol, with shifts in negative affect associated with higher cortisol.

Interestingly, Smyth and colleagues also found considerable individual differences in the size of the cortisol response to daily stressors, with about a quarter of participants showing little to no reactivity, and with men showing greater reactivity than women. A study of middle-class parents (Adam, 2005) also found wide variability in the degree to which mood states and cortisol levels were associated, with some individuals experiencing up to a 92% increase in cortisol (over the level expected for that time of day) for each standard deviation increase in negative emotion, and others experiencing little change in cortisol. Although, on average, both mothers and fathers tended to show higher cortisol levels when feeling more irritated and stressed, this effect was much larger for men, suggesting that males’ HPA axis may be more sensitive to negative emotion. Positive moods were associated with lower-than-expected cortisol for both men and women. Finally, a study of both depressed and non-depressed adults (Peeters, Nicholson, & Berkhof, 2003) found negative events and mood states to be associated with higher-than-expected cortisol, but that this reactivity appeared blunted in depressed participants. In summary, the literature exploring state factors and cortisol is still small, but has been fairly consistent, showing negative experiences and emotions to be associated with higher-than-usual and positive affect with lower-than-expected cortisol levels. However, participants’ “reactivity” appears highly variable (van Eck et al., 1996; Smyth et al., 1998; Adam, 2005), with some evidence that depression may attenuate the cortisol response to stress (Peeters et al., 2003).

One challenge for researchers attempting to capture naturalistic stress reactivity is fixing the timing of the stressor and cortisol sampling: in the lab, the interval between stress induction and cortisol measurement can be tightly controlled, but experience sampling or daily diary paradigms typically ask participants to recount experiences that may have occurred within the last several hours, despite evidence that the recency of a stressful experience may be important (Smyth et al., 1998). Additionally, in some studies, cortisol is sampled concurrently with the completion of diary questionnaires (e.g. Peeters et al., 2003), while other studies ask participants to sample saliva after a set interval, such as 20 minutes, to allow time for the HPA axis response to become measurable in saliva (e.g. Smyth et al., 1998; Adam, 2005). Given that many everyday stressors are mild and likely to exert subtle effects on cortisol, variability in stressor-sampling intervals, whether due to participant compliance or to study design, may obscure linkages between cortisol and daily experiences. The finding that mood states appear to be more consistently associated with cortisol may be due in part to this issue of timing: mood states may be more likely to linger and to infuse ongoing behavior, while some stressful experiences may have more discrete
boundaries and be more sensitive to time delays. This contention is supported by the Smyth et al. (1998) study that found current and anticipated stressors, but not recently experienced stressors, to be associated with cortisol increases when saliva was sampled 20 minutes after experience sampling reporting, with effects mediated by negative affect. At the same time, current affect was more strongly related to cortisol than stressful experience in two studies that asked participants to collect saliva concurrently with an experience sampling measure (van Eck et al., 1996; Hanson et al., 2000), suggesting that emotion may have a stronger effect on cortisol levels than the presence of a discrete stressor, regardless of the interval between experience sampling and cortisol sampling.

The timing of stressful experience and cortisol sampling is also relevant in terms of understanding the direction of effects. It is often assumed that stressful experiences trigger changes in cortisol levels, but cortisol levels may also inform participants’ appraisals of mood and stress. Studies of glucocorticoid administration of the lab support this hypothesis, as does recent evidence that cortisol levels are related to subsequent subjective perceptions (Abercrombie, Kalin, & Davidson, 2005). A study of older adults (Adam et al., 2006) found that low morning cortisol levels predicted higher fatigue ratings later that day. Further exploration of the causal associations between mood, daily experiences, and cortisol is warranted.

Given that they often require several days of repeating sampling of both cortisol and experience measures, studies of momentary states and cortisol impose fairly high demands on participants, a participant burden comparable to that of diurnal slope studies. However, these studies offer perhaps the best chance for researchers to capture cortisol reactivity to everyday situations, an issue with relevance to allostatic load and the long-term health effects of normative stressors. In terms of data analysis, state-focused naturalistic cortisol studies typically require the use of a multilevel or latent state-trait model to account for the nesting of moments within days and the nesting of days within participants.

5) Single timepoint or time period studies

Some cortisol studies have focused on specific values or time periods (other than the morning hours), although this approach is limited for reasons described above, e.g. the difficulty of finding effects with few sampling occasions given the volatility and diurnal rhythm of cortisol. In general, high cortisol late in the day and a smaller-than-average drop in cortisol at the end of the day have both been associated with higher stress and fewer psychosocial resources. For example, a comparison of divorcing and stably married women found elevated 9 pm cortisol levels among the divorcing group (Powell et al., 2002). Another study found high financial strain among unemployed individuals to be associated with higher evening cortisol among women, but not among male participants (Grossi, Perski, Lundberg, & Soares, 2001). In a large, racially diverse study (Cohen et al., 2006) evening cortisol levels were lower among participants with higher income and education, and these effects were partially mediated by smoking, social network diversity, depression, and sleep quality. Very few studies have examined the physiology of “unwinding” from work in terms of the HPA axis, in part because it may be difficult to tease apart the effects of transitioning from work to home from the normal diurnal rhythm of cortisol. An early study of Volvo employees found that males showed more of a drop in (urinary) cortisol at the end of the workday than women (Frankenhaeuser et al., 1989). A study examining recovery from work among dual-income couples found that evening cortisol levels were lower than expected after a busier workday, suggesting an apparent “overcorrection” of the HPA axis, and that marital satisfaction appeared to augment this effect among women (Saxbe et al., 2008).
Conclusions and future directions

Researchers’ attempts to incorporate cortisol sampling into everyday life have not always yielded intuitive or elegant results. Indeed, finding reliable associations between cortisol patterns and naturalistic stressors may be an intrinsically limited enterprise: the HPA axis is a dynamic, complex system that self-regulates through multiple short- and long-term adaptations to stress, so it may be no surprise that its end-product, cortisol, is not always linked with subjective experience. That said, the only way to distinguish between the intrinsic limitations and the methodological limitations of HPA axis research is to conduct studies that follow consistent guidelines for data collection, statistical power, and the proper conceptual framing of analyses. Most of the naturalistic cortisol studies to date, including many of those reviewed in this paper, conduct between-subject analyses on samples under 100 participants, and within-subject analyses with protocols that may include only a few days or even a few timepoints of saliva collection. Given that the “signal to noise” ratio of HPA activity in everyday life is likely to be high, finding reliable effects might simply require more data. It is also imperative to screen participants carefully (eliminating or adjusting for possible confounds) and to make design choices that encourage participants’ compliance. It would also behoove researchers to report absolute values of cortisol and begin to develop standard definitions for such ephemeral phenomena as a “flat slope” or a “large CAR.” Cortisol values may vary widely between studies, due to participant selection, study protocols, and sampling or assay conditions (Clow et al., 2004), and the field would benefit from benchmarks to help contextualize results. See Table 2 for a list of additional guidelines for naturalistic cortisol researchers to follow. Ultimately, in order to develop a more fine-grained understanding of how stress and coping influence pathways to physical health, HPA axis research must move (at least partially) out of the lab and into the field.

Several cortisol indices have begun to generate consistent findings, as summarized in Table 1, including diurnal cortisol slope and the Cortisol Awakening Response (CAR). The study of within-person changes in cortisol remains preliminary but offers great potential for building models of stress reactivity in everyday life. The size of the cortisol response to daily stressors appears to show wide variability (van Eck et al., 1996; Smyth et al., 1998; Adam, 2005), suggesting that individual differences in everyday stress responding warrant further study.

New directions

Researchers focusing on day-to-day HPA axis functioning have their work cut out for them, as many areas remain underexplored. For example, laboratory studies have reported that men show greater acute stress reactivity than men (Kudielka & Kirschbaum, 2005; Kajantie & Phillips, 2006), except in particular cases, such as during relationship conflict (Kiecolt-Glaser & Newton, 2001). The question of how the genders diverge in everyday life is particularly interesting given that the sexes often occupy different roles, with men typically spending more hours engaged in paid work but women devoting more time to childcare and domestic duties (Bianchi, Milkie, Sayer, and Robinson, 2000). Despite evidence that women may show attenuated physiological unwinding from work in the evening (Frankenhaeuser et al., 1989), or that women’s recovery from work might be influenced by the marital context (Saxbe et al., 2008), the question of gender differences in everyday HPA axis functioning has not been thoroughly addressed. For that matter, the
physiology of recovery from work or other forms of daily challenge is still not well understood, despite its importance for allostatic load.

Another interesting question for naturalistic cortisol researchers is how and whether HPA axis functioning may be interrelated among families, cohabitating adults, or even coworkers. Most people dwell and work in the close company of others, so the possibility of biobehavioral coregulation has implications for health and allostatic load. Intriguing new research on social zeitgebers suggests that daily activities and social contacts may exert a regulatory influence on diurnal cortisol (Stetler, Dickerson, & Miller, 2004), but little is known about the specific mechanisms underlying this process. There is some evidence that cortisol levels are positively correlated among parents and children, and between spouses (Schreiber, et al., 2006), but the question of how one family member’s physiological stress state might influence others has not been explored in depth. A related issue involves genetic influences on cortisol levels; preliminary research suggests some heritability of HPA axis functioning, perhaps with genetic influences contributing more to variance in cortisol.

Table 2. Eleven simple rules for conducting field research on cortisol

1) Select outcome variable(s) appropriate for phenomena of interest (e.g., normative baseline levels, changes in baseline, or momentary reactivity to stress); clearly define and distinguish within- and between-subjects predictions
2) Conduct power analyses to determine a sufficient ‘n,’ given the typically modest effect sizes found in ambulatory cortisol research
3) If conducting between-subjects analyses, screen and/or standardize for possible confounds (e.g., medical or psychiatric conditions, medications, birth control use, pregnancy, smoking, scheduling of daily routines, frequency of strenuous exercise, alcohol intake, body mass, average intensity of home and work demands). If possible, do not recruit participants taking steroid medications and/or with endocrine disorders.
4) To whatever extent possible, standardize menstrual cycle phase among female participants and/or collect this information (e.g., by measuring sex hormone levels) to use as a covariate in analyses
5) Assess compliance with saliva collection procedures by using a time-stamped collection system (such as MEMScaps) or electronic diary device that records time (like a PalmPilot). Educate participants about the diurnal rhythm of the hormone and the importance of adhering to the study’s saliva sampling schedule.
6) Develop clear protocols regarding participant behaviors that may affect cortisol concentrations, such as the timing of eating and drinking, sleep habits, smoking, taking medications, caffeine intake, exercising, and vigorous flossing or brushing of teeth (which may cause bleeding gums). Consider measuring confounds that may be of particular concern (for example, additional assays to test for blood contamination in saliva, or actigraphic measurements of sleep quality and waking time).
7) Schedule sufficient sampling occasions to measure the phenomena of interest (at least 3-6 samples per day collected over at least 2 days) and make sure sampling occasions are appropriately spaced (e.g. every 15 minutes over the first two hours after waking in a study of the CAR, every four hours in a study of diurnal slope)
8) Use a consistent procedure for the storage, shipping, and assaying of cortisol
9) Use a statistical approach (e.g., multilevel modeling, latent state-trait modeling) that adjusts for the nesting of cortisol samples within participants, within days and, if applicable, within families, classrooms, and worksites; and that measures the contribution of both state-level and trait-level variance to cortisol
10) Expect missing data and skipped or corrupted sampling occasions, and plan accordingly when calculating power and developing a data analysis strategy
11) Rather than simply comparing participants’ cortisol levels to each other, report absolute values of the hormone to allow comparison with other studies using similar participants
during the morning hours and environmental factors becoming more important in the afternoon and evening (Bartels et al., 2003; Wust et al., 2000; Kupper et al., 2005, Schreiber et al., 2006).

Sleep has also been understudied by HPA axis researchers. Awakening time and nighttime noise exposure have both been linked with the size of the cortisol awakening response, and it is likely that sleep quality and duration play a role in determining both the CAR and the subsequent patterning of cortisol across the day. A recent study from the experimental literature supports the possibility of a link between sleep and cortisol, finding blunted responses to an acute stressor after a night of poor sleep quality (Wright, Valdimarsdottir, Erblich, and Bovbjerg, 2007). Given that both work and parenting demands can place adults at risk for sleep deprivation and disturbance, it is important to understand whether and how sleep might mediate previously observed linkages between psychosocial adjustment, chronic stress, and HPA axis functioning.

Most of the research reviewed in this paper concerns the influence of stressful experience and affect on cortisol, but this influence is likely to be both bidirectional and reciprocal; in other words, cortisol levels may inform subsequent perceptions and behaviors (Abercrombie, Kalin, & Davidson, 2005; Adam et al., 2006), which may in turn feed back to affect the HPA axis. For example, the “boost hypothesis” contends that high morning cortisol levels represent a mustering of energy levels in preparation for a challenging day (Adam et al., 2006). In addition to identifying predictors of changes in cortisol, researchers ought to study the impact of cortisol levels on subjective well-being, an intriguing issue with implications for health and psychosocial adaptation. Relatedly, little is known about whether and how stress-reducing and psychotherapeutic interventions can alter cortisol patterns outside the lab, despite a few exceptions (e.g. Gaab et al., 2006). Intervention research might help to clarify the clinical significance of different cortisol parameters and their intra-individual stability.

The importance of context

While it is understandable that researchers and readers wish to class particular cortisol levels or patterns as, on some level, “good” or “bad,” cortisol researchers must be mindful that apparent dysregulations of the HPA axis may have adaptive significance when considered in a developmental, social, or even situational context. For example, hyperarousal and exaggerated reactivity to stress may be beneficial in a risky or chaotic environment, just as a blunted stress response may be adaptive in the wake of unavoidable trauma. Elevated cortisol may be beneficial in situations that warrant a burst of short-term energy and deleterious in situations that do not permit an active response. Additionally, the ability to recover, or “turn off” the cortisol response, may have just as much adaptive significance as the initial stress response. Interestingly, a recent study found that neither chronic nor episodic stress predicted HPA axis outcomes, but their interaction did; for example, cortisol output was higher among participants experiencing severe episodic stressors and concurrent chronic interpersonal stress (Marin, Martin, Blackwell, Stetler, & Miller, 2007). Among participants with low chronic interpersonal stress levels, however, episodic stress was linked with lower morning and whole-day cortisol output, suggesting that participants with greater social support may show physiological resilience to acute stressors. The importance of context underscores the value of conducting cortisol research in everyday life, as participants encounter situations that are meaningful to them. However, it places an additional burden on researchers to appropriately represent the social and emotional dimensions of participants’ daily lives, both in terms of ongoing, trait-level
functioning (e.g., personality, relationship satisfaction, occupational status) and on a momentary, state-level basis (e.g., changes in mood, the quality of particular social interactions, fluctuations in everyday work demands).

Just as psychosocial context is critical to fully interpreting cortisol patterning, so too is physiological context; that is, cortisol is a single hormone released as part of a complex, multifaceted stress response. Studying cortisol in isolation from other components of this response is an inherently limited enterprise. For example, cortisol levels may look identical in participants who have different levels of glucocorticoid sensitivity, cortisol-evoked genomic activity, and glucocorticoid receptor expression. Or patterns of sympathetic nervous system activity may differ between participants who show similar HPA axis activity. As research methods and ambulatory sampling technology increases in sophistication, researchers will be able to complement both field and laboratory studies of cortisol with analyses of tissues, receptors, genes, and additional biomarkers, as well as statistical approaches that consider the circadian rhythm of cortisol across the day. Additionally, given that everyday stressors produce cortisol changes that are often quite small and sometimes undetectable, the clinical significance of these changes must be better established by more longitudinal and holistic research. While cortisol is often posited as a mediator between psychological states and long-term health outcomes, we know little about the actual effect of small alterations in HPA axis functioning on many mental and physical disorders, so more prospective studies are needed to establish the relevance of cortisol parameters to future disease risk.

In conclusion, HPA axis research is at an exciting juncture. After several decades of work, experimental cortisol research is reaching maturity. At the same time, its younger branch, formed by researchers who integrate cortisol sampling into everyday life, is heading into adolescence and establishing its own identity, in step with a larger movement within psychology towards more ecologically valid research. Despite obstacles to naturalistic cortisol research, many of which have been enumerated here, this line of inquiry has much to teach us about the interconnections between daily experiences, psychological resources, and physical health.

Acknowledgements
The author was supported by a National Science Foundation Graduate Research Fellowship and a UCLA Chancellor’s Dissertation Fellowship during the writing of this paper. She thanks Drs. Rena Repetti, Dauglas Granger, Gregory Miller, Theodore Robles, and Chris Dunkel Schetter, as well as anonymous reviewers, for valuable guidance and feedback.

References


Nicolson, N.A. (2004). *Childhood parental loss and cortisol levels in adult men* *Psychoneuroendocrinology, 29*(8), 1012–1018.


